This dissertation has been
microfilmed exactly as received

69-2269

ANDERSON, Sabra Sullivan, 1939-

A GRAPH-THEORETIC SURVEY OF THE FINITE STRUCTURES OF CLASSICAL COMBINATORIAL THEORY.

The University of Michigan, Ed.D., 1968
Mathematics

University Microfilms, Inc., Ann Arbor, Michigan

A GRAPH-THEORETIC SURVEY OF THE FINITE STRUCTURES OF CLASSICAL COMBINATORIAL THEORY

by

Sabra Sullivan Anderson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Education in
The University of Michigan
1968

Doctoral Committee:

Assistant Professor Thomas Storer, Chairman
Professor Dorwin Cartwright
Professor Philip S. Jones
Professor Roger C. Lyndon
Professor Frank Raymond

A B S T R A C T

A GRAPH-THEORETIC SURVEY OF THE FINITE STRUCTURES
OF CLASSICAL COMBINATORIAL THEORY

by

Sabra Sullivan Anderson

Chairman: Thomas Storer

Graph theory is a simple basic study of structure. Disregarding all other physical factors, graph theory concentrates on the basic members of a given situation together with their interrelationships, which may be determined by nearness, kinship, dominance, or any other relation, and studies the abstract sturcture of the configuration which results when the members of the situation are viewed as points in the plane, and the relation between pairs of them as arcs joining the points. Thus graph theory includes the definition and study of such concepts as distance, centrality, connectedness, and isomorphism together with the characterization, classification, and enumeration of graphs according to certain structural criteria which arise from the demands of a given application or from the requirements of a careful logical approach.

The versatility of graph theory is clearly indicated by the great varity of areas in which graph theory was independently invented. Euler was the first to invent graph theory when he published a solution to the Konigsberg

bridge problem in 1736. In physics Kirchhoff invented graphs in 1847 in order to study the interconnections in electrical networks without the additional physical considerations of inductances and resistances, lengths of wires, etc. In the late nineteenth century, Cayley studied an enumeration problem in chemistry by viewing the molecular structure of hydrocarbons as a special type of graph, called a tree, and more recently, social psychologists have invented directed graphs to study groups of people and certain relationships between them. Within mathematics graphs have been introduced to study partial and complete orders, lattices, Markov chains, Ramsey numbers, finite topologies, and many other topics.

The wide use of graph theory, however, has led to great variation in terminology and notation. In the first chapter a systematic presentation of various types of graphs, diferentiated by certain relational properties, provides a foundation for the other topics in the dissertation. In the remaining chpaters a wide range of applications has been collected and analyzed, and an extensive bibliography is given.

One of the most famous unsolved problems of mathematics--the Four Color Conjecture--lends itself easily to a graphical formulation, and its many equivalent forms, which include conjectures on factorability, chromatic polynomials, the existence of a binary function, Hadwiger's conjecture, and others, are described and the equivalences demonstrated within the graphical context. The existence of systems of distinct

representatives is closely related to a problem concerning the connectivity of graphs, and P. Hall's theorem on systems of distinct representatives has a surprisingly wide range of equivalent theorems. Those which have been collected and discussed in the dissertation include Menger's theorem on 1-dimensional simplexes, König's theorems on bipartite graphs and (0,1)-matrices, Ford and Fulkerson's max-flow min-cut theorem, Dulmage and Mendelsohn's matrix theorem, Dilworth's theorem on partially ordered sets, and several others.

Graphs and matrices provide each other with a variety of methods and simplifications including, in particular, a remarkably simple proof of a theorem of Amitsur and Levitski. Cayley developed a graphical representation of groups which gives insight into some of the basic group concepts and also raises new structural questions about groups. Conversely, the study of the automorphism groups of graphs is an important topic in graph theory with many questions still to be answered.

A certain type of bipartite graph occurs in many disguises; those which have been collected and studied here are Latin squares, finite projective planes, symmetric block designs, Hadamard matrices, and perfect difference sets. These topics provide each other with various existence theorems, and they share many unsolved problems of current interest.

Other graphical structures which are analyzed include Markov chains, Ramsey numbers and related topics, mathematical

games, and finite topologies. Some research of the author on the lattice topologies of digraphs is included. Several topics from nonmathematical areas, in particular, Kirchhoff's work in circuit theory, Cayley's chemical research, and studies from social psychology are described in detail, together with brief mention of several other interesting applications. In each of these areas, mathematical and nonmathematical, graph theory provides a sound logical framework for viewing and analyzing structural properties; graphical diagrams furnish clarity and graphical theorems furnish information about the systems and suggest new directions of inquiry.

ACKNOWLEDGEMENTS

For his inspiration and continued interest in directing this thesis, I wish to gratefully acknowledge the help and support of Dr. Thomas Storer.

I am also indebted to my colleagues, Dr. Gary Chartrand, Dr. Edgar M. Palmer, and Dr. Michael Plummer for their many helpful suggestions, and to Dr. Roger C. Lyndon for bringing several topics to my attention and for reading the manuscript with great care.

TABLE OF CONTENTS

		Page
ACKNOWLEDGEMENTS		ii
LIST OF TABLES		v
LIST OF ILLUSTRATIONS		vi
INTRODUCTION		1
Chapter I.	GRAPH THEORY AND FINITE RELATIONS	9
	1. Relations	9
	2. Digraphs	16
	3. Functional digraphs	24
	4. Graphs	28
	5. Trees	40
	6. Oriented graphs	45
	7. Tournaments	48
	8. Orders	52
	9. Lattices	57
II.	COLORABILITY	61
III.	MENGERIAN THEOREMS	76
IV.	GRAPHICAL METHODS IN MATRIX THEORY	94
V.	GROUPS AND GRAPHS	109
VI.	FINITE PROJECTIVE PLANES AND RELATED SYSTEMS	140
VII.	MISCELLANEOUS MATHEMATICAL APPLICATIONS	177
	1. Markov chains	177
	2. Ramsey numbers	181
	3. Number theory	187
	4. Topology	195
	5. Mathematical puzzles	204
VIII.	NONMATHEMATICAL APPLICATIONS OF GRAPH THEORY	209
	1. Circuit theory	210

2. Chemistry 214
3. Social Psychology 215
4. Miscellaneous topics . . 219

APPENDIX 229

BIBLIOGRAPHY 230

LIST OF TABLES

Table		Page
5.1.	Cycle lengths of two-color Cayley diagrams with N points	117
5.2.	Cycle lengths of plane three-color Cayley diagrams with N points . . .	123
6.1.	The number of Latin squares of small order	143
7.1.	Ramsey numbers with t=2 and r=2 . .	182
7.2.	Number of labeled topologies on n points	199

LIST OF ILLUSTRATIONS

Figure		Page
1.1.	Relagraph	10
1.2.	Relagraph with one point of each type .	11
1.3.	Isomorphic relagraphs	11
1.4.	Relagraph with cycles and acyclic relagraph	13
1.5.	Relagraph with its adjacency matrix .	13
1.6.	Relagraph, its converse and complement	16
1.7.	All digraphs with three points . .	17
1.8.	Digraphs in each of the connectedness categories	18
1.9.	Digraph with three strong and two weak components	19
1.10.	Adjacency matrix of a labeled acyclic digraph	20
1.11.	Trees to and from a point . .	21
1.12.	Digraph with three spanning trees from v_1	22
1.13.	Eulerian digraph	24
1.14.	All functional digraphs with four points	25
1.15.	Contrafunctional digraphs . .	26
1.16.	All graphs with four points . .	28
1.17.	Disconnected graph and matrix . .	30
1.18.	Graph with bridge x, cutpoint v .	31
1.19.	Nonhamiltonian and hamiltonian blocks .	33
1.20.	Eulerian and noneulerian, hamiltonian and nonhamiltonian graphs . .	33
1.21.	Graphs resulting from operations on K_n	35

Figure		Page
1.22.	Union, intersection, and symmetric difference of graphs . . .	36
1.23.	Graphs having 1-factors . . .	37
1.24.	Planar and nonplanar graphs . .	38
1.25.	Graphs homeomorphic to K_5 and $K_{3,4}$.	38
1.26.	Planar and plane graphs . .	39
1.27.	Representation of a cube as a plane graph	40
1.28.	Trees with up to six points . .	41
1.29.	Trees with one and two central points .	42
1.30.	Trees with a given number of central and centroid points . . .	43
1.31.	Graph with two independent cycles .	44
1.32.	Oriented graphs with three points .	45
1.33.	Strongly oriented graph . . .	46
1.34.	Tournaments with two, three, and four points	48
1.35.	Tournament labeled by strong components	51
1.36.	Strong tournament . . .	51
1.37.	Quasi-order	52
1.38.	Hasse diagrams and digraphs of partial orders	54
1.39.	Lattices with five points . .	58
2.1.	Bipartite graphs . . .	63
2.2.	Graph and its (Whitney) dual . .	64
2.3.	Replacing point of degree > 3 . .	66
2.4.	Removing faces of degree 4 . .	66
2.5.	Petersen graph	68
2.6.	Correspondence between a face-coloring and a line-coloring . . .	69
2.7.	Line graph of K_4 . . .	69

Figure		Page
2.8.	Assignment of a color function . .	70
2.9.	Contraction of a planar graph . .	72
2.10.	Construction of P(G,n) . . .	74
2.11.	Three-dimensional figure with chromatic number n	75
3.1.	Graphs with cut sets indicated . .	77
3.2.	Graph with minimal cut set {x,y,z} .	79
3.3.	Graph with covering set {u,v,w,x} .	81
3.4.	Network with flow value 7 . .	87
3.5.	Bipartite graph and its associated network	88
3.6.	Construction of a maximal set of independent 1's	90
4.1.	Matrix and associated relagraph .	94
4.2.	Characteristic polynomials and eigenvalues of graph	100
4.3.	Graph with polynomial $\frac{x(x+2)}{2}$ . .	102
4.4.	Directed graph with $[A_1,A_2,A_3] \neq 0$.	105
4.5.	Construction of a digraph in which q = 2p	106
4.6.	Digraph with point v with three incident arcs	106
4.7.	Construction of digraph D_i . .	107
4.8.	Construction of D_j . . .	107
5.1.	Cayley diagrams of S_3 and A_4 . .	110
5.2.	Cayley diagram of S_4 . . .	113
5.3.	Cayley diagram of C_3 . . .	114
5.4.	Intersecting cycles of different colors	115
5.5.	Plane two-color Cayley diagrams with 2j, 6, and 12 points . . .	118
5.6.	Plane two-color Cayley diagrams with 24 and 60 points . . .	119

Figure		Page
5.7.	Plane two-color Cayley diagram with N=60, i=2, j=2, k=5	120
5.8.	Plane two-color Cayley diagram with N=60, i=5, j=3, k=2	121
5.9.	Plane three-color Cayley diagrams with 4k and 24 points . . .	124
5.10.	Plane three-color Cayley diagrams with 48 and 120 points . . .	125
5.11.	Graphs with abstractly, but not permutationally, isomorphic graphs .	129
5.12.	Graphs and their groups . . .	130
5.13.	Cubic graph with points of type (3,3,3) and $(4,4,\infty)$	133
5.14.	Graph whose group is cyclic . .	134
5.15.	Graph with preassigned group . .	137
5.16.	Smallest graphs whose groups are C_3, C_4, C_5	138
6.1.	Latin squares of order 4 . .	141
6.2.	Euler squares of orders 3 and 10 .	144
6.3.	Graphs of π_2 and π_3 . .	172
6.4.	Graph and its line graph . . .	172
7.1.	Relagraph of a Markov chain . .	175
7.2.	Relagraph R_1 of a regular Markov chain and relagraph R_2 of an ergodic Markov chain	179
7.3.	Relagraph of an absorbing Markov chain .	180
7.4.	Graphs determining the Ramsey numbers N(3,3,2), N(3,4,3), and N(3,5,2) .	184
7.5.	Graphs illustrating $N'(3,3) > 5$ and $N'(4,3) > 7$	186
7.6.	Graph showing $N''(4,4) \geq 7$ . .	186
7.7.	Cycle graphs of $M_{11}, M_{15}, M_{16}, M_{55}, M_{21}$.	187
7.7.	Cycle graphs of M_5, M_8, M_{16}, M_{24} . .	193

Figure Page

7.9. Cycle graph of M_{63} 194

7.10. Digraphs of labeled topologies . . 198

7.11. Transitive digraph 198

7.12. Cliques in digraphs . . . 201

7.13. Two digraphs with the same topological lattice 202

7.14. Tournament associated with lattice L . 203

7.15. Map of Konigsberg 204

7.16. Graph of Konigsberg 205

7.17. Dodecahedron and its graph . . . 205

7.18. Tower of Hanoi puzzle . . . 206

7.19. The 3-cube Q_3 208

8.1. Graph of a network 210

8.2. Oriented graph of an electrical network with its matrices F and C . . 211

8.3. A maximal set of independent cycles for the graph in Figure 8.2. . . 213

8.4. Assignments of currents and voltages . 214

8.5. Hydrocarbon molecules . . . 214

8.6. Communication channels . . . 217

8.7. Signed digraphs 218

8.8. Teleprinter diagrams . . . 222

8.9. Kinship systems 225

8.10. Analysis and Hasse diagram of a project . 227

INTRODUCTION

Graph theory is a simple basic study of structure. Disregarding all other physical factors, graph theory concentrates on the basic members of a given situation together with their interrelationships, which may be determined by nearness, kinship, dominance, or any other relation, and studies the abstract structure of the configuration which results when the members of the situation are viewed as points in the plane, and the relation between pairs of them as arcs joining the points. Thus graph theory includes the definition and study of such concepts as distance, centrality, connectedness, and isomorphism together with the characterization, classification, and enumeration of graphs according to certain structural criteria which arise from the demands of a given application or from the requirements of a careful logical approach.

The versatility of graph theory is clearly indicated by the great variety of areas in which graph theory was independently invented. Euler was the first to invent graph theory when he published a solution to the Konigsberg bridge problem in 1736. In physics Kirchhoff invented graphs in 1847 in order to study the interconnections in electrical networks without the additional physical considerations of inductances and resistances, lengths of wires,

etc.; the graphical properties of his networks aided him in gaining knowledge about voltage and current equations. In the late nineteenth century, Cayley studied an enumeration problem in chemistry by viewing the molecular structure of hydrocarbons as a special type of graph, called a tree, and more recently, social psychologists have invented directed graphs to study groups of people and certain relationships between them. Within mathematics graphs have been introduced to study groups, Markov chains, Ramsey numbers, finite topologies, and many other topics. In each of these areas, mathematical and nonmathematical, graph theory provides a sound logical framework for viewing and analyzing structural properties; graphical diagrams furnish clarity and graphical theorems furnish information about the systems and suggest new directions of inquiry.

In this thesis we first present a systematic survey of finite relations from a graph-theoretic viewpoint. We collect and systematize many types of structures--graphs, digraphs, orders, lattices, etc.--and structural concepts which may properly be considered graph-theoretical, and indicate how other topics can be brought into this framework.[1] Starting with the general class of relagraphs, we introduce various structural properties and proceed to specialized subclasses of graphs exhibiting certain structural

---

[1]A diagram which displays the categories and subcategories which we study, ordered from the most general to the more specific, is given in Appendix I. It serves as a basic framework in which other classes of relations may be placed, giving them the benefits of all the results preceding them in the diagram.

characteristics, such as symmetry or completeness. In the course of this discussion, we develop a body of basic graphical theorems which apply to a given class and all of its subclasses. Certain specific simple graphs are studied in detail, and then techniques for building up more complicated graphs from these by various operations are introduced, together with a study of which properties are preserved under these operations. It is the purpose of Chapter I to give a systematic description of various types of graphs and relations and a collection of basic structural theorems which form the foundation of the rest of the thesis. Theorems are not proved, but many examples are given to illustrate and lend plausibility to them, and an extensive list of references is provided.

The major part of the thesis deals with the interrelationship between graph theory and other remarkably diverse mathematical topics. Many of these have, to a large extent, been studied separately in the past and are widely scattered throughout the literature; they are collected here and their similarities and, in some cases, their equivalences, are analyzed. Several unsolved problems are stated and interpreted in different contexts. In Chapter II colorability of graphs is discussed and the Four Color Conjecture is stated in its graphical form; the many other mathematical problems which are actually equivalent to the Four Color Conjecture are then introduced and discussed. Research of a more successful nature which has been inspired by the Four Color Conjecture is also mentioned.

The existence of systems of distinct representatives is closely related to a problem concerning the connectivity of graphs. In Chapter III, Menger's theorem on the connectivity of graphs is presented together with a surprisingly wide range of equivalent theorems, which include Hall's theorem on systems of distinct representatives and Ford and Fulkerson's max-flow min-cut theorem. These theorems deal with everything from graphs, matrices, and sets to lattices and network flows; and the list will undoubtedly be extended to include many other areas in the future.

Graphs and matrices are very closely related; a graph can be represented by a matrix which records adjacencies between points, and a matrix can be represented by a digraph whose arcs are assigned the numbers of the matrix. In Chapter IV graph theory is applied to matrix computations and theorems to yield, in many cases, a considerable simplification. In particular, we show how to simplify the process of finding determinants, inverses, and eigenvalues, how to solve graphically a matrix equation, and how to prove elegantly an abstract matrix theorem.

Graphs and groups are also closely associated; in Chapter V we show how each plays a significant role in the study of the other. A graphical representation of a group, first introduced by Cayley, provides insight into some of the basic group concepts, such as normality and commutativity, and also raises a new structural question about groups; namely, when is the graph of a group planar? The study of the automorphism groups of graphs is an important topic in

graph theory, especially since it is a central problem in the enumeration of classes of graphs. We present the basic theorems on the groups of graphs and discuss in detail a question (still unsolved in general) posed by König: For which groups Γ are there graphs whose automorphism group is isomorphic to Γ?

The closely related systems of Latin squares, finite projective planes, symmetric block designs, Hadamard matrices, and perfect difference sets can also be viewed as bipartite graphs. In Chapter VI these systems and their interrelationships are described in detail. Several unsolved problems common to them all are discussed, together with a famous conjecture made by Euler which was not completely resolved until 1959 (and then only with the aid of computers).

A thorough discussion of all the mathematical topics in which graph theory plays a role is impossible in any work of reasonable length. In Chapter VII, several other topics are described more briefly. First a graphical description of Markov chains and some of their basic properties is given; it is hoped that this description may some day prove to be of more than simple representational value to probabilists. Secondly, Ramsey numbers are presented in a graphical context together with some recent results on determining the exact values of some of these numbers; this research has been very successful in the graphical framework. A little-known application of graph theory to

number theory occurs in the representation of the groups of numbers less than and relat. ely prime to a given number m; a complete characterization of these groups is described. There is a close connection between graph theory and topologies on finite sets; in fact, graph theory has been used to compute the number of labeled topologies on sets with at most 7 points. A lattice may be associated with a topology on a finite set and some research of the author's on the graphical properties of these lattices is presented. Not insignificant in graph theory is the role which mathematical puzzles have played in inspiring and stimulating research. Several mathematical puzzles together with their solutions or nonsolutions complete Chapter VII.

The nonmathematical applications of graph theory cover a tremendous range of interests. In order to illustrate the important function which graph theory performs in representation and structural analysis in this area, we have selected, somewhat arbitrarily, topics from circuit theory, chemistry, social psychology, sociology, economics, critical path analysis, information theory, and bell-ringing. These topics demonstrate how graph theory can be applied to any situation which can be viewed as a set of points related to each other in some binary fashion, and how graph theory is remarkably adaptable in accommodating certain special aspects of a situation through the device of weighting, signing, or coloring the arcs of a relagraph. In some applications the major function of graph theory is to provide a representation; in other areas, the full power of graphical

theorems is utilized.

The emphasis throughout this thesis lies on the role that graph theory plays in structural analysis in general. Our purpose is to provide first a background and foundation for graphical study by collecting, organizing, and illustrating the many structures and concepts which belong under the heading of graph theory. Secondly, we attempt to provide an outlook and an approach to graphical research by presenting basic graphical techniques, by discussing an extensive collection of important and widely applicable theorems, by discussing general graphical classifications and decompositions, and by applying all of these to many examples. Thirdly, we discuss in detail many widely diverse topics, mathematical and nonmathematical, from a graphical viewpoint; we describe the structural categories in which these topics belong, indicate how some of their basic theorems are graphical theorems in disguise, and raise new questions of a graphical nature. It is our hope that this may give direction, motivation, and insight into future research both in graph theory and in its applications.

While the treatment in this thesis is not historical, the historical origins of graph theory are mentioned and original sources are cited where possible. An extensive bibliography, given by chapters, also lists the most accessible sources.

Prerequisite to the reading of this thesis is a rudimentary knowledge of matrix theory. Matrices are used extensively in computation and classification, and we

assume that the reader can multiply two matrices together and can find determinants, inverses, and eigenvalues.

I. GRAPH THEORY AND FINITE RELATIONS

We begin with a graph-theoretic discussion of finite relations and relational systems. The purpose of this chapter is to provide an introduction to graphical notation and terminology, an analysis of the transformation of any given collection of relational properties into a graphical configuration, and a presentation of the concepts and theorems basic to the discussion in later chapters. The graphs and relations which we consider are defined only on finite sets, although most of the concepts and theorems can easily be generalized to the infinite case.

1. Relations

A relation R on a finite nonempty set S is a collection of ordered pairs of elements of S, that is, a subset of the cartesian product S x S. We may also interpret R to be a configuration in the plane by letting the elements of S be points $s_1, s_2, \ldots, s_n$ and by drawing a directed line from s_i to s_j whenever (s_i, s_j) is in R. The relation $R = \{(s_1,s_2), (s_2,s_1), (s_2,s_3), (s_1,s_3), (s_4,s_3), (s_4,s_4)\}$ on the set $S = \{s_1,s_2,s_3,s_4,s_5\}$ is shown in Figure 1. The word "graph" is often used as a general term for such a representation, but to avoid ambiguity with a more restricted definition of "graph", we call this pictorial representation of an arbitrary relation a relagraph. We use the terms

"relation" and "relagraph" interchangeably.

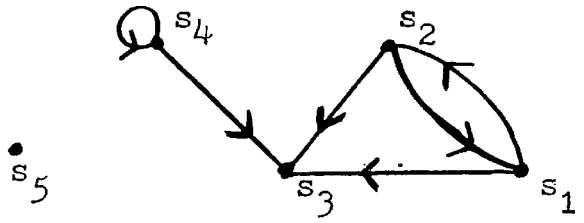


Figure 1. Relagraph

Conforming with graph-theoretic notation and terminology, we now denote the finite set by V and the set of ordered pairs by X. The elements v_i of V are called <u>points</u> while the elements $x_i = (v_i, v_j)$ of X are called <u>arcs</u>.[1] If arc (u,v) is in X, then u is said to be <u>adjacent to</u> v and v is <u>adjacent from</u> u. The <u>outdegree</u> od(v) of a point v is the number of arcs in X whose initial point is v, the <u>indegree</u> id(v) is the number of arcs whose terminal point is v. Since each arc contributes 1 to the sum of the outdegrees at its initial point and 1 to the sum of the indegrees at its terminal point, we have the following basic theorem.

<u>Theorem 1</u>. If p is the number of points in a relagraph and q the number of arcs,

$$\sum_{i=1}^{p} id(v_i) = \sum_{i=1}^{p} od(v_i) = q \quad .$$

It is convenient to classify every point according to its indegree and outdegree. A <u>transmitter</u> is a point with positive outdegree and zero indegree, while a <u>receiver</u> has

---

[1]The reader should by no means infer that there is widespread uniformity of terminology in graph theory. For example, a "point" is also called a "vertex" or "node" by a great many graph theorists, and still others call it a "dot", "cell", "0-simplex", or "junction".

zero outdegree and positive indegree. A carrier is a point whose outdegree and indegree are both one, and an isolate is a point with zero outdegree and indegree. Any other point is called an ordinary point. Figure 2 shows a relagraph with exactly one point of each of these five types.

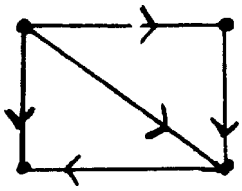


Figure 2. Relagraph with one point of each type

Two relagraphs R_1 and R_2 with point sets V_1 and V_2 are isomorphic, written $R_1 \cong R_2$, if there exists a one-to-one correspondence between V_1 and V_2 which preserves adjacency. A pair of isomorphic relagraphs, with their points labeled to indicate the correspondence, is shown in Figure 3. Isomorphism of relagraphs is an equivalence

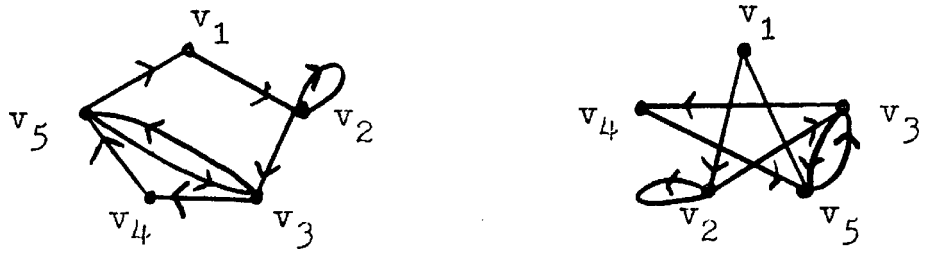


Figure 3. Isomorphic relagraphs

relation, and henceforth we consider all relagraphs belonging to a given equivalence class to be the same. The problem of determining when two relagraphs are isomorphic is more easily discussed in terms of a second representation of relations.

Given a relation R on a set V of p points, we define

the adjacency matrix A(R) to be the p x p matrix whose (i,j)-entry is 1 if (v_i,v_j) is in R and 0 otherwise. The adjacency matrix for the relagraphs in Figure 3 is

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

It is clear that two relagraphs are isomorphic if and only if their points can be labeled in such a way that their adjacency matrices are identical.

We now introduce some basic concepts of graph theory whose properties are closely associated with various functions of the adjacency matrix. A walk in a relagraph is a nonempty sequence of points and arcs, beginning and ending with points, such that each arc is directed from the point preceding it to the point following it. A walk may be denoted by its point sequence, e.g., $v_1v_2\ldots v_n$, the arcs being evident. A walk in which no point occurs more than once is called a path. The point v is said to be reachable from the point u if there is a path from u to v; in particular, every point v is reachable from itself by the trivial path consisting of the single point v. A walk is closed if its initial and terminal points coincide. A cycle is a closed walk $v_1v_2\ldots v_nv_1$ of length at least 1 in which the n points are distinct; a cycle of length 1 is called a loop. If a relagraph has no cycles, it is acyclic. In Figure 4, the relagraph R_1 contains cycles of lengths 1, 2, and 3, whereas R_2 is acyclic.

The distance d(u,v) from u to v is the length of a

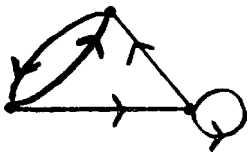
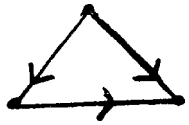
R_1:  R_2: 

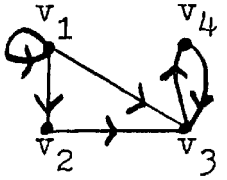
Figure 4. Relagraph with cycles and acyclic relagraph

shortest walk from u to v if such a walk exists; otherwise we define $d(u,v) = \infty$. We remark that d satisfies the ubiquitous "triangle inequality",

$$d(u,w) \leq d(u,v) + d(v,w) \quad ,$$

whenever there is a walk from u to v and a walk from v to w.

In the relagraph of Figure 5 there are two walks of length 2 from v_1 to v_3, one which includes the loop and one which does not. The usefulness of the adjacency matrix becomes apparent in the following algorithm which enables us to count the number of walks of a given length from one point to another.

R: 

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

$$A^2 = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 5. Relagraph with its adjacency matrix

Theorem 2. If A is the adjacency matrix of a relagraph R, the (i,j)-entry in A^n is the number of walks of length n from v_i to v_j in R.

There are two other matrices[2] associated with a

---

[2]We remark here that by a matrix we mean any ordered array of elements. These elements need not come from a field.

relagraph which, even though they do not completely characterize the relagraph, are useful in studying particular properties. The reachability matrix N(R) has $n_{ij} = 1$ if v_j is reachable from v_i and $n_{ij} = 0$ otherwise. In the distance matrix B(R), the entry b_{ij} is the distance from v_i to v_j. The matrices associated with the relagraph R in Figure 5 are

$$N(R) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \text{ and } B(R) = \begin{pmatrix} 0 & 1 & 1 & 2 \\ \infty & 0 & 1 & 1 \\ \infty & \infty & 0 & 1 \\ \infty & \infty & 1 & 0 \end{pmatrix}.$$

Both N(R) and B(R) can be obtained from A(R) by applying Theorem 2.

Corollary 2a. The entries of N(R) and B(R) can be obtained from the powers of A(R) as follows: Letting $a_{ij}^{(n)}$ denote the (i,j)-entry of A^n,

$n_{ij} = 1$ for $i \neq j$ if and only if $a_{ij}^{(n)} > 0$ for some n.

$n_{ii} = 1$ for all i.

b_{ij} for $i \neq j$ is the smallest n such that $a_{ij}^{(n)} > 0$, if any, and ∞ otherwise.

$b_{ii} = 0$ for all i.

Many of the definitions we have given occur in pairs, such as indegree and outdegree, adjacent to and adjacent from, transmitter and receiver. Two such directionally opposite ideas are called converse concepts. The converse R^c of a relagraph R is the relagraph having the same point set as R and having the arc (u,v) whenever (v,u) is in R. Thus R^c is obtained by reversing the directions of all the arcs of R, as shown in Figure 6, and $A(R^c)$ is precisely the transpose of A(R). A relagraph is self-converse if $R^c \cong R$.

Clearly every cycle is self-converse, as is every symmetric relagraph. The concepts of carrier, isolate, and ordinary point are all self-converse. Converse concepts are related by a powerful principle which provides a set of instant corollaries for subsequent theorems.

Directional Duality Principle. For each theorem on relagraphs, there is a corresponding theorem obtained by replacing every concept by its converse concept.

We use Theorem 3 below, and its directional dual, Theorem 3', to illustrate how this principle generates new theorems.

Theorem 3. An acyclic relagraph has at least one point of outdegree 0.

The Directional Duality Principle immediately implies:

Theorem 3'. An acyclic relagraph has at least one point of indegree 0.

There is a second type of duality which we define here for completeness. The complement $\overline{R}$ of a relagraph R is the relagraph having the same set of points as R and containing the arc (u,v) whenever (u,v) is not in R. The adjacency matrix for $\overline{R}$ is clearly seen to be J - A(R), where J is the matrix whose entries are all 1's. A relagraph is self-complementary if $\overline{R} \cong R$. A relagraph, its converse, and its complement are shown in Figure 6. The Boolean Duality Principle, analogous to the Directional Duality Principle relates complementary concepts. The dearth of such concepts in graph theory, however, makes it of little use here, and

consequently future reference to the dual concept will mean the directional dual.

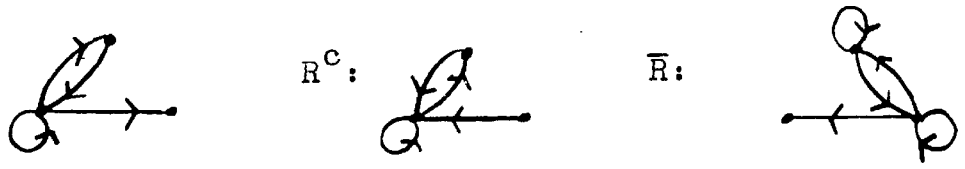


Figure 6. Relagraph, its converse and complement

2. Digraphs[3]

A relation R on the set S is reflexive if (s,s) is in R for every s in S; it is irreflexive if it contains none of these pairs. There is no significant difference in structure between reflexive and irreflexive relations; reflexive relations have a loop at every point and irreflexive relations have no loops at all. For all concepts involving a relationship between points, the presence or absence of loops is immaterial. Therefore, we now deal primarily with irreflexive relations.

A relagraph corresponding to an irreflexive relation is a directed graph, or, more briefly, a digraph. All the digraphs (up to isomorphism) with three points are shown in Figure 7. With one exception, all of the concepts and theorems pertaining to relagraphs which we have mentioned above apply immediately to digraphs. The complement $\overline{D}$ of a digraph D is defined to be the digraph having the same point set as D and having the arc (u,v) where $u \neq v$ whenever (u,v) is not in D. Thus loops are eliminated from the complement

---

[3]Many of the ideas discussed in this section may be found in greater detail in F Harary, R. Norman, and D. Cartwright's Structural Models, Wiley, New York, 1965.

also.

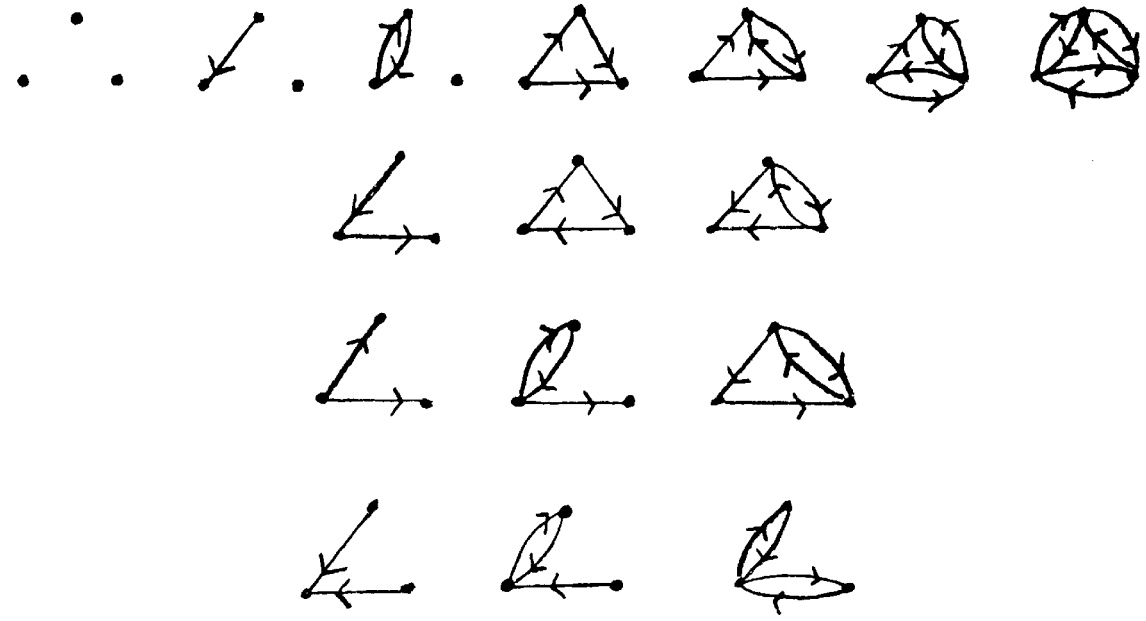


Figure 7. All digraphs with three points

A primary structural aspect of a digraph is its degree of connectedness. To define the various categories of connectedness, it is necessary to introduce the concept of an undirected walk. A semiwalk joining u and v is a nonempty alternating sequence of points and arcs, $v_1, x_1, v_2, x_2, \ldots, v_n$, where x_i is either (v_i, v_{i+1}) or (v_{i+1}, v_i); a semipath and a semicycle are defined analogously. A walk or a semiwalk in a digraph D is called spanning if it contains all the points of D.

There are four ways in which two points may be related to each other in a digraph; there may be a path from each to the other, from one to the other, a semipath between them, or none of these. Viewing the digraph as a whole, we say that a digraph D is strongly connected, or strong, if every two points are mutually reachable; D is unilaterally

connected, or unilateral, if for any two points, at least one is reachable from the other. A digraph is weakly connected, or weak, if every two points are joined by a semiwalk. A digraph is disconnected if it is not weak, and a digraph with no arcs at all is called totally disconnected. Clearly every strong digraph is unilateral and every unilateral digraph is weak. We say that a digraph is strictly unilateral if it is unilateral but not strong, strictly weak if it is weak but not unilateral. In Figure 8 we show examples of graphs in each of the connectedness categories. These examples suggest the characterization given in Theorem 4 of the connectedness types.

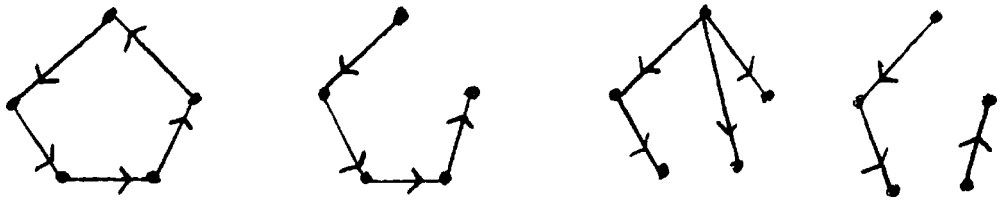


Figure 8. Digraphs in each of the connectedness categories

Theorem 4. A digraph D is strong if and only if D has a spanning closed walk; D is unilateral if and only if D has a spanning walk; and D is weak if and only if D has a spanning semiwalk.

The connectedness of the complement of a digraph can sometimes be inferred from knowledge of the digraph itself. For example, the complement of a disconnected graph is always strong. On the other hand, the complement of a strong digraph may be in any of the connectedness categories. The

complement of a strictly weak digraph is always weak, but not necessarily strictly weak.

If a digraph is not strong, we may wish to study a portion of the digraph in which all points are mutually reachable. A subgraph of a digraph D is a digraph whose points and arcs are points and arcs of D. The subgraph of D induced by a subset V_1 of the point set V has V_1 as its point set and contains all the arcs of D which join two points of V_1. A strong component of a digraph is a maximal strong digraph. The strong components partition the points of a digraph, but not the arcs; in fact, an arc is in a strong component if and only if it belongs to some cycle. The digraph in Figure 9 has three strong components, namely, the cycle of length 2, the cycle of length 3, and the single point v_4. The weak components of a digraph D, that is, the maximal weak subgraphs, partition both the points and the arcs of D. The digraph in Figure 9 has as its two weak components the subgraph induced by the points v_1, v_2, v_3, v_4, and the subgraph on the points v_5 and v_6.

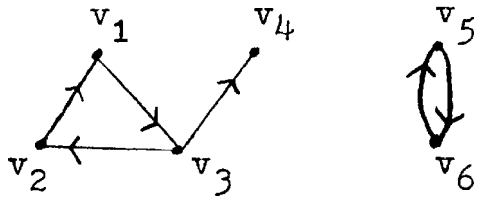


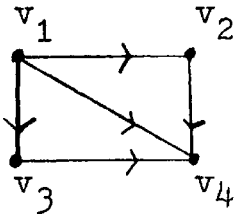
Figure 9. Digraph with three strong and two weak components

A particular type of digraph which plays an important role in the applications of graph theory is the acyclic digraph, which we already know from Theorem 3 has at least

one point with indegree 0 and one point with outdegree 0. we give several characterizations of acyclic digraphs and illustrate them with the acyclic digraph in Figure 10.

Theorem 5. The following statements are equivalent for a weak digraph D.

1) D is acyclic.
2) Every strong component of D consists of a single point.
3) Every walk of D is a path.
4) The points of D may be labeled in such a way that the adjacency matrix A(D) is upper triangular.

D: 

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

Figure 10. Adjacency matrix of a labeled acyclic digraph

An acyclic digraph of special interest is one which has no semicycles. A tree from a point v is a weak digraph with no semicycles in which every point is reachable from v; the point v is called the source of the digraph. A sink, dually, is a point which is reachable from all points. A tree to a point v is a weak digraph with no semicycles having v as a sink. In Figure 11, D is a tree from the point v and the converse of D is a tree to v. In the following theorem we indicate how a tree from a point may be characterized by the indegrees of its points.

Theorem 6. A weak digraph is a tree from the point v if and only if id(v) = 0 and every other point has indegree 1.



Figure 11. Trees to and from a point v

Taking the converse of a tree from a point, we interchange indegrees and outdegrees and obtain a tree to a point. Theorem 6 in its dual form characterizes these digraphs.

Theorem 6'. A weak digraph is a tree to the point v if and only if od(v) = 0 and every other point has outdegree 1.

The number of spanning trees of a digraph D from one of its points can be calculated by a few simple matrix operations from the adjacency matrix A(D) and the indegree matrix id(D), which has the indegrees of the points of D on its diagonal. Defining the matrix $M_{id}(D) = id(D) - A(D)$, we can now state the matrix tree theorem for digraphs, first proved by Tutte.[4]

Theorem 7. For any labeled digraph D, the value of the cofactor of any element in the j'th column of M_{id} is the number of spanning trees from v_j.

Applying Theorem 7 to the digraph in Figure 12, we

---

[4]This theorem was first proved in R.L.Books, C.A.B. Smith, A.H.Stone, and W.T.Tutte, The dissection of rectangles into squares, Duke Math. J. 7(1940), 312-240. It is implicit in G. Kirchhoff, Uber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung Galvanischer Ströme geführt wird, Annalen der Physik 70(1847), 497-508.

obtain

$$id(D) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad A(D) = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \quad M_{id}(D) = \begin{pmatrix} 1 & -1 & -1 & -1 \\ 0 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

The cofactor of each element in the first column is 3, and the cofactors of the elements in the remaining three columns are 0, 0, and 3, respectively. We note that the digraph does, indeed, have three spanning trees from v_1 and v_4, and none from v_2 or v_3. To compute the number of spanning trees to each point, we define od(D) to be the matrix with the outdegrees of the points on its diagonal, $M_{od}(D)$ to be od(D) - A(D), and calculate the cofactors of elements in the rows of M_{od}. The digraph D in Figure 12 has two spanning trees to each of v_2 and v_3 and none to v_1 or v_4, and

$$M_{od}(D) = \begin{pmatrix} 3 & -1 & -1 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}$$

has 0, 2, 2, 0 as the cofactors of the elements in its rows. Theorem 7' gives this dual calculation.

Theorem 7'. In the matrix M_{od}, each cofactor of the i'th row equals the number of spanning subtrees of D to v_i.

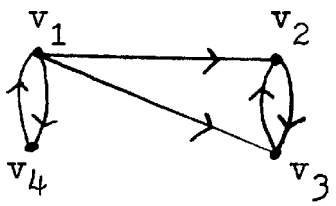


Figure 12. Digraph with three spanning trees from v_1

Another type of digraph of special interest is one

containing a walk which spans the arcs of D. A digraph is eulerian[5] if it has a closed spanning walk (called an eulerian trail) in which each arc of D occurs exactly once; clearly every cycle is eulerian. Theorem 8 below characterizes eulerian digraphs in terms of their degrees or cycles.

Theorem 8. The following statements are equivalent for a weak digraph D with more than one point.

1) D is eulerian.

2) For every point of D, $id(v) = od(v)$.

3) The set of arcs of D can be partitioned into cycles.

Matrices may also be used to calculate the number of eulerian trails of an eulerian digraph, as indicated in the next theorem, called the BEST theorem after the names of deBruijn and Ehrenfest[6], Smith and Tutte[7], the first two and last two of whom proved the theorem independently. Since $id(v) = od(v)$ for every point v in an eulerian digraph D, the matrices $M_{id}(D)$ and $M_{od}(D)$ are equal. Therefore, for all i and j, the value of the cofactor of the i'th row equals the value of the cofactor of the j'th column, since both are the cofactor of the entry m_{ij}. Hence the values of all the cofactors of M_{od} are equal, and this common value is called Δ.

---

[5]Euler originate the study of graph theory by solving a problem involving graphs of this type; see p. 203.

[6]T. van Aardenne-Ehrenfest and N.G.deBruijn, Circuits and trees in oriented graphs, Simon Stevin 28(1959), 203-217.

[7]C.A.B.Smith and W.T.Tutte, On unicursal paths in a network of degree 4, Amer. Math. Monthly 48(1941), 233-237.

Theorem 9. In an eulerian digraph, the number of eulerian trails is

$$\Delta \cdot \prod_{i=1}^{p} (d_i - 1)!$$

where $d_i = id(v_i) = od(v_i)$.

Applying this theorem to the digraph of Figure 13, we have

$$M_{od}(D) = \begin{pmatrix} 2 & -1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 0 \\ -1 & 0 & 2 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & -1 & 2 \end{pmatrix}, \quad \Delta = 3,$$

and the eulerian trails denoted by their arcs are

$x_4x_5x_8x_7x_3x_1x_2x_6$, $x_4x_1x_2x_3x_5x_8x_7x_6$, and $x_4x_1x_2x_6x_8x_7x_3x_5$.

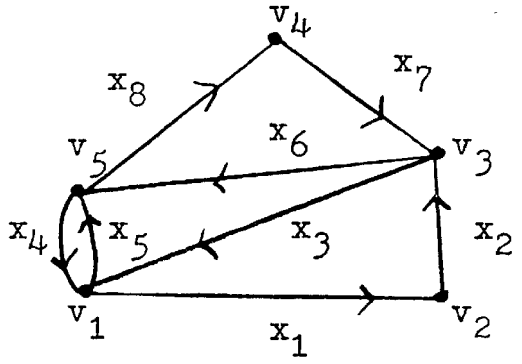


Figure 13. Eulerian digraph

3. Functional digraphs

We devote this section to a graphical discussion of one of the most widely studied relations, the function. By a function on a set S we mean a relation $R \subset S \times S$ which contains precisely one ordered pair (x,y) for each x in S. Since we are interested in the structural properties of such relations, we again assume irreflexivity, that is, we

consider functional digraphs. Formally, a functional digraph is a digraph in which every point has outdegree 1. The six functional digraphs on four points are shown in Figure 14. We note that with the exception of the disconnected one, each of these has exactly one cycle, and the arcs which are not in the cycle are all directed towards points in the cycle. These two observations lead to the structural characterizations of functional digraphs given in the next theorem.[8]

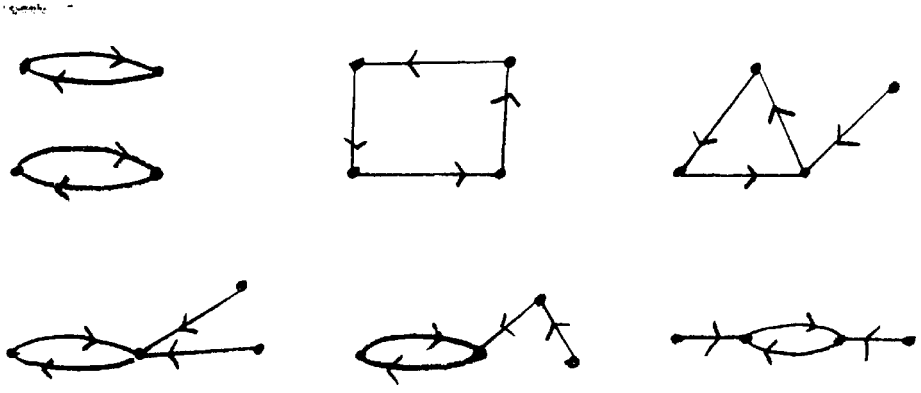


Figure 14. All functional digraphs with four points

Theorem 10. The following statements are equivalent for a weak digraph D.

1) D is functional.
2) D contains an arc x = (u,v), lying on a cycle, such that D - x is a tree to u.
3) D has exactly one cycle Z. Let D' be the subgraph of D obtained by the deletion of the arcs of Z. Then each weak component of D' is a tree to a point of Z.

---

[8] F.Harary, The number of functional digraphs, Math. Ann. 138(1959), 203-210.

Clearly a disconnected digraph is functional if and only if each of its weak components satisfies the conditions of this theorem.

Since a weak functional digraph can be viewed as a cycle together with trees to the points of the cycle, we can identify its connectedness type at a glance.

Theorem 11. A functional digraph D is strong if and only if it is a cycle; D is strictly unilateral if and only if it consists of a cycle and a nontrivial path to one point of that cycle; D is strictly weak if and only if it contains exactly one cycle and has at least two transmitters.

This theorem says that in a strong digraph every point is a carrier. A strictly unilateral functional digraph has exactly one transmitter and one ordinary point, the rest being carriers. Referring to Figure 14, we observe that there are exactly one strong, two strictly unilateral, two strictly weak, and one disconnected functional digraphs with four points.

In the dual of a functional digraph, called a contrafunctional digraph, every point has indegree one. In the examples in Figure 15, we see that contrafunctional digraphs resemble functional digraphs except that the arcs which are



Figure 15. Contrafunctional digraphs

not on the cycle are directed away from it. The dual of Theorem 10 verifies this.

Theorem 10'. The following statements are equivalent for a weak digraph D.

1) D is contrafunctional.

2) D contains an arc $x = (u,v)$, lying on a cycle, such that $D - x$ is a tree from v.

3) D has exactly one cycle Z. Let D' be the subgraph of D obtained by deletion of the arcs of Z. Then each weak component of D' is a tree to a point of Z.

Similarly the criteria for contrafunctional digraphs to belong to certain connectedness categories differ from those for functional digraphs only in the directions of the noncyclic paths.

Theorem 11'. A contrafunctional digraph D is strong if and only if it is a cycle; D is strictly unilateral if and only if it consists of a cycle and a path from one point of that cycle; D is strictly weak if and only if it contains exactly one cycle and has at least two receivers.

The converse D^c of a functional digraph D plays an important role in the study of functions. If D^c is itself a function, then it is called the inverse of D. It is clear that this occurs only when D is also contrafunctional. In a digraph which is both functional and contrafunctional, every point must be a carrier, and therefore every weak component must be a cycle. We summarize this in Theorem 12.

Theorem 12. The converse D^c of a functional digraph D is also functional if and only if every weak component of D is a cycle.

4. Graphs

A relation R on a set S is symmetric, if, for elements s and t in S, (s,t) is in R, whenever (t,s) is in R. In an asymmetric[9] relation at most one of the pairs (s,t) and (t,s) can occur. Turning now from functional digraphs to symmetric digraphs, we get another interesting class of relations. It is this type of relation on which graph theorists have largely concentrated their attention. Its basic structure is simple and provides a versatile model for applications.

A irreflexive symmetric relation is called a graph. Since the arcs of a symmetric digraph occur in directionally opposite pairs, we draw a graph using a single undirected line x between the points u and v whenever (u,v) and (v,u) are in G. All of the graphs with four points are shown in Figure 16.

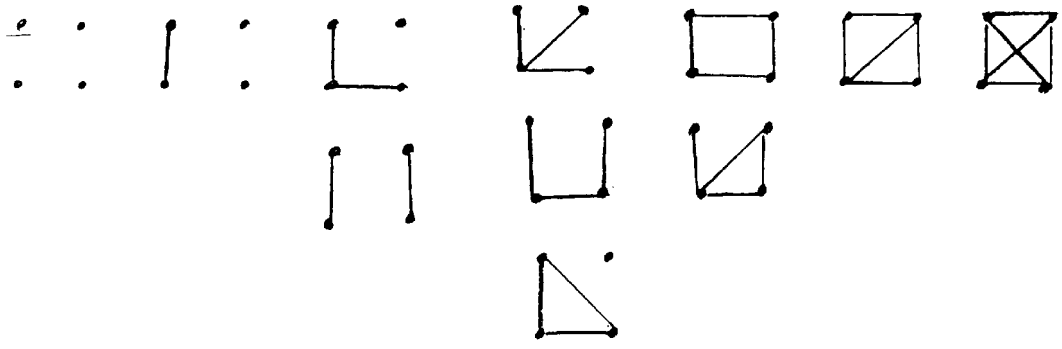


Figure 16. All graphs with four points

[9]The term "antisymmetric" is also in common usage.

Some new terminology for graphs is necessary to reflect the use of undirected lines in place of arcs. If the line $x = (u,v)$ is in the graph G, the points u and v are called adjacent and the line x is said to be incident with u and v. The degree of a point u, denoted deg(u), is the number of lines incident with u, and as a corollary to Theorem 1 we have that

$$\sum_{i=1}^{p} \deg(v_i) = 2q .$$

This leads us to the somewhat unexpected result that the number of points of odd degree in a graph is always even. A walk in a graph is a nonempty alternating sequence of points and lines, beginning and ending with points, in which each line is preceded by one of its endpoints and followed by the other. Paths and cycles in graphs are defined as in digraphs with the word "line" replacing the word "arc"; a path between u and v is often called a u-v path. By definition, a graph contains at most one line between any pair of points and no loops, and hence contains no cycles of lengths 1 or 2.

Since the graph G is irreflexive, the adjacency matrix A(G) has zeros on the diagonal; since G is symmetric, the i'th row and i'th column are identical, and their common sum is $\deg(v_i)$. Clearly if A is the adjacency matrix of a graph G, Theorem 2 implies that the number of walks of length n from v_i to v_j is the (i,j)-entry of A^n.

In a graph there is no difference between a walk and a semiwalk; consequently, strong, unilateral, and weak are

all equivalent. We say that a graph is connected if there is a path between every pair of points. A connected graph is characterized in terms of partitions of the point set V in Theorem 13.

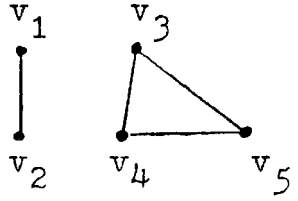
Theorem 13. A graph G is connected if and only if for every partition $V = V_1 \cup V_2$, there is a line joining some point of V_1 with some point of V_2.

This theorem implies that in a disconnected graph there is some partition of the points $V = V_1 \cup V_2$ such that no line of G joins a point of V_1 with a point of V_2. If the points of G are labeled so that $v_1, v_2, \ldots, v_k$ are in V_1 and $v_{k+1}, v_{k+2}, \ldots, v_p$ are in V_2, then the adjacency matrix can be written as

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

where A_i is the adjacency matrix of the subgraph induced by V_i. An example is shown in Figure 17. The components of a disconnected graph G are the maximal connected subgraphs, and they define a natural partition of the lines and points of G into k subgraphs $G_1, G_2, \ldots, G_k$, where k is the number of components of G. Labeling the points by subgraphs, the adjacency matrix of G has $A(G_i)$ on its diagonal and 0's elsewhere.

A:



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

Figure 17. Disconnected graph and matrix

In a connected graph G there are certain subgraphs, lines, and points which are closely related to the connectivity properties of G. A cutpoint of a connected graph G is a point whose removal, together with all lines incident with it, results in a disconnected graph; similarly, a bridge is a line whose removal disconnects G. For example, in the graph G of Figure 18, the point v is a cutpoint and the line x is a bridge. Cutpoints can be characterized as follows.

Theorem 14. The following statements are equivalent for a connected graph G with more than two points.

1) The point v is a cutpoint.
2) Points u and w exist such that v is on every u-w path.
3) There is a partition of V - {v} into subsets U and W such that for any u in U and w in W, the point v is on every u-w path.

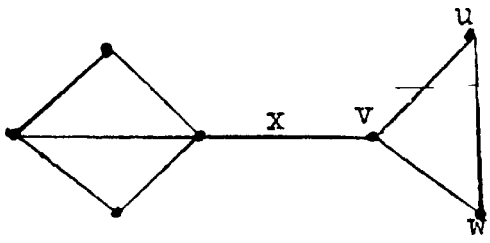


Figure 18. Graph with bridge x, cutpoint v

A nontrivial connected graph which has no cutpoints is called a block; a block of a graph G is a maximal connected subgraph of G containing no cutpoints (of itself). In Figure 18, the cycle containing u, v, and w, is a block of G. The concept of a block is very useful since in many

cases the proof of a theorem involving the arbitrary graph G can be reduced to the case in which G is a block. Hence the two characterizations of a block given in the next theorem are very important.

Theorem 15. The following statements are equivalent for a connected graph G with more than two points.

1) G is a block.

2) Every pair of points of G lie on a common cycle.

3) Every pair of lines of G lie on a common cycle.

There are two types of graphs, important both in the history of graph theory and in modern applications, which are always blocks. A graph G which has a closed spanning walk which contains every line of G exactly once is called eulerian. Two characterizations of eulerian graphs, corresponding to those for eulerian digraphs are now given.

Theorem 16. The following statements are equivalent for a connected graph G with more than two points.

1) G is eulerian.

2) Every point of G has even degree.

3) The set of lines of G can be partitioned into cycles.

A hamiltonian graph is one which has a spanning cycle, that is, a closed walk containing each point exactly once.[10] Theorems 17 and 18 give a sufficient condition[11] and a necessary condition, respectively, for a graph to be hamiltonian.

---

[10]The study of hamiltonian graphs began with a game invented by the mathematician, Sir William Hamilton, see p. 205.

[11]See O. Ore, Theory of Graphs, Amer. Math. Soc. Colloq. Publs., 38, 1962, p. 55.

A single theorem giving useful conditions which are simultaneously necessary and sufficient has yet to be found. The graph G_2 in Figure 19 is hamiltonian without satisfying the hypothesis of Theorem 17, while G_1 is a nonhamiltonian block.



Figure 19. Nonhamiltonian and hamiltonian blocks

Theorem 17. If G is a graph with $p \geq 3$, and if for all nonadjacent points u and v in G, $\deg(u) + \deg(v) \geq p$, then G is hamiltonian.

Theorem 18. If G is hamiltonian, then G is a block.

It may seem that eulerian and hamiltonian graphs are very closely related, and that information about one could perhaps provide information about the other. Actually the two concepts are completely independent, and this is illustrated in Figure 20 by four graphs each having eight points and twelve lines but each satisfying a different combination of the properties of eulerian and noneulerian, hamiltonian and nonhamiltonian.

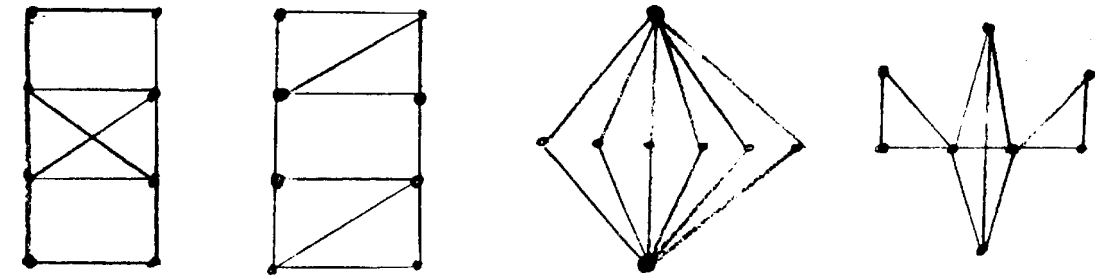


Figure 20. Eulerian and noneulerian, hamiltonian and nonhamiltonian graphs

Several basic types of graphs have been singled out by graph theorists for special study. One of these is the complete graph on p points, denoted by K_p, in which every two points are adjacent; its complement $\overline{K}_p$ is, of course, the totally disconnected graph on p points. A second type of graph is the cycle with p points, denoted by C_p. Other special graphs are most efficiently described by performing certain operations on complete graphs and cycles. The union $G_1 \cup G_2$ of two disjoint graphs G_1 and G_2 having point sets V_1 and V_2 and line sets X_1 and X_2, respectively, is the graph with the set of points $V = V_1 \cup V_2$ and the line set $X = X_1 \cup X_2$; the union of n disjoint copies of G is denoted by nG. We remark that a disconnected graph can be written as the union of its components. The join $G_1 + G_2$ is the graph on $V = V_1 \cup V_2$ having the line set $X = X_1 \cup X_2 \cup \{(u,v) \mid u \in V_1, v \in V_2\}$. In particular, the complete bipartite graph[12] $K_{m,n}$ is defined as $\overline{K}_m + \overline{K}_n$. The cartesian product of G_1 and G_2 is the graph $G_1 \times G_2$ whose set of points is $V_1 \times V_2$ where adjacency is defined by (u_1,u_2) adj (v_1,v_2) for u_1,v_1 in V_1 and u_2,v_2 in V_2 if either $u_1 = v_1$ and u_2 adj v_2 in G_2 or u_1 adj v_1 in G_1 and $u_2 = v_2$. The cycle C_4 is the cartesian product of K_2 with itself. The composition $G_1[G_2]$ also has the point set $V_1 \times V_2$, but adjacency is defined by (u_1,u_2) adj (v_1,v_2) if

[12]This type of graph is also called "bicolored", the name deriving from the fact that m of its points can be colored in one color and the remaining m points with a second color in such a way that no two points of the same color are adjacent and every pair of points of different colors are joined.

u_1 adj v_1 or $u_1 = v_1$ and u_2 adj v_2. Thus K_4 may be expressed as $K_2[K_2]$. Several examples of operations on graphs are shown in Figure 21.

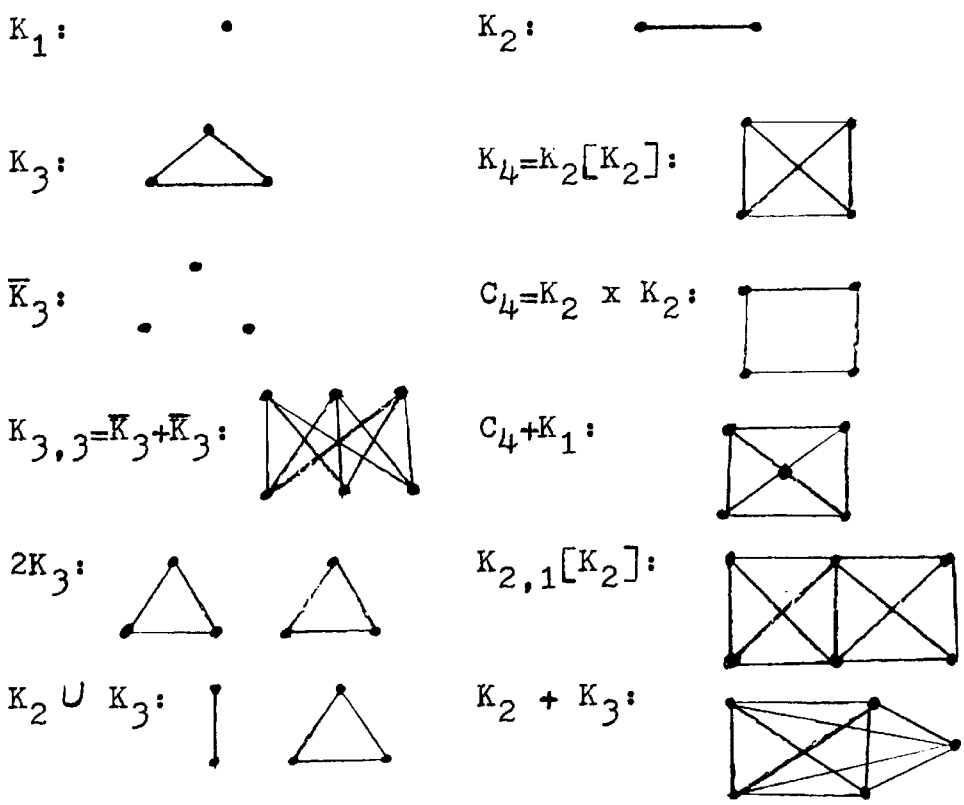


Figure 21. Graphs resulting from operations on K_n

The adjacency matrices for these graphs are easily found. Clearly $A(K_p)$ is the p x p matrix J - I, where I is the matrix with 1's on its diagonal and 0's elsewhere, since every point is adjacent to every other point, but not to itself, and $A(\overline{K}_p)$ is the p x p matrix of zeros. Letting A_1 and A_2 be the adjacency matrices for G_1 and G_2, $A(G_1 \cup G_2)$ is the p_1 x p_2 matrix

$$\begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}.$$

Since all points of G_1 are adjacent to all points of G_2 in

$G_1 + G_2$, the $p_1 \times p_2$ adjacency matrix of the join is given by

$$\begin{pmatrix} A_1 & J \\ J & A_2 \end{pmatrix} .$$

The operations of join and union described above were defined only for disjoint graphs. When two graphs have the same point set V, we also speak of their union, intersection, and symmetric difference. The union $G_1 \vee G_2$ of two graphs G_1 and G_2 with the same point set V and respective line sets X_1 and X_2 is the graph having the point set V and line set $X = X_1 \cup X_2$. The intersection $G_1 \wedge G_2$ is the graph with point set V which contains the line x if x is in both X_1 and X_2. The symmetric difference $G_1 \oplus G_2$ is also defined on V and contains the line x if x is in X_1 or X_2 but not in both. Some examples are shown in Figure 22.

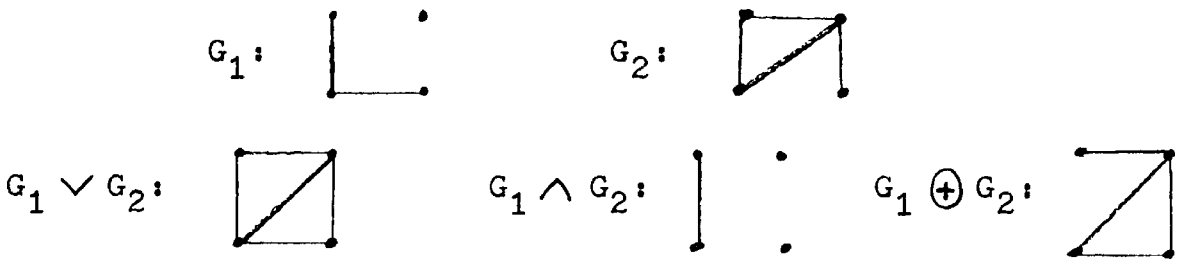


Figure 22. Union, intersection, and symmetric difference of graphs

This second operation of union is the basis for one kind of decomposition of graphs, called factorization. An n-factor of a graph G is a spanning subgraph of G which is regular of degree n, and we say that G is n-factorable (or has an n-factorization) if it can be written as the union of mutually disjoint n-factors. As an example of the case n=1,

in Figure 23 both G_1 and G_2 have a 1-factor, but only G_1 is 1-factorable. We observe that a 1-factor is a collection of disjoint lines and therefore a graph having a 1-factor must have an even number of points. Conversely, it is not difficult to show that every complete graph on an even number of points has a 1-factor and is, in fact, 1-factorable. Tutte[13] has characterized graphs having a 1-factor in the following way.

Theorem 19. A graph has a 1-factor if and only if it has an even number of points and if there is no subset S of points such that the number of odd components of G - S exceeds |S|.

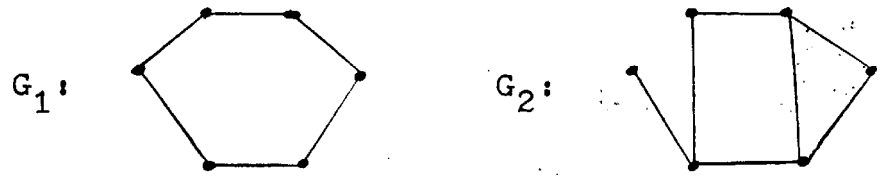


Figure 23. Graphs having 1-factors

A 2-factor is a collection of disjoint cycles. If it is a single cycle, then the graph is hamiltonian. Every complete graph on an odd number of points is 2-factorable, and Beineke [14] has shown that K_{2n+1} can always be written as the disjoint union of spanning cycles.

In Figure 24, the complete graph K_4 has been drawn in two ways. In the first example the diagonals cross each other, while in the second the lines intersect only at points

---

[13] W.T.Tutte, The factors of graphs, Canad. J. Math. 4(1952), 314-328.

[14] L. Beineke and F. Harary, On the thickness of the complete graph, Bull. Amer. Math. Soc. 70(1964), 618-620.

of the graph. Any graph which can be drawn so that its lines intersect only at points of the graph is called planar. Two graphs which are not planar, K_5 and $K_{3,3}$ are also shown in Figure 24, with one intersection apiece.

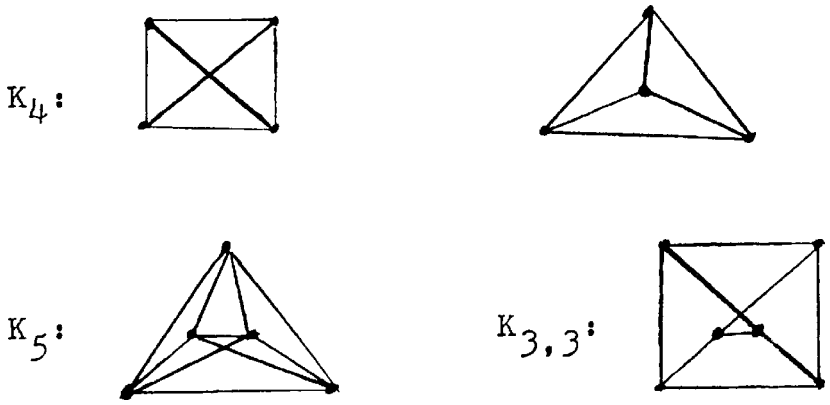


Figure 24. Planar and nonplanar graphs

An elegant theorem by Kuratowski[15] characterizes all nonplanar graphs, and equivalently, all planar graphs. It states that every nonplanar graph must contain as subgraphs K_5 or $K_{3,3}$ or graphs which resemble them in the following way. Two graphs G_1 and G_2 are homeomorphic if isomorphic graphs can be obtained by inserting new points of degree 2 on the lines of G_1 and G_2. For example, in Figure 25, G_1 is homeomorphic to K_5, and G_2 to $K_{3,3}$. Kuratowski's result is now given in Theorem 20.

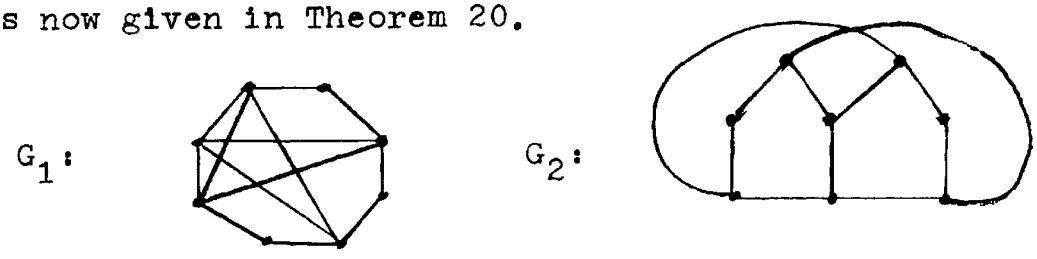


Figure 25. Graphs homeomorphic to K_5 and $K_{3,3}$

---

[15] G. Kuratowski, Sur les problème des courbes gauches en topologie, Fund. Math. 15(1930), 271-283.

Theorem 20. A graph is nonplanar if and only if it has a subgraph homeomorphic to K_5 or $K_{3,3}$.

A graph which is actually drawn in the plane so that its lines intersect only at points of the graph is called a plane graph. In drawing plane graphs, it is often convenient to use curved lines, as in Figure 26. A theorem by Fary[16], however, states that this is not necessary.



Figure 26. Planar and plane graphs

Theorem 21. Every planar graph can be drawn as a plane graph with straight line segments.

A plane graph partitions the plane into several regions, one of which is the infinite exterior region. Letting r denote the number of regions in the partition, we have the following theorem.

Theorem 22. If G is a connected plane graph with p points and q lines which partitions the plane into r regions, then $p - q + r = 2$.

This theorem is an extension of Euler's classical formula for spherical polyhedra, $V - E + F = 2$, where V is the number of vertices, E the number of edges, and F the number of faces of a given polyhedron. Euler's formula

---

[16] I.Fary, On straight line representation of planar graphs, Acta Sci. Math. (Szeged), 11(1948), 229-233.

becomes immediately apparent when the polyhedron is represented as a plane graph, by puncturing any face of the polyhedron and expanding the hole until the rest of the polyhedron lies flat. The punctured face thus corresponds to the exterior region of a graph. The cube with a punctured face and its plane graph are shown in Figure 27.

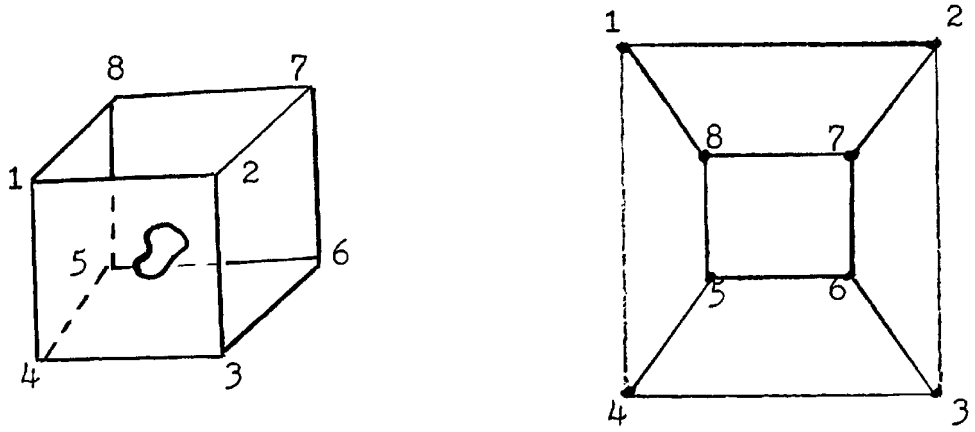


Figure 27. Representation of a cube as a plane graph

5. Trees

The acyclic graph, particularly simple in structure, is perhaps the most widely used graph in applications. A connected acyclic graph is called a tree. All trees with up to six points are shown in Figure 28.

Since trees have found such widespread use in diverse areas, many different definitions have arisen. Four of these are given in the next theorem.[17]

---

[17]A longer list of equivalences may be found in a forthcoming book by F. Harary, Theory of Graphs.

Theorem 23. The following statements are equivalent for a graph G.

1) G is a tree.

2) G is connected and p = q+1.

3) G is acyclic and p = q+1.

4) Every two points of G are joined by a unique path.

5) G is connected and every line of G is a bridge.

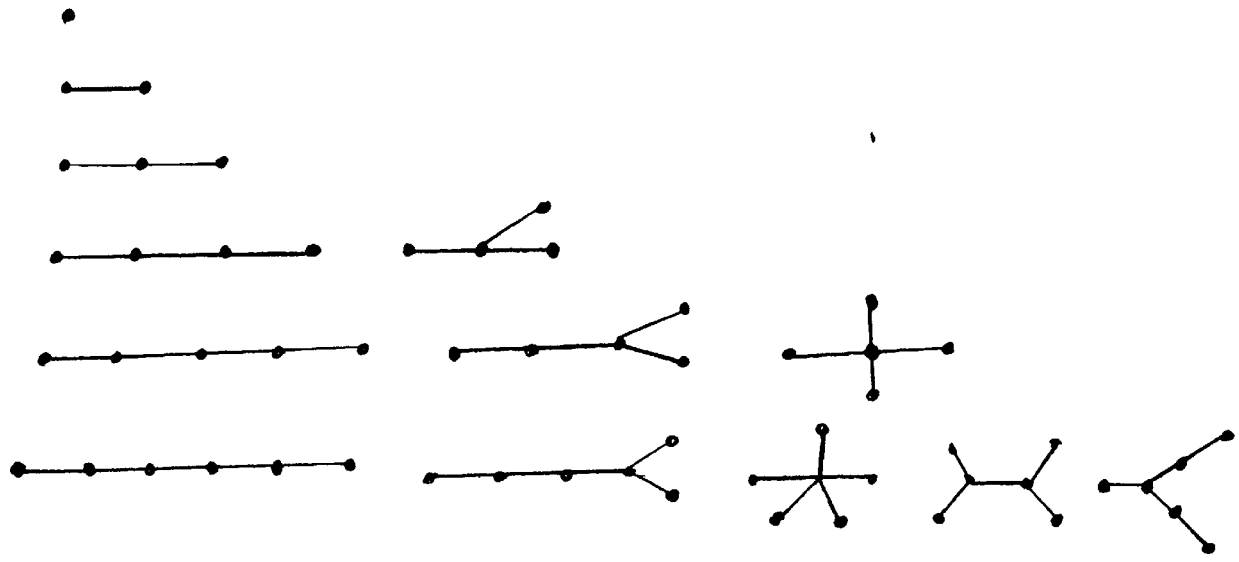


Figure 28. Trees with up to six points

An endpoint in a tree is a point of degree 1. The endpoints are the only points of a tree which are not cut-points. The following lemma provides a very useful tool in the study of trees, particularly in proofs by induction.

Lemma 1. Every tree has at least two endpoints.

The eccentricity e(v) of a point v in a graph G is the maximum of the distances d(u,v) between v and all other points u in G, and the radius r(G) of a graph G is the smallest eccentricity of its points. A point v is a central point if e(v) = r(G); the set of central points is called

the center of the tree. In Figure 29 two trees with five points are drawn with the eccentricity written at each point. Both trees have radius 2, but T_1 has one central point while T_2 has two. Theorem 24 states that no tree can have more than two central points.

Theorem 24. The center of a tree contains either one point or two adjacent points.

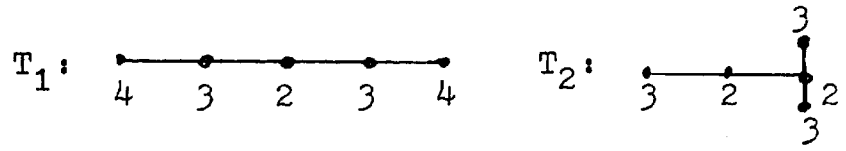


Figure 29. Trees with one and two central points

The center of a tree is "central" with respect to distance, but there is also a concept which involves the center with respect to mass. A subtree of a tree T is any connected subgraph of T. A branch at a point u of a tree T is a maximal subtree containing u as an endpoint, and the maximum number of lines in any branch at u is called the weight at u. The weight at any endpoint is, of course, the total number of lines in the tree itself. In Figure 30 the weights of the non-endpoints have been indicated. A point v is a centroid point of a tree if v has minimum weight. The centroid of a tree is the set of all centroid points. Just as in the case of centers, every tree has either one or two centroid points.

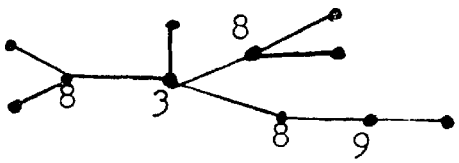


Figure 30. Tree with one centroid point

Theorem 25. Every tree has a centroid consisting of either one point or two adjacent points.

Despite the seeming similarity, the concepts of centroid and center are not closely related. The trees in Figure 31 show that the number of points in the centroid and center of a tree need not even be equal.[18]

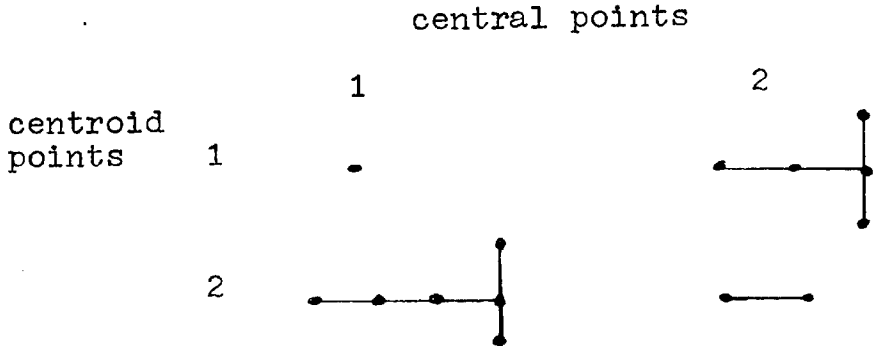


Figure 31. Trees with a given number of central and centroid points

A spanning tree T of a graph G is a subgraph of G which is a tree and which contains all the points of G. Spanning trees play a significant role in the study of graphs in general. We illustrate their use with a theorem which, at first glance, does not involve the concept of a tree, after stating a basic theorem and a few necessary definitions.

Theorem 26. Every connected graph has a spanning tree.

A cycle C is said to depend on a set S of cycles if C can be expressed as the symmetric difference of cycles in S, and the set S is independent if none of its cycles depends

---

[18]Jordan studied these two concepts along with two other centrality concepts in Sur les assemblages de lignes, J. Reine Angew. Math. 70(1869), 185-190.

on the others. In Figure 32 the cycles C_1 and C_2 are independent, while the outside cycle is the symmetric difference $C_1 + C_2$. The maximum number m(G) of independent cycles of a graph G is called the cycle rank of G. For a tree T, m(T) = 0. A formula for m(G) in general, first noted by Kirchhoff[19], is given in the next theorem.

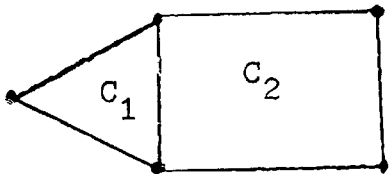


Figure 32. Graph with two independent cycles

Theorem 27. For any connected graph G, m(G) = q-p+1.

The existence of a spanning tree of G is guaranteed by Theorem 26, and this spanning tree has p-1 lines. A maximal set of independent cycles is obtained by adding each of the remaining q - (p-1) lines of G to the tree, creating a single cycle.[20]

The matrix tree theorem for digraphs can be adapted to give the number of spanning trees of a (labeled) graph G. In this form the theorem was first discovered by Kirchhoff.[21]

---

[19]Kirchhoff, 497-508.

[20]Graphs with cycle rank 1 have also been studied. Such graphs consist of a single cycle C together with trees to the points of C, and are called unicyclic. Several characterizations similar to those for trees may be found in S.S.Anderson and F. Harary, Trees and unicyclic graphs, Math. Teacher 60(1966), 345-348.

[21]Kirchhoff, 497-508.

Theorem 28. The number of spanning trees of a graph G is the value of any cofactor in D - A, where D and A are the degree and adjacency matrices of G, respectively.

Applying this theorem to the graph K_p, we obtain the total number of (nonisomorphic) labeled trees with p points, namely, p^{p-2}.

6. Oriented graphs

An orientation of a graph G is an assignment of exactly one direction to each of its lines, and the resulting asymmetric configuration is called an oriented graph. In Figure 33 all oriented graphs with three points are drawn.

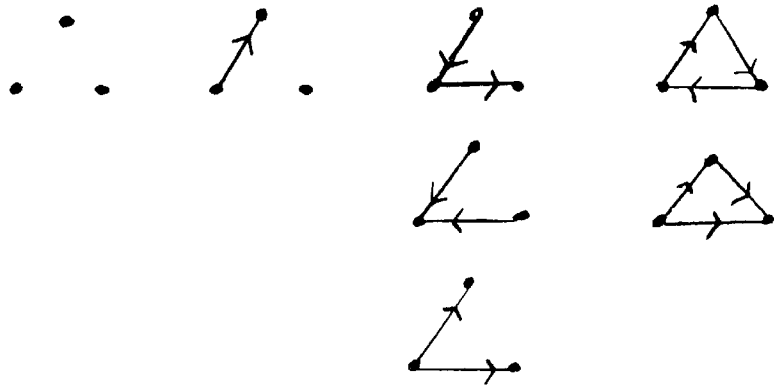


Figure 33. Oriented graphs with three points

A discussion of strong oriented graphs first appeared in the context of a city's street system.[22] An orientation of a network of streets has the effect of converting every street to a one-way street. Robbins considered the problem of deciding when this can be done so that no part of the city is inaccessible from any other. A graph is called

---

[22] H.E.Robbins, A theorem on graphs, with an application to a problem of traffic control, Amer. Math. Monthly 46(1939), 281-283.

strongly orientable if there exists an orientation of G which is strongly connected. Robbins found the following criterion for a graph to possess this property.

Theorem 29. A graph is strongly orientable if and only if it is connected and bridgeless.

We give the construction of a strong orientation for a connected bridgeless graph. Let C_1 be any cycle in G and orient its lines so that a directed cycle is produced. If G and C_1 have the same point sets, the remaining lines may be oriented arbitrarily and the construction is complete. If G contains other points, then there must be a point v not in C_1 which is adjacent to a point of C_1. Let x be the line joining v to C_1, and note that since G is bridgeless, x lies on another cycle C_2. Now orient the lines of C_2 which have not already been oriented consistently in either direction. The oriented subgraph $C_1 \cup C_2$ is strongly connected by construction. This process may be repeated until all points of G are included. The graph G in Figure 34 has been strongly oriented by this construction.

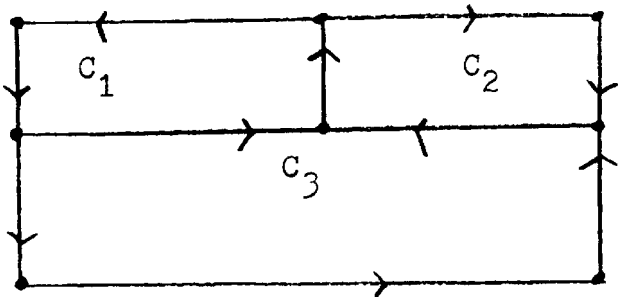


Figure 34. Strongly oriented graph

A second type of orientation involves a transitive relation, that is, a relation R on a set S in which, for distinct elements s, t, and u in S, (s,u) is in R whenever

(s,t) and (t,u) are in R.[23] The problem of determining which graphs have a transitive orientation T(G) is more difficult and, as yet, unsolved. Wolk[24] has characterized those having a transitive orientation which also satisfies the property P_1: If neither (u,v) nor (v,u) is in T(G), then there is no w such that (u,w) and (v,w) are in T(G). A graph G is said to have the diagonal property if whenever (u,v), (v,w), and (w,t) are in G, then (u,w) or (v,t) is in G.

Theorem 30. A graph G has a transitive orientation satisfying property P_1 if and only if G has the diagonal property.

In addition, Wolk[25] has shown that the complement of a graph with the diagonal property also has that property.

Finally, some graphs have the property that every orientation is self-converse. Obviously K_1 and K_2 are in this class. The graph K_3 admits two orientations, one cyclic and one transitive, both of which are self-converse. It has been shown[26] that the only graphs having this property are K_1, K_2, K_3, and cycles of lengths 4 and 5.

---

[23] The usual definition of "transitive" omits the word "distinct". However, we do not wish to require that a transitive relation include the pair (s,s) whenever it includes (s,t) and (t,s).

[24] E.S.Wolk, The comparability graph of a tree, Proc. Amer. Math. Soc. 13(1962), 789-795.

[25] E.S.Wolk, A note on "The comparability graph of a tree", Proc. Amer. Math. Soc. 16(1965), 17-20.

[26] F.Harary, E.M.Palmer, and C.Smith, Which graphs have only self-converse orientations?, Canad. Math. Bull. 10 (1967), 425-429.

7. Tournaments

A complete oriented graph is called a tournament, from its connection with round-robin competition. All the tournaments with two, three, and four points are shown in Figure 35. The first with three points is called a transitive triple, the second is a cyclic triple. We note that

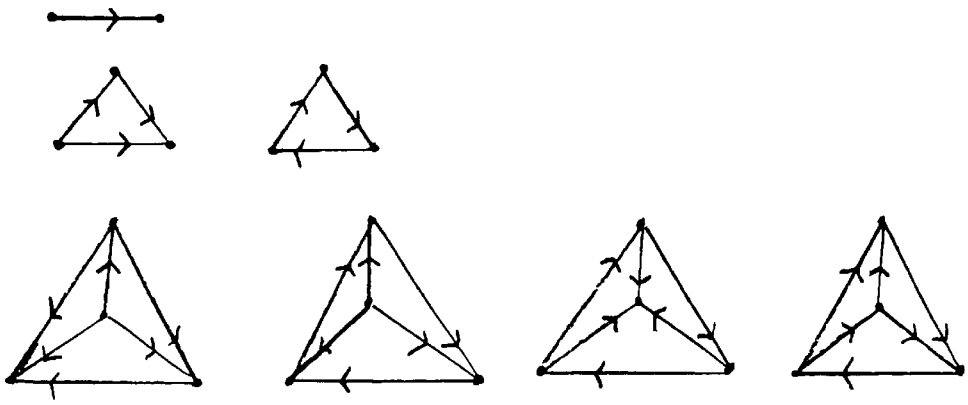


Figure 35. Tournaments with two, three, and four points

since a tournament T is complete, every point v in T satisfies the degree equation

$$id(v) + od(v) = p - 1$$

and hence that

$$q = \frac{p(p-1)}{2} .$$

In the adjacency matrix A(T), the entry $a_{ij} = 0$ if and only if $a_{ji} = 1$, for $i \neq j$. Conversely, a square (0,1)-matrix A is the adjacency matrix of a tournament if and only if $A + A' = J - I$. We remark that the converse T^c of a tournament is also a tournament, and the complement $\overline{T}$, whose adjacency matrix is also the matrix $J - A - I$, is identical with the converse T^c.

It is clear that every complete graph can be oriented in such a way that the resulting tournament is strongly connected, since a complete graph has many spanning cycles. Not so obvious is the fact that every orientation of a complete graph yields a tournament with a spanning path. This theorem, proved by Redei[27] in 1934, was the first theorem on tournaments.

Theorem 31. Every tournament has a spanning path.

It follows from this theorem that every tournament is unilateral. In addition, every tournament has at least one point which is "central" in the following sense.

Theorem 32. In a tournament, the distance from a point of maximum outdegree to every other point is 1 or 2. Dually, the distance to a point of maximum indegree from every other point is 1 or 2.

In empirical situations such as round-robin tournaments, the major objective is often to rank the teams in order of their competence. When a tournament has a unique spanning path, this can be done easily and unambiguously, since each team has beaten every subsequent team in the path; such a tournament is transitive. Transitive tournaments, also called "complete orders", will be studied in detail in the next section. When a tournament has several spanning paths, however, there will be several possible rankings in which each team defeats the team following it

---

[27] L. Redei, Ein Kombinatorischer Satz, Acta Sci. Math. (Szeged) 7(1934), 39-43.

in the list, but not necessarily every subsequent team. In fact, if the tournament contains a spanning cycle, any team can be ranked first. In these cases a different kind of ranking is desirable, usually ranking with respect to outdegree (wins), the winner being the central point of Theorem 32.

When a tournament is strong, it contains not only a spanning closed walk, but also a spanning cycle. Harary and Moser[28] have shown that a strong tournament contains cycles of every length.

Theorem 33. A strong tournament contains cycles of every length $k = 3,4,\ldots,p$.

Furthermore, Moon[29] has shown that every point of a strong tournament lies on a cycle of each length $k = 3,4,\ldots,p$. The reader may verify this for the case $p = 4$ by referring to Figure 35.

Theorem 33 also holds, of course, for a strong component of a tournament, and a common method of ranking the points of a tournament is to rank the strong components, considering the points in each component equivalent in some sense. In every tournament there is a unique path which spans the components, and the next theorem indicates how this may be constructed.

---

[28] F.Harary and L.Moser, The theory of round robin tournaments, Amer. Math. Monthly 73(1966), 231-246.

[29] J.W.Moon, On subtournaments of a tournament, Canad. Math. Bull. 9(1966), 297-301.

Theorem 34. For any two strong components S_1 and S_2 of a tournament T, the outdegrees of the points of one, say S_1, are all less than the outdegrees of the points of the other. Moreover, the arcs joining the points of S_1 to the points of S_2 are all oriented from S_2 to S_1.

As an illustration, the tournament in Figure 36 has strong components S_1 with points u_1,u_2,u_3, all having outdegree 1, and S_2 with points v_1,v_2,v_3, all having outdegree 4.

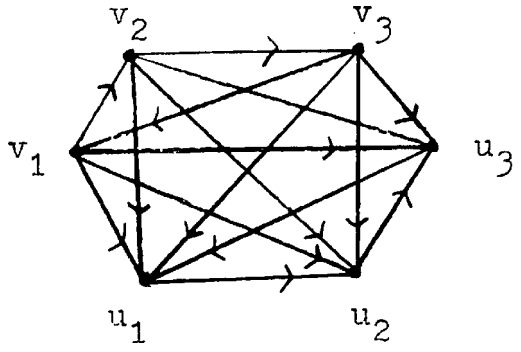


Figure 36. Tournament labeled by strong components

Theorem 34 immediately implies that in a tournament all points having the same outdegree are in the same component. The converse, however, is not true. The strong tournament T is Figure 37, with outdegrees labeled, provides a counterexample. Thus a ranking by outdegrees would be a refinement of a ranking by strong components.

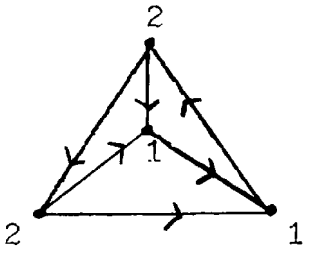


Figure 37. Strong tournament

8. Orders

Some of the problems encountered in attempting to rank the points of a tournament can be resolved by requiring that the relation be transitive, for in a transitive relation R, if u is ranked higher than v (i.e., if (u,v) is in R), then u is ranked higher than every point below v. We discuss several types of transitive relations, also called order relations.

A quasi-order is a reflexive transitive relation. The relagraph of the quasi-order consisting of the set $\{1/3, 1/2, 2/4, 2/3\}$ under $\leq$ is shown in Figure 38. This example illustrates the fact that a quasi-order is not necessarily acyclic.

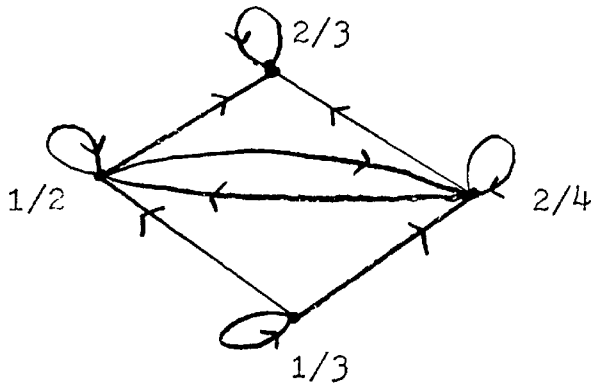


Figure 38. Quasi-order

A strict partial order is an irreflexive, asymmetric, transitive relation or, equivalently, a transitive oriented graph. The word "strict" refers to the irreflexivity of the relation and henceforth is omitted. The relation of proper set inclusion, $A \subset B$, is a partial order on any collection of sets, and, in particular, the n-cube Q_n, consisting of all subsets of a set with n elements, is a partial order.

A partial order P is usually represented by a Hasse diagram. We construct such a diagram by a strategic placement of the points of the underlying set in the plane so that whenever the pair (u,v) is in the relation, u is "higher" than v. A line is drawn from u to v if and only if (u,v) is in P and there is no point w such that (u,w) and (w,v) are in P. (The arrowheads are omitted since they are implied by the fact that u is higher than v.) Since a partial order P is transitive, (u,v) is in P whenever there is a path from u to v in the diagram. Hence there are many more pairs in the partial order than there are lines in the Hasse diagram. In Figure 39 the Hasse diagrams of the cube Q_3 and the proper divisors of 12 are drawn. Although these diagrams resemble graphs, we emphasize that an orientation is understood. The digraphs corresponding to the same relations are also shown; their complexity correctly suggests that in many applications, the Hasse diagram representation is more useful.

The existence of a Hasse diagram for every partial order is ensured by the following fundamental property of partial orders.

Theorem 35. Every partial order is acyclic.

Now applying Theorem 5 on acyclic digraphs, we get several corollaries.

Corollary 35a. Every strong component of a partial order consists of a single point.

Corollary 35b. Every walk of a partial order is a path.

Corollary 35c. The points of a partial order may be labeled in such a way that its adjacency matrix is upper triangular.

Finally, Theorem 5 together with Theorem 3 (and 3') implies the following result.

Corollary 35d. A partial order has at least one transmitter and at least one receiver.

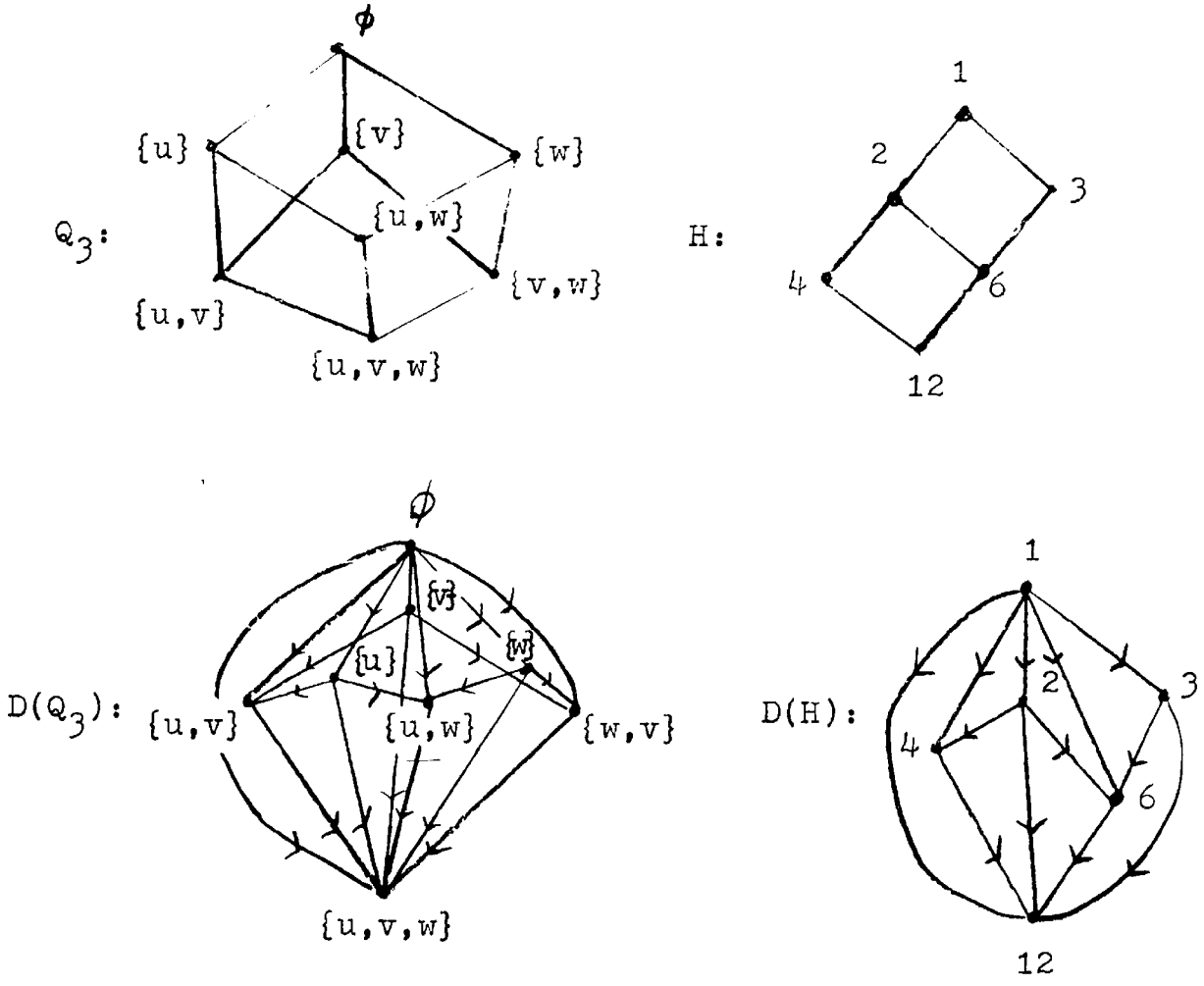


Figure 39. Hasse diagrams and digraphs of partial orders

The fact that a partial order is acyclic allows us to assign a "level" m_i to each point v_i in such a way that $n_i < n_j$ whenever (v_i, v_j) is in P. One method of assigning

levels is to assign the level 0 to each transmitter and to assign to every other point v the maximum length of all paths to v from a transmitter; the levels of the partial order in Figure 40 have been assigned in this manner. The following theorem asserts that, in fact, this is the most "efficient" method of assigning levels.

Theorem 36. If n is the length of the shortest path in an acyclic digraph D, then n+1 is the smallest number of levels in any level assignment of D.

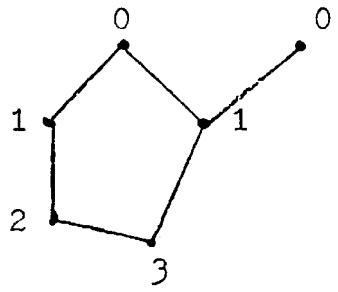


Figure 40. Partial order with level assignment

If two points v_i and v_j of a partial order P have been assigned the same level, then necessarily neither (v_i,v_j) nor (v_j,v_i) is in P; v_i and v_j are called incomparable if neither (v_i,v_j) nor (v_j,v_i) is in P. A (strict) complete order is a partial order in which every two elements are comparable; thus it is an irreflexive, asymmetric, transitive, complete relation. As in the case of partial orders, the word "strict" refers to the irreflexivity of the relation and is omitted in the following discussion. The Hasse diagram H and the digraph D of the complete order $<$ on the set of points $\{1,2,3,4\}$ are shown in Figure 41.

H: 

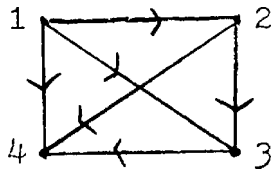
D(H): 

Figure 41. Hasse diagram and digraph of a complete order

It is clear that a complete order can also be defined as a transitive tournament or as an acyclic tournament. Employing a few facts about tournaments, we have the following characterizations of complete orders.

Theorem 37. The following statements are equivalent for a tournament T with p points.

1) T is a complete order, that is, T is transitive.

2) T is acyclic.

3) T has a unique spanning path.

4) T has exactly one point with outdegree i,

for $i = 0,1,\ldots,p-1$.

Corollary 37a. Every complete order has precisely one transmitter and one receiver.

Complete orders may also be characterized by emphasizing their acyclic properties.

Theorem 38. The following statements are equivalent for an acyclic digraph D with p points.

1) D is a complete order.

2) D has $\frac{p(p-1)}{2}$ arcs.

3) Any three points induce a transitive triple subgraph.

4) D is a maximal acyclic digraph, that is, the addition of any new arc creates a cycle.

The outdegrees of the points of a complete order provide a natural level assignment from p-1 to 0, and not only is this assignment unique, but all complete orders with p points are isomorphic.

Theorem 39. Every complete order with p points is isomorphic to the set of integers 1,2,...,p under the relation of <.

This theorem allows us to say that the complete order on four points illustrated in Figure 41 is the complete order on four points. Using some of the properties of the set of integers 1,2,...,p, under the relation <, we obtain the following corollaries.

Corollary 39a. The points of a partial order can be labeled in such a way that v_i is adjacent to v_j if and only if i<j.

Corollary 39b. The adjacency matrix of a complete order has all 1's above the diagonal and all 0's on and below the diagonal.

9. Lattices

In this section we discuss partial orders having two special points s and t such that s is higher than all points of the partial order and t lower. To be more precise, we define an upper bound of the set V of a partial order P to be any element x such that (v,x) is in P for all v in V. The upper bound x is called the least upper bound (l.u.b) of V if for any other upper bound y, (x,y) is in P. Similarly, we define lower bound and greatest lower bound (g.l.b.).

A partial order in which every pair of elements has a l.u.b. and a g.l.b. is called a lattice.[30] Every complete order is a lattice, as are the cubes Q_n. The Hasse diagrams of the five lattices on five points are shown in Figure 42. We note that graphically a lattice is connected, acyclic, and has a source and sink.

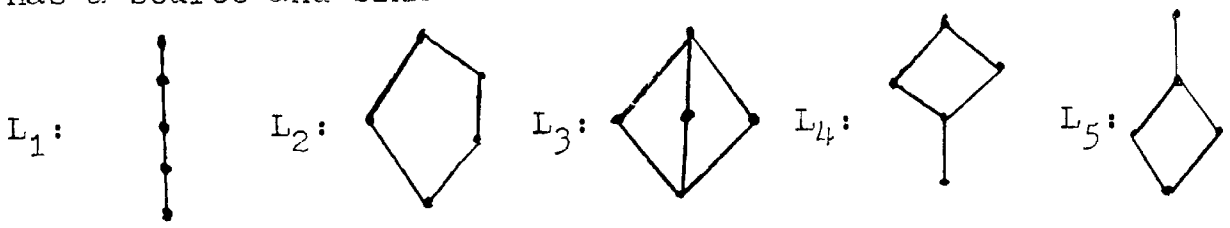


Figure 42. Lattices with five points

Equivalently, a lattice may be defined as a set with two binary operations $\cap$ and $\cup$, which are associative, commutative, idempotent, and satisfy the absorption law $a \cap (a \cup b) = a \cup (a \cap b) = a$. It is not difficult to show that the two definitions are equivalent, by defining $a \cup b$ = l.u.b.(a,b) and $a \cap b$ = g.l.b.(a,b). The identity element for $\cup$ is the source of the partial order, and the identity element for $\cap$ is the sink.

A lattice which in addition satisfies the law

$$(x,z) \text{ in } P \Rightarrow x \cap (y \cup z) = (x \cap y) \cup z$$

is called modular. We notice that all but the lattice L_2 in Figure 42 are modular. This lattice in fact characterizes all modular lattices, in the sense of the following theorem.

Theorem 40. A lattice is modular if and only if it contains no sublattice isomorphic to the pentagonal lattice L_2.

---

[30] D.E.Rutherford, Introduction to Lattice Theory, Hafner Publishing Co., 1965, provides a good reference for lattice definitions and properties.

A second type of lattice is the distributive lattice, that is, a lattice which satisfies the property that

$$x \cap (y \cup z) = (x \cap y) \cup (x \cap z).$$

If (x,z) is in P, we have $x \cap z = z$ and thus every distributive lattice is modular. The converse is not true, however, since L_3 in Figure 42 is not distributive. The next theorem states that modular lattices having a sublattice isomorphic to L_3 are the only ones which are not distributive.

Theorem 41. A lattice is distributive if and only if it has no sublattice isomorphic to either L_2 or L_3.

Alvarez[31] has examined the question of when a given graph is the Hasse diagram of a lattice. His graphical characterization of modular lattices is given in the next theorem.

Theorem 42. The graph G is the Hasse diagram of a modular lattice if and only if G satisfies the following three properties.

1) G is connected and has no cycles of odd length.

2) There are two points a_1, a_2 in G such that the distance $d(a_1,a_2)$ is the maximum distance between any two points in the graph. Further, if (c,d), (d,e) are distinct arcs of G with $d(a_i,c) = d(a_i,d) = d(a_i,e) + 1$, then there is a unique point f_i in G such that $d(a_i,f_i) = d(a_i,e) + 2$ and (f_i,c), (f_i,d) are in G, for $i = 1,2$.

---

[31] L.R.Alvarez, Undirected graphs realizable as graphs of modular lattices, Canad. J. Math. 17(1965), 923-932.

3) If v is a point of Q_k, then Q_k is a subgraph of G whenever $Q_k - v$ is.

Lattices complete our study of specific classes of relations from a graph-theoretic viewpoint. In the following chapters we utilize the concepts, theorems, and methods developed here in the study of other topics, including colorability, finite groups, and finite projective planes. In certain applications a more specialized type of relation than the ones we have described is necessary, but the techniques of this chapter may be applied to define and describe them as well.

II. COLORABILITY

The Four Color Conjecture--the conjecture that four colors are sufficient to color any geographical map so that no neighboring countries have the same color--has been the single most important stimulus to graphical research. The conjecture is attributed to ancient cartographers, but its first appearance in mathematical circles seems to have occurred in DeMorgan's classes[1], and a formal presentation of it was given to the London Mathematical Society by Cayley in 1878. Since that time, the most determined efforts of many well-known mathematicians--G.D.Birkhoff, H. Whitney, P.J.Heawood, O.Ore, W.T.Tutte, to name a few--have not succeeded in producing a solution to this deceptively simple problem.

Although the original statement of the Four Color Conjecture is given in terms of coloring the regions of a map, it is not difficult to show that it is equivalent to the problem of coloring the points of a planar graph. One of the fascinations of the Four Color Conjecture is its many equivalent formulations. For example, Tait has related

---

[1]A.B.Kempe, On the geographical problem of the four colors, Amer. J. Math. 2(1879), 193-194.

the problem to one involving the existence of 1-factors or the existence of a line-coloring; and Hadwiger has reduced it to a question about the contractions of a graph. It can further be shown that the existence of a certain binary function on the points of a planar graph is equivalent to a 4-coloring. These, and other equivalences, have all led to a great deal of graphical research much of which has independent interest and value.

In this chapter we first discuss briefly the general concept of colorability, and then we describe the Four Color Conjecture in detail. Various results which have reduced the conjecture to a problem concerning cubic planar blocks are mentioned, and several equivalent forms of the conjecture are presented. In closing we state several peripheral results, such as the chromatic number of Euclidean 3-space, inspired by the Four Color Conjecture.

A graph is <u>k-colorable</u> if its points can be colored with k colors in such a way that no two adjacent points have the same color. The smallest number k for which such a coloring is possible is called the <u>chromatic number</u> $\chi(G)$ of the graph G. Clearly any 1-colorable graph is totally disconnected. If $\chi(G) = 2$, then it is possible to partition the set V of points into two subsets in such a way that every line of the graph joins a point in one subset with a point in the other. Therefore a graph with chromatic number 2 is a bipartite (bicolorable) graph as defined in Chapter I. Some examples of bipartite graphs are given in Figure 1.

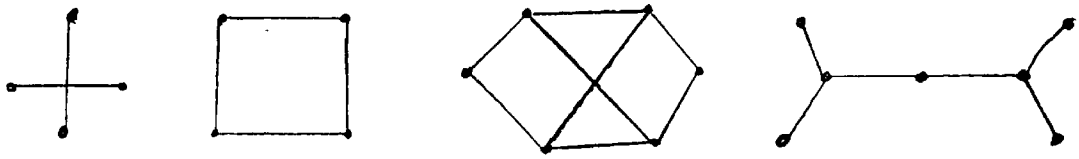


Figure 1. Bipartite graphs

It is clear that all trees are bipartite, as are all cycles of even length. Theorem 1 gives a characterization of bipartite graphs in terms of their cycles.

Theorem 1. A graph is bipartite if and only if each of its cycles has even length.

No simple structural characterization of graphs with chromatic number three has yet been found. Brooks[2], however, has proved an important theorem on k-colorability for arbitrary k which gives a partial result for $\chi = 3$.

Theorem 2. Let G be a connected graph, not K_p, with maximal degree $d \geq 3$. Then G is d-colorable.

Corollary 2a. Every connected cubic graph except K_4 is 3-colorable.

The characterization of graphs with chromatic number four brings us to the Four Color Conjecture, the most famous outstanding problem in graph theory. In its original form as a map-coloring problem, the Four Color conjecture states that no two regions sharing a common boundary of more than a point have the same color. Its graphical equivalent asserts

---

[2]R.L.Brooks, On colouring the nodes of a network, Proc. Cambridge Philos. Soc. 37(1941), 194-197.

that every planar graph is 4-colorable.

Four Color Conjecture I. Every planar graph is 4-colorable.

The exact nature of the equivalence is now shown. A map can be viewed as a planar graph or, permitting multiple lines and loops, multigraph. A coloring of the map corresponds to a coloring of the faces of the graph, including the exterior face. To relate the faces of a graph to its points, we define the Whitney dual G* of a planar graph G to be the graph (or multigraph) obtained by placing a point v(F) in each face F of G (including the exterior face) and by drawing a line x* between $v(F_i)$ and $v(F_j)$ whenever F_i and F_j have the line x as a common boundary. A graph and its (Whitney) dual are shown in Figure 2, with the construction of the dual indicated in the middle. The dual graph is always planar, and the operation of taking the dual is idempotent, $(G^*)^* = G$.

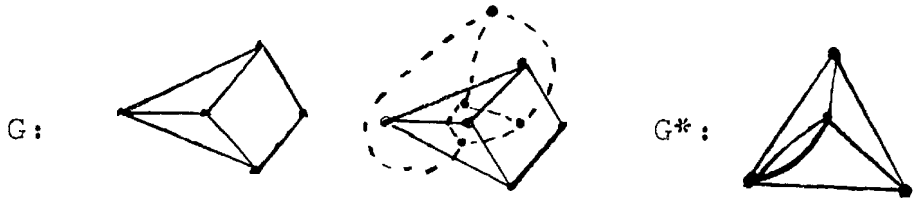


Figure 2. Graph and its (Whitney) dual

The concept of a dual can also be defined abstractly in terms of relations so that it does not depend on a particular plane representation of a graph. Using such a definition, Whitney was able to characterize planar graphs in terms of their (Whitney) dual graphs.[3]

---

[3] H. Whitney, Non-separable and planar graphs, Trans. Amer. Math. Soc. 34(1932), 339-362.

Theorem 3. A graph is planar if and only if it has a (Whitney) dual.

Since a coloring of the faces of a planar graph G induces a coloring of the points of G*, Theorem 3 implies that the map coloring problem and the graph coloring problem are equivalent. Calling the least number of colors necessary to color the faces of a planar graph G the face-chromatic number $\varphi(G)$, we have the equalities $\varphi(G) = \chi(G^*)$ and $\varphi(G^*) = \chi(G)$. We refer to the face-coloring form of the problem as the Four Color Conjecture II.

Four Color Conjecture II. Every planar graph has face-chromatic number $\varphi(G) \leq 4$.

In considering the face-chromatic number of a planar graph G, we first note that $\varphi(G)$ is equal to the largest of the chromatic numbers of the blocks of G. It is therefore sufficient to consider only planar blocks. Furthermore we now show that it is possible to restrict our study of $\varphi(G)$ to cubic graphs. Obviously points of degree 2 may be suppressed without affecting the face-coloring of a graph. If v is a point with degree greater than 3, it may be replaced by a small "patch" without altering any of the previous face-adjacencies (as shown in Figure 3), so that all of the new points have degree 3. If the new graph can be face-colored with four colors, then by shrinking the patch back to a point, we obtain a face-coloring of the original graph with four colors, and therefore we need to consider only cubic graphs.

Corresponding to the concept of the degree of a point, we define the degree deg(F) of a face F to be the number of lines bounding F. We now show that, without loss of generality, we may restrict our study to graphs in which deg(F) ≥ 5 for every face F. First, if a graph G has a face F with deg(F) ≤ 3, adjacent to at most three other faces, then there is always a fourth color available for F. Now suppose the graph G has a face F_0 with exactly four adjacent faces F_1, F_2, F_3, F_4, as shown in Figure 4. By the Jordan Curve Theorem it cannot be the case that both F_1 adj F_3 and F_2 adj F_4, and since we may assume that F_1 is not adjacent to F_3, we may merge F_0 with F_1. If the resulting graph can be colored with four colors, then a 4-coloring of the original graph can be obtained by reinstating the face F_1 and giving it the color of F_3. Thus all faces having degree less than 5 may be suppressed. Summarizing these results, we have the following theorem.

Theorem 4. It is sufficient to prove the Four Color Conjecture II for cubic planar blocks in which every face has degree at least 5.

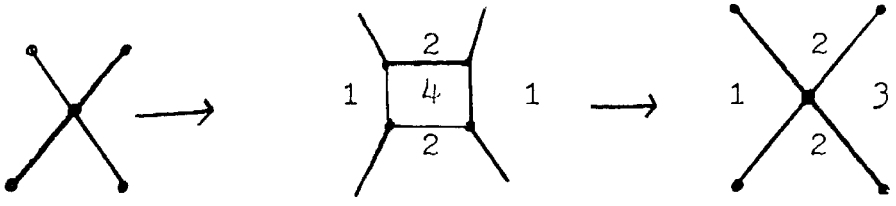


Figure 3. Replacing points of degree >3

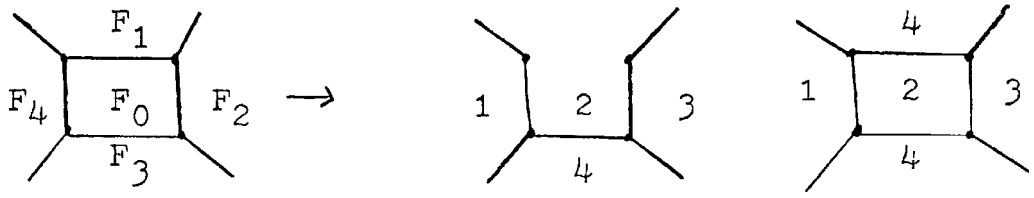


Figure 4. Removing faces of degree 4

The necessity of four colors is demonstrated by the complete graph K_4 which has four mutually adjacent faces. The sufficiency of five colors has long been known, and may be proved[4] by reasoning similar to that used above in suppressing regions with degree 4.

Five Color Theorem. Every planar graph can be face-colored with 5 colors.

Other approaches to the Four Color Conjecture[5] have led to many equivalent forms of the problem. Tait thought he had solved the problem in 1880, but in his "proof" he assumed the following "theorem", now known as the Tait conjecture, which is equivalent to the Four Color Conjecture.[6] Its equivalence becomes apparent in the discussion below.

Equivalence 1. Every cubic planar graph can be written as the union of three 1-factors.

Petersen[7] has proved that cubic blocks can always be written as the union of a 1-factor and a 2-factor, and he has also shown[8] that the planarity is essential in Tait's conjecture by constructing the "Petersen graph"--a nonplanar

---

[4] R. Courant and H. Robbins, What is Mathematics?, Oxford University Press, New York, 1941, 264-267.

[5] For a detailed study of the Four Color Conjecture in general, see O. Ore, The Four-Color Problem, Academic Press, New York, 1967.

[6] P. G. Tait, On the colouring of maps, Proc. Roy. Soc. Edinburgh Sect. A 10(1880), 501-502, 729.

[7] J. Petersen, Die Theorie der regularen Graphen, Acta Math. 15(1891), 193-220.

[8] J. Petersen, Sur le theorème de Tait, L'Intermediare des Mathematiciens, 5(1898), 225-227.

cubic block which cannot be written as the union of three 1-factors. See Figure 5.

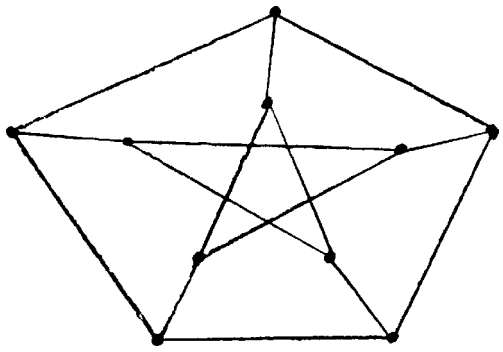


Figure 5. Petersen graph

If a graph G is expressible as the union of three 1-factors, say $G_1 \cup G_2 \cup G_3$, then each G_i, $i = 1,2,3$, is a collection of disjoint lines, and G admits a 3-coloring of the lines in which no two lines with the same color have a common endpoint. Thus Tait's conjecture may also be stated using the concept of a line-coloring.

Equivalence 2. Every cubic planar graph can be line-colored with three colors.

A direct correspondence between a line-coloring and a face-coloring is easy to construct. Suppose the faces of the graph have been colored with the colors A, B, C, and D. Then the lines may be colored with the colors X, Y, and Z by the following rule: Any line separating faces colored A and B or C and D is colored with the color X, any line separating faces colored A and C or B and D with the color Y, and any line separating faces colored A and D or B and C with the color Z. Since the graph is cubic, any two lines with a common endpoint share one face and not the other, and hence have different colors. The correspondence is

illustrated for K_4 in Figure 6. The reverse correspondence is similar.

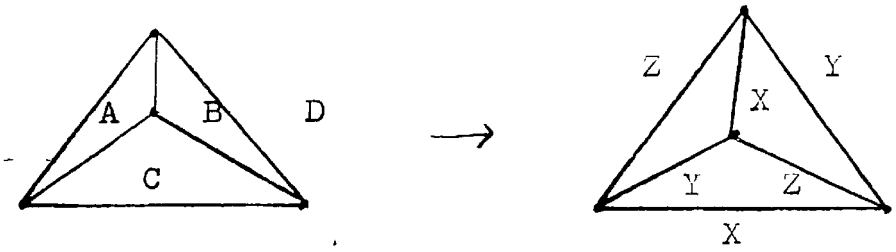


Figure 6. Correspondence between a face-coloring and a line-coloring

The line graph L(G) of a graph G is the graph whose points are the lines of G. Two points in L(G) are adjacent if the corresponding lines in G have a common endpoint. The line graph of K_4 is shown in Figure 7. Equivalence 2 can also be restated in terms of line graphs.

Equivalence 2'. There is a 3-coloring of the points of L(G) for every planar graph G.

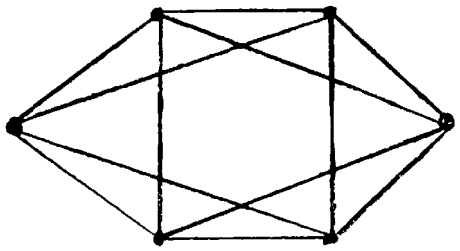
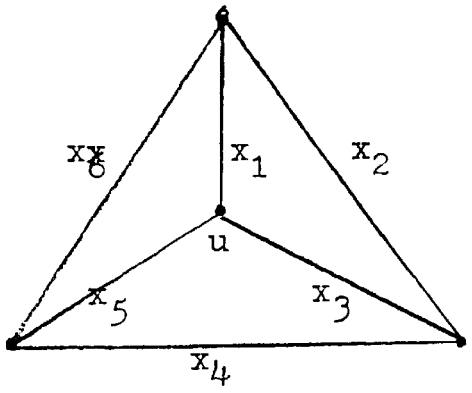


Figure 7. Line graph of K_4

A 3-coloring of the lines of a planar cubic graph may be expressed in terms of the color function γ, where $\gamma(x) = 0$, 1, or -1, according to the color of line x. We consider the values 0,1,-1 to be cyclically oriented in that order. An angle is defined to be a pair (x_1,x_2) of lines which meet at a point; the character of the angle (x_1,x_2), denoted $\tau(x_1,x_2)$, is defined to be 1 if $\gamma(x_1)$

immediately precedes $\gamma(x_2)$ in the cycle, and -1 if $\gamma(x_1)$ is the successor of $\gamma(x_2)$ in the cycle. We illustrate this with the graph in Figure 8. The function $\gamma(x)$ has been assigned according to the line-coloring of G given in Figure 6.



$\gamma(x_1)=\gamma(x_4)=0$ $\quad \kappa(x_1,x_2)=1$

$\gamma(x_2)=\gamma(x_5)=1$ $\quad \kappa(x_2,x_1)=-1$

$\gamma(x_3)=\gamma(x_6)=-1$ $\quad \kappa(x_5,x_3)=1$

Figure 8. Assignment of a color function

By the definition of the character function κ, it is clear that the coloring function γ must satisfy the rule $\gamma(x_1) = \gamma(x_2) + \kappa(x_2,x_1) \pmod 3$, where x_1 and x_2 are lines which meet at a point. Therefore the sum of the characters of the angles around any face must be 0 (mod 3), since

$$\gamma(x_0) = \gamma(x_0) + \kappa(x_0,x_1) + \kappa(x_1,x_2) + \cdots + \kappa(x_n,x_0) \pmod 3.$$

Furthermore, in a cubic graph, the lines at any point v are oriented either clockwise or counterclockwise according to the color function. For example, in the graph of Figure 8, the lines at every point are oriented in a counterclockwise direction. If the lines at v are ordered in a clockwise direction, then the angles at v, also taken in a clockwise

direction, will all have character 1, and we therefore define $\kappa(v) = 1$. Similarly, if the lines are oriented counterclockwise, we define $\kappa(v) = -1$. Thus we may also say that $\Sigma \kappa(v_i) = 0 \pmod 3$, where the v_i are the points on the boundary of any face of the graph. The existence of such a function of the points provides a third equivalent of the Four Color Conjecture.[9]

Equivalence 3. A character $\kappa(v) = \pm 1$ can be assigned to each point in a cubic planar graph in such a way that the points of any minimal cycle satisfy the conditior $\Sigma \kappa(v) = 0 \pmod 3$.

A third characterization of planar graphs leads to another equivalent form of the Four Color Conjecture. We say that G_1 is an elementary contraction of G if G_1 is the result of identifying two adjacent points of G. A contraction is the configuration resulting from a sequence of such identifications. Figure 9 shows how a particular graph may be contracted to K_3. It can be shown that the Kuratowski criterion[10] for planarity holds for contractions as well as homeomorphisms.[11]

Theorem 5. A graph G is planar if and only if it has no subgraph contractible to K_5 or $K_{3,3}$.

This theorem suggests that homeomorphisms and

---

[9] Ore, The Four Color Problem, 112-116.

[10] See Chapter I, p. 39.

[11] F. Harary and W.T. Tutte, A dual form of Kuratowski's theorem, Bull. Amer. Math. Soc. 71(1965), 168.

contractions might be equivalent. That this is not the case, however, is shown by the Petersen graph in Figure 5, which is clearly contractible to K_5, but has no subgraph homeomorphic to K_5. (The Petersen graph is nonplanar since it contains a subgraph homeomorphic to $K_{3,3}$.)

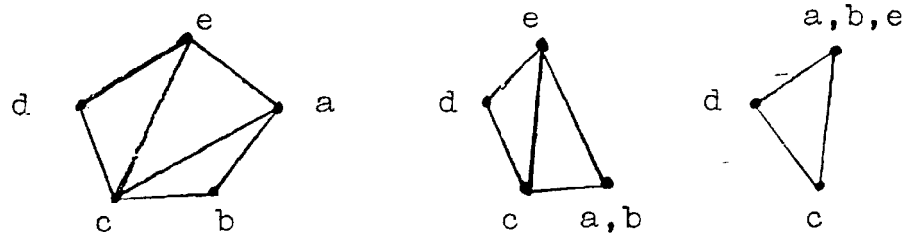


Figure 9. Contraction of a planar graph

A very general conjecture by Hadwiger, which contains the Four Color Conjecture as a special case, is stated in terms of contractions.

Hadwiger's Conjecture. If G is an arbitrary graph with chromatic number $\chi(G) = n$, then G can be contracted to K_n.

For n=5 Hadwiger's conjecture immediately implies the 4-colorability of planar graphs, since a graph with chromatic number 5 would then be contractible to K_5 and hence would be nonplanar by the Kuratowski criterion. It is interesting to note that Hadwiger's conjecture makes no reference to $K_{3,3}$ and thus suggests that perhaps planarity is not the crucial condition for 4-colorability but rather contractibility to K_5.

The converse, that the Four Color Conjecture implies Hadwiger's conjecture for n=5, is far more difficult to

establish; the first proof was given by Wagner.[12]

Equivalence 4. Every graph with chromatic number $\chi(G) = 5$ can be contracted to K_5.

Hadwiger's conjecture can be proved without difficulty for $n < 5$. Still more tantalizing is the theorem[13] that whenever Hadwiger's conjecture is true for n, it is also true for all $k < n$.

A different approach to the Four Color Conjecture was devised by Birkhoff[14], who defined the chromatic polynomial $P(G,n)$ of a graph G, which counts the number of ways G can be colored with n colors. For example, $P(\overline{K}_p,n) = n^p$, since any of the p points may have any of the n colors, and $P(K_p,n) = n(n-1)\ldots(n-p+1)$. There are several ways of constructing this polynomial. Birkhoff developed a set of linear equations whose determinant is the chromatic polynomial. Since that time, chromatic polynomials have been the subject of much research, and simpler ways of obtaining the polynomial have been found. One method[15] reduces all chromatic polynomials to the sums of polynomials of complete graphs; the construction uses the fact that $P(G,n) = P(G + (u,v),n) + P(G_{u=v},n)$, where

+

---

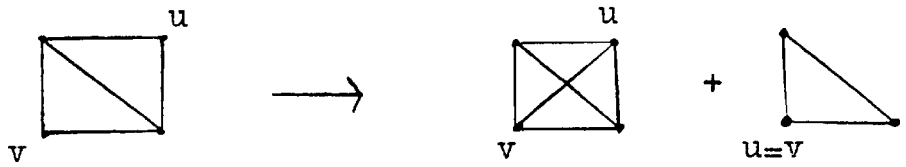
[12] K.Wagner, Uber eine Eigenschaft der ebenen Komplexe, Math. Ann. 114(1937), 570-590.

[13] Ore, The Four Color Problem, p.158

[14] G.D.Birkhoff, A determinant formula for the number of ways of coloring a map, Ann. of Math. 14(1912), 42-46.

[15] See, for example, F.Harary, Theory of Graphs, forthcoming.

$G + (u,v)$ is the graph formed by adding the line (u,v) to G for a pair of nonadjacent points u and v in G and $G_{u=v}$ is the graph formed from G by identifying u and v. For example, applying this construction to the graph in Figure 10, we obtain the results that $P(G,n) = n(n-1)(n-2)^2$ and $\chi(G)$, the minimal value of n for which the polynomial does not vanish, is 3.



G $\qquad$ G+(u,v) $\qquad$ $G_{u=v}$

Figure 10. Construction of P(G,n)

The principal aim of the research on the chromatic polynomial has been, of course, to prove the following equivalent of the Four Color Conjecture.

Equivalence 5. The chromatic polynomial $P(G,4)$ is greater than 0 for all planar graphs G.

The problem of the 4-colorability of planar graphs has been irresistable to countless mathematicians. Heawood[16] even tried a probabilistic argument which resulted in the following assertion. The probability of being unable to color a regular map of F regions with four colors is of the order of $(1 - 3^{-F})^{2^V}$, or, when F is fairly large, $e^{-(4/3)F}$. For example, if $F > 35$, there is less than 1 chance in $10^{10,000}$.

---

[16]This fact is reported in P.Franklin, The four color problem, Galois Lectures, Scripta Mathematica Library, No. 5, (1941), 49-85.

The problem of coloring other surfaces and higher dimensional spaces has met with more success. In his essay on colorability, Franklin[17] reports the following results.

Theorem 6. The chromatic number of the projective plane is 6, the chromatic number of the Klein bottle is 6, and the chromatic number of the torus is 7.

Stäckel[1] solved the colorability problem for three-dimensional space, and thus for all spaces of higher dimension, when he showed that the chromatic number of E^3 is infinity. One may construct a three-dimensional figure with chromatic number n as follows. Place n long blocks numbered from 1 to n side by side. On top of them place another layer of such blocks at right angles to the first, as shown in Figure 11. Considering the two blocks labeled i as a single object, we now have n mutually adjacent objects.

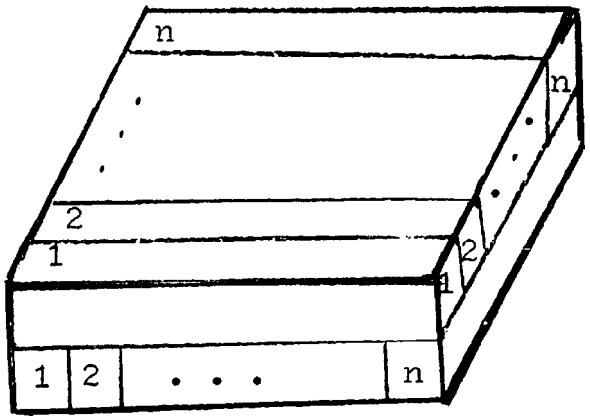


Figure 11. Three-dimensional figure with chromatic number n

---

[17]Franklin, 80-81.

[18]P.Stäckel, Uber nachbargebiete im Raum, Zeitschrift fur Math. und Physik,42(1897), 275-276.

III. MENGERIAN THEOREMS

The diversity of the origins and applications of graph theory within mathematics can be demonstrated by considering a basic theorem on the connectivity of graphs. First proved by the topologist K. Menger, this theorem has variations which include Hall's theorem on systems of distinct representatives and Ford and Fulkerson's max-flow min-cut theorem; several of these will be discussed in this chapter. Most of them can be translated without difficulty into statements about bipartite graphs, and therefore it is this setting in which we discuss Menger's theorem and its variants. Although the many versions of this theorem are not all equivalent in the strictest sense, they are very closely related in that the methods used to prove one will suffice to prove all.

Menger's theorem is stated in terms of the <u>connectivity</u> $\kappa(G)$ of the graph G where $\kappa(G)$ is the minimum number of points whose removal results in a disconnected graph (or the graph K_1). A graph G is said to be <u>n-connected</u> if $\kappa(G) \geq n$. A set of points whose removal, together with the lines terminating at these points, results in a disconnected graph (or K_1) is called a <u>cut set</u> of points. In Figure 1 we show graphs with connectivity 1, 2, and 3, with cut sets

of 1, 2, and 3 points indicated.

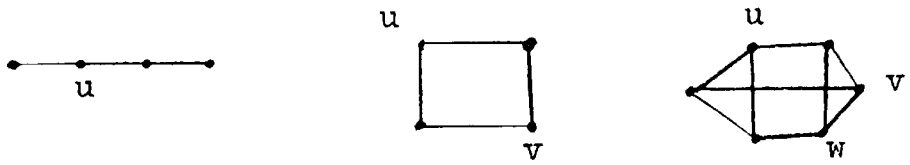


Figure 1. Graphs with cut sets indicated

Menger[1] proved a theorem in 1927 which relates the connectivity of a graph with the number of mutually disjoint paths (that is, distinct paths having in common no points other than, possibly, their initial or terminal points). We state it below in its original form before giving a graphical statement of it.

Theorem 1. If K is a compact one-dimensional space which is n-connected between the finite sets P and Q, then K includes n pairwise disjoint paths each of which joins a point of P and a point of Q.

The theorem holds, of course, when P and Q each contain a single point, and it is this weakened version that is of importance to graph theory.

Theorem 2. The maximum number of disjoint paths connecting two nonadjacent points u and v in a graph G is equal to the minimum number of points whose removal separates u and v.

As an illustration of this theorem we consider the graph in Figure 2 which contains three disjoint paths from

---

[1]K.Menger, Zur Allgemeinen Kurventheorie, Fund. Math. 10(1927), 96-115.

u to v and a minimal cut set {x,y,z}.

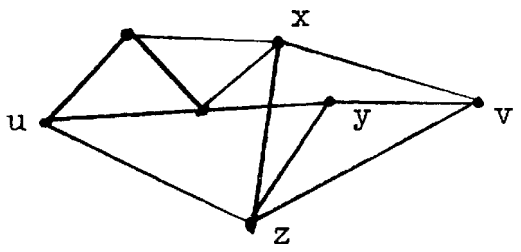


Figure 2. Graph with minimal cut set {x,y,z}

Whitney[2] extended Theorem 2 to describe the overall connectivity of a graph as follows.

Theorem 3. The graph G is n-connected if and only if there are at least n disjoint paths joining every pair of points of G.

The first specifically graphical statement of Menger's theorem was given for bipartite graphs by König[3] in 1931. To state König's theorem, we need the concept of a covering of the lines of G. A set S of points of a graph G is said to cover the lines of G if every line has at least one of its endpoints in S. Clearly, if the sets P and Q are the two color sets of a bipartite graph G, then a subset of points whose removal separates P and Q must cover the lines of G. König's theorem is given in terms of a minimal covering set.

---

[2]H.Whitney, Congruent graphs and the connectivity of graphs, Amer. J. Math. 54(1932), 150-168.

[3]D. König, Graphok es matrixok, Matematikai es Fizikai Lapok 38(1931), 116-119.

Theorem 4. The minimum number of points which cover the lines of a bipartite graph G is equal to the maximum number of lines of G which are pairwise disjoint.

König's proof[4] of this theorem which we now give, serves as a model for the proofs of all of the variations of Menger's theorem. Since G is bipartite, its point set is partitioned into two sets U and V such that every line of G joins a point of U with a point of V. Let the minimum number of points in a covering set be n and suppose that $X = \{(u_1,v_1), (u_2,v_2), \ldots, (u_k,v_k)\}$ is a maximal set of pairwise disjoint lines of G. It is clear that $k \leq n$, and we now show that $k = n$ by exhibiting k points whose removal separates U and V. Define $U' = \{u_i \mid i=1,2,\ldots,k\}$ and $V' = \{v_i \mid i=1,2,\ldots,k\}$ and call a path $w_1w_2\ldots w_{2n}$ in G an X-path if the lines $(w_2,w_3), (w_4,w_5), \ldots, (w_{2n-2},w_{2n-1})$ are in X. A point in U-U' cannot be joined to a point in V-V' by an X-path for if W were such a path, then one could remove the second, fourth, ..., (2n-2)th lines of W from the set X and replace them with the first, third,...,(2n-1)th lines of W. But this would increase the number of lines in X, contradicting its maximality.

Now construct the subset $R = \{r_1,r_2,\ldots,r_k\}$ of $U' \cup V'$ as follows. Take $r_i = v_i$ if v_i is joined to a point in U-U' by an X-path, and take $r_i = u_i$ otherwise. We show that R covers the lines of G by proving that every line has an endpoint in R.

---

[4] D.König, Theorie der Endlichen und Unendlichen Graphen, Academische Verlagsgesellschaft M.B.H., Leipzig, 1936, 232-233.

Case 1. If the line (p,q) has p in U-U', q in V-V', then (p,q) is disjoint from every line in X, contradicting the maximality of X.

Case 2. If (p,q) has p in U-U', q in V', then q is in R, since the line (p,q) itself gives an X-path from q to a point in U-U'.

Case 3. If (p,q) has p in U', q in V-V', then either $p=u_i$ is in R or there is an X-path joining v_i with a point in U-U'. In the latter case we could construct an X-path from U-U' to V-V' by adding the lines (u_i,v_i) and (u_i,q) to the end of the path from U-U' to v_i. Since this is impossible, we have that p is in R.

Case 4. If (p,q) has $p=u_i$ in U' and $q=v_j$ in V', then either u_i is in R or v_i is in R. If v_i is in R, then there is an X-path from U-U' to v_i and we can extend this to an X-path from U-U' to v_j by adding the lines (v_i,u_i) and (u_i,v_j). Thus either $p=u_i$ is in R or $q=v_j$ is in R.

This completes König's argument.

To show that Menger's theorem immediately implies König's theorem we simply note that a set of points which covers the lines of a bipartite graph is equivalent to a set of points whose removal separates the two color classes of points. For example, in the bipartite graph of Figure 3, the points u,v,w,x cover the lines while the lines (p,u), (q,v), (s,w), and (t,x) form a maximal set of pairwise disjoint lines.

Konig's theorem does not directly imply Menger's theorem, but the latter can be proved by an argument very similar

to that just demonstrated for bipartite graphs. In order to reduce the general case to a bipartite argument, we start with the graph G and two disjoint sets of points, P and Q, which are n-connected, and construct the graph G* by removing as many lines as possible from G without destroying the n-connectedness of P with Q. In G* the points of P and Q each induce a totally disconnected subgraph, and each remaining point of G* that is not an isolate lies on a single path from P to Q. If G* has no points other than those of P, Q, and isolates, then it is bipartite with color classes P and Q and König's theorem applies directly. Otherwise G* is "essentially" bipartite relative to P and Q and a straightforward argument completes the proof.

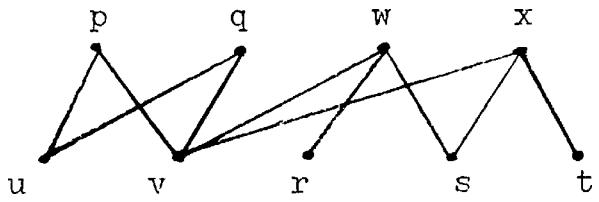


Figure 3. Graph with covering set {u,v,w,x}

A well-known form of Menger's theorem in the context of matrices was also discovered by König.[5] It is given here for (0,1)-matrices, although "nonzero entry" may be substituted for "one" without altering the truth of the assertion. We say that two 1's of a matrix are <u>independent</u> if they do not share the same row or column.

---

[5]D.König, <u>Theorie der Endlichen und Unendlichen Graphen</u>, 240.

Theorem 5. The minimum number of rows and columns which contain all the 1's of a (0,1)-matrix is equal to the maximal number of pairwise independent 1's.

The relationship between this theorem and König's theorem is immediately apparent when we associate with the matrix M a bipartite graph G whose points u_i represent the rows of M and whose points v_i represent the columns of M. Then the line (u_i,v_j) is in G if the (i,j)-entry of M is 1. Thus the matrix which would give rise to the bipartite graph of Figure 3 is

$$M = \begin{array}{c} \\ u \\ v \\ r \\ s \\ t \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} p & q & w & x \end{array} \\ \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array} .$$

The rows corresponding to u and v together with the columns corresponding to w and x cover all the 1's, while the (p,u), (q,v), (s,x), and (t,x) entries form a maximal set of independent 1's.

Undoubtedly the most celebrated of the Mengerian theorems is Philip Hall's theorem on systems of distinct representatives. A system of distinct representatives (SDR) of a collection $\{T_1, T_2, \ldots, T_m\}$ of subsets of a finite set S is a set $\{s_1, s_2, \ldots, s_m\}$ of distinct elements of S such that s_i is in T_i for i=1,2,...,m. Hall's theorem[6] gives a necessary and sufficient condition for the existence

---

[6] P.Hall, On representatives of subsets, J.London Math. Soc. 10(1935), 26-30.

of an SDR.

Theorem 6. A collection $T = \{T_1, T_2, \ldots, T_m\}$ of subsets of a finite set S has an SDR if and only if, for each $k = 1,2,\ldots,m$, the union of any k sets in T has at least k elements.

The similarity between Hall's theorem and König's theorem is not immediately evident; however, each can be proved directly from the other. To prove Hall's theorem, we construct a bipartite graph G in which the points u_i represent the elements s_i of S and the points v_i represent the sets T_i. The line (u_i, v_j) is in G whenever s_i is in T_j. An SDR corresponds to a set of m disjoint lines in G and König's theorem then says that an SDR exists only if the minimum number of points whose removal separates U from V is also m. To show that the condition of Hall's theorem implies that a minimal cut set has m points, suppose that there is a set of k points whose removal separates U from V, where k_1 of these are in U, k_2 in V, and $k < m$. Then the $m - k_2$ sets T_j represented by the remaining points in V have at most those k_1 elements of S in common. But $k_1 = k - k_2$ which is less than $m - k_2$, since $k < m$, thus there are $m - k_2$ sets T_j having less than $m - k_2$ elements in common, contradicting the hypothesis of Hall's theorem. Therfore $k < m$, and the sufficiency condition is proved; the necessity is obvious. The converse construction proves König's theorem from Hall's.

It would be of interest to be able to calculate the

number of different SDR's of a collection R of subsets of S. A lower bound, given in the next theorem[7] can be obtained by induction on the number of subsets T_i.

Theorem 7. Let $T = \{T_1, T_2, \ldots, T_m\}$ be a collection of subsets of a set S such that $|T_i| \geq t$ for each i. If $t \leq m$, then T has at least t! SDR's and if $t > m$, then T has at least $\frac{t!}{(t-m)!}$ SDR's.

Hall's theorem can be used to determine when disjoint subsets X_i of preassigned cardinality p_i can be found to represent the sets T_i by applying it to the family of sets consisting of p_i copies of T_i, for $1 \leq i \leq m$. This extension of the theorem was first stated by Halmos and Vaughan.[8]

Theorem 8. Let $\{T_1, T_2, \ldots T_m\}$ be a family of subsets of S and let $p_1, p_2, \ldots, p_m$ be positive integers. Then there exist pairwise disjoint sets X_i such that $X_i \subset T_i$ and $|X_i| < p_i$ if and only if

$$|T_{i_1} \cup T_{i_2} \cup \cdots \cup T_{i_k}| \geq p_{i_1} + p_{i_2} + \cdots + p_{i_k}$$

whenever $1 \leq k \leq m$ and $1 \leq i_1 < \cdots < i_k \leq m$.

Dulmage and Mendelsohn[9] have extended the matrix form of Hall's theorem by introducing a criterion similar to Hall's condition. To make the connection clear, we first restate Hall's theorem in matrix form.

---

[7] H.J.Ryser, Combinatorial Mathematics, Carus Mathematical Monographs, 14, Mathematical Association of America, Buffalo, 52-53.

[8] P.Halmos and H.E.Vaughan, The marriage problem, Amer. J. Math. 72(1950), 214-215.

[9] L.Dulmage and N.S.Mendelsohn, Some generalizations of the problem of distinct representatives, Canad. J. Math. 10 (1958), 320-241.

Theorem 6'. The (0,1)-matrix A possesses a set of independent 1's incident with all the columns if and only if each set of k columns contains 1's in at least k rows.

Dulmage and Mendelsohn's theorem is concerned with a set of independent 1's which are incident with a given subset of the rows and columns of a (0,1)-matrix. The case r=0, s=n is Hall's theorem.

Theorem 9. Let A be an m x n matrix, $0 \leq r \leq m$, $0 \leq s \leq n$. Let R be a set of r rows, C a set of s columns in A. Then A possesses a set of independent 1's incident with both R and C if and only if both of the following conditions are satisfied:

1) The 1's of each subset R* of R belong to at least $|R^*|$ columns.
2) The 1's of each subset C* of C belong to at least $|C^*|$ rows.

The above theorem can also be stated in terms of lines instead of points; to do so we need the concepts of the line-connectivity of a graph and line-disjoint paths. The line-connectivity $\lambda(G)$ of a graph G is the minimum number of lines whose removal results in a disconnected graph. Two paths are line-disjoint if they have no lines in common. The line-form of Menger's theorem, first published by Elias, Feinstein, and Shannon,[10] is given in Theorem 10.

---

[10] P.Elias, A.Feinstein, and C.Shannon, A note on maximum flow through a network, IEEE Trans. Information theory 2(1956), 117-119.

Theorem 10. The maximum number of line-disjoint paths joining two points u and v in a graph G is the minimum number of lines whose removal separates u from v.

The line-analogue of Theorem, first stated by Whitney[11], characterizes the line-connectivity of a graph in terms of its line-disjoint paths.

Theorem 11. The graph G is n-line connected if and only if any pair of points of G is joined by at least n line-disjoint paths.

Theorem 10 is trivial when the graph G is bipartite, since the number of line-disjoint paths from U to V and the number of lines whose removal separates U from V both are simply the number of lines in G. The point-form of the bipartite theorem is, however, a consequence of a very general line-form of this theorem, which we discuss next.

In linear programming, Ford and Fulkerson solved a problem of network flows which directly implies Hall's and König's theorems. A network is a directed graph with a source s and sink t in which a positive real number (or ∞), called the capacity $c(x,y)$, is assigned to each arc (x,y). A positive real number $f(x,y)$, called the flow of (x,y), is also assigned to each arc in such a way that the following two conditions are satisfied:

1) $f(x,y) \leq c(x,y)$.

2) The sum of the flows of the arcs into a point v equals the sum of the flows of the arcs out of v for $v \neq s,t$.

---

[11] H.Whitney, Congruent graphs and the connectivity of graphs, 150-168.

The flow from s equals the flow to t, and this common value is called the flow value of the network. Finally, the capacity of a minimal cut set is the sum of the capacities of the arcs in the set. We can now state the max-flow min-cut theorem of Ford and Fulkerson.[12]

Theorem 12. For any network, the maximal flow value from s to t is equal to the minimal cut capacity of all cut sets separating s and t.

A network is shown in Figure 4 with the flow and capacity of each arc indicated by the ordered pair of numbers on the arc. The flow value of this network is 7 and the minimal cut capacity is 7, and thus the flow value shown is maximal. It can be shown that whenever the capacity function is integral-valued, then there is a maximal flow function which is also integral-valued.

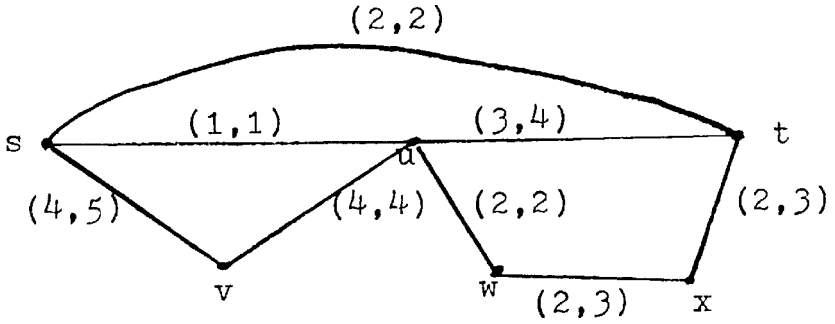


Figure 4. Network with flow value 7

---

[12] L.R.Ford, Jr., and D.R.Fulkerson, Maximal flow through a network, Canad. J. Math. 8(1956), 399-404.

To derive König's bipartite theorem from the max-flow min-cut theorem, we take a bipartite graph G with color sets U and V and orient all the lines from U to V. We then adjoin a point s and a point t to G together with arcs from s to each point of U and arcs from each point of V to t. (See Figure 5.) By assigning unit capacity to the new arcs and infinite capacity to the old arcs, we insure that every minimal cut set of arcs will include only new arcs. Therefore, a cut set S* of arcs of the new digraph G* corresponds to a cut set S of points of G by taking the one endpoint in G of each arc in S*. If f is any integral maximal flow, the arcs of G having flow value 1 are pairwise disjoint. The max-flow min-cut theorem says that the number of such arcs equals the number of arcs in a minimal cut set, and this implies König's theorem.

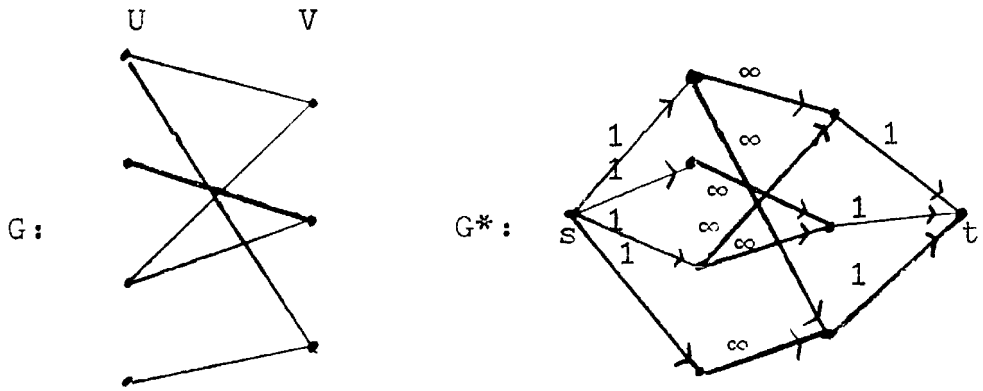


Figure 5. Bipartite graph and its associated network

Philip Hall's original proof of the existence of a system of distinct representatives was of a nonconstructive nature. An algorithm for actually constructing a maximal

set of independent 1's or, equivalently, a system of distinct representatives of a collection of subsets has been given more recently by several people. We give Ford and Fulkerson's procedure[13] for the matrix case, and show how it may be used to obtain an SDR. Given a (0,1)-matrix, we select any set of independent 1's in the matrix and designate its members by 1*. We may assume that the 1's come from the first k rows of the matrix. To determine whether this set is maximal, label the rows and columns in the following manner. Label every row containing none of the 1*'s with a dash. Then select any labeled row R_i and scan it for 1's, labeling the (unlabeled) columns with the number i. Repeat until all labeled rows have been scanned. Then scan the labeled columns C_j for 1*'s, labeling the previously unlabeled rows in which they lie with the number j. Continue by scanning rows for 1's and columns for 1*'s until no more labels can be assigned or until a column containing no 1* is labeled. In the first case, the original set of independent 1's is maximal, since the 1's in every row fall in columns which have 1*'s. In the latter case, we are able to construct a larger set as follows. Take the 1 in the j-th row of the column which has just been labeled j and put it in S. In the row in which this 1 lies, proceed to the position indicated by the label of the row. There is a 1* in this position. Replace it by the 1 in that column

---

[13] L.R.Ford, Jr., and D.R.Fulkerson, Flows in Networks, Princeton University Press, Princeton, 1962, 55-57.

in the position indicated by the column's label. In the row in which this new 1* lies, proceed to the position indicated by the row's label and continue as before. This procedure will end when one of the rows labeled with a dash is reached. A larger set of independent 1's is thus obtained. A diagram of this procedure for a specified matrix is given in Figure 6; the members of the original independent set are designated by 1* and the new members of the larger independent set by 1**.

	1	2	3	4	5	6	7	
1	1*	0	1	0	1	1**	1	1
2	0	1*	0	1	1	0	0	2
3	1	0	0	0	1*	1	0	5
4	0	1	0	0	1	0	1*	7
5	0	0	1*	0	0	1	0	3
6	1**	1	1	0	0	0	0	-
7	0	0	0	0	1	0	1	-
	6	6		7	1	7		

Figure 6. Construction of a maximal set of independent 1's.

In the context of a collection of subsets $\{T_1, T_2, \ldots, T_m\}$ of the set S satisfying Hall's condition, this algorithm becomes a method for obtaining an SDR. We first select a representative from one set at a time, x_i in T_i, continuing as long as it can be done, keeping the representatives distinct. We may assume that the sets with representatives are $T_1, \ldots, T_k$. If $k = m$, then an SDR has been obtained. Otherwise consider the sets $T_{k+1}, \ldots, T_m$ which have no representatives. Their elements are all contained in the set $X = \{x_1, x_2, \ldots, x_k\}$. Examine the sets

T_i whose representatives are in $T_{k+1} \cup \cdots \cup T_m$. We may assume that they are $T_1, T_2, \ldots, T_{k_1}$ (or we could relabel them). If any of these sets, say T_j, has an element s which is not in X, then we may use s to represent T_j, freeing x_j to represent one of $T_{k+1}, \ldots, T_m$. If not, then $(T_1 \cup T_2 \cup \cdots \cup T_{k_1}) \subset X$. Examine the sets among T_{k_1+1}, $\ldots, T_k$ whose representatives occur in the set $T_1 \cup T_2 \cup \cdots \cup T_{k_1}$. There must be some, since otherwise $|T_1 \cup \cdots \cup T_{k_1} \cup T_{k+1} \cup \cdots \cup T_m| = k_1$, contradicting the hypothesis of Hall's theorem. We may assume that these sets are $T_{k_1+1}, \ldots, T_{k_2}$. If any of these sets, say T_j, has an element s not in X, then s may be used to represent T_j and x_j to represent some T_i, where $1 \leq i \leq k_1$, thereby freeing x_i to represent one of $T_{k+1}, \ldots, T_m$. Hall's condition guarantees that this process cannot terminate before a larger set of representatives is obtained.

As an illustration, consider the subsets $T_1=\{1,2,3,5\}$, $T_2=\{2,3,4\}$, $T_3=\{1,3,5\}$, $T_4=\{1,2\}$, and $T_5=\{1\}$ of the set $S = I_6$. Taking 1 in T_1, 2 in T_2, and 3 in T_3, there is no element remaining to represent T_4 or T_5. The set X of the algorithm is $\{1,2,3\}$. The sets T_1 and T_2 have their representatives in $T_4 \cup T_5$. $T_2 - X = \{4\}$, so we may use 4 to represent T_2, freeing the element 2 to represent T_4. Beginning again, $X = \{1,2,3,4\}$. T_1 is the only set having its representative in T_5. Since $T_1 - X = \{5\}$, we let 5 represent T_1 and free the element 1 to represent T_5.

The Ford-Fulkerson theorem is a statement concerning

directed graphs; Menger's theorem can also be stated for directed graphs. Dirac[14] was the first to give a directed form of Menger's theorem.

Theorem 13. The maximum number of (arc-)disjoint paths from u to v in a digraph D equals the minimum number of points (arcs) whose removal separates u from v.

We mention one more variation of Menger's theorem, also stated in terms of digraphs, which was discovered by Dilworth[15] in lattice theory in 1950, with the first correspondence between it and the max-flow min-cut theorem being given in 1956 by Fulkerson.[16] We say that a subset S of elements in a partially ordered set is independent if for every pair of elements u and v in S, (u,v) is not in the relation. Alternately, we say that the elements of an independent set are pairwise incomparable. A chain in a partially ordered set is a path, and a set-sum of chains is the subgraph induced by the points in the chains. Dilworth's theorem relates the number of independent elements of a partially ordered set to the number of disjoint chains.

Theorem 14. Let every set of k + 1 elements of a partially ordered set P be dependent, while at least one set of k elements is independent. Then P is a set-sum of k disjoint chains.

---

[14]G.A.Dirac, Extensions of Menger's theorem, J.London Math. Soc. 38(1963), 148-161.

[15]R.P.Dilworth, A decomposition theorem for partially sets, Ann. of Math. 51(1950), 161-166.

[16]D.R.Fulkerson, On Dilworth's theorem, Proc. Amer. Math. Soc. 7(1956), 701-702.

In max-min terminology, this asserts that the maximum number of pairwise incomparable elements of a partially ordered set P is equal to the minimum number of disjoint paths which span P.

For a survey of results concerning the infinite case of Menger's theorem, the reader is referred to a study by Mirsky and Perfect.[17]

---

[17] L.Mirsky and H.Perfect, Systems of representatives, J.Math. Anal. Appl. 15(1966), 520-268.

IV. GRAPHICAL METHODS IN MATRIX THEORY

We have already seen that matrices are an important tool in the study of graphs; for example, matrices have been used to record adjacencies between points, to recognize acyclic digraphs, and to compute the number of spanning trees of a graph. Graphs, in turn, can be used to simplify proofs and computations in matrix theory. In this chapter we discuss graphical methods related to the problems of finding determinants, eigenvalues, and matrix inverses, solving a particular matrix equation, and proving a rather abstract matrix theorem.

A relagraph R(M) can be associated with an arbitrary n x n matrix $M = ((m_{ij}))$ by taking n points $v_1, v_2, \ldots, v_n$ and the arc (v_i, v_j) whenever $m_{ij} \neq 0$. The adjacency matrix of the resulting relagraph is the (0,1)-matrix obtained from M by reducing every nonzero entry to 1. It is convenient, however, to "weight" the arc (v_i, v_j) with the number m_{ij} and thus deal only with the matrix M. An example is shown in Figure 1.

$$M = \begin{pmatrix} 1 & 0 & -2 & 1 \\ 0 & 0 & 4 & 0 \\ 0 & 4 & 0 & 0 \\ 3 & 0 & 2 & -1 \end{pmatrix}$$

R(M):

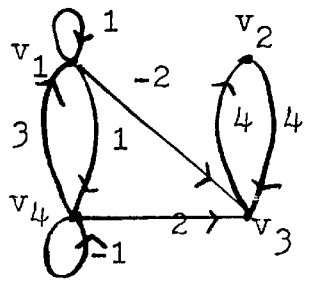


Figure 1. Matrix and associated relagraph

There is a straightforward graphical procedure[1] for simplifying the process of finding the determinant, eigenvalues, and inverse of a matrix whose relagraph has more than one strong component. We first relabel the points of the relagraph with respect to its strong components, and note that this effects a permutation of the rows and columns of the matrix M, that is, M is replaced by PMP^t, which has the same eigenvalues and determinant as M. Moreover, since the strong components of a relagraph are partially ordered, the points may be relabeled in such a way that

$$PMP^t = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ * & & & A_n \end{pmatrix},$$

where A_i is the (weighted) adjacency matrix of the i-th strong component. The eigenvalues of this matrix are the eigenvalues of the smaller matrices on the diagonal, and the determinant is, up to sign, the product of the determinants of the A_i. Applying this method to the matrix M in Figure 1, we obtain

$$PMP^t = \begin{pmatrix} 0 & 4 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 0 & -2 & 1 & 1 \\ 0 & 2 & 3 & -1 \end{pmatrix}$$

---

[1] F. Harary, A graph theoretic method for the complete reduction of a matrix with a view toward finding its eigenvalues, J. Math. Physics 38(1959), 104-111.

with

$$A_1 = \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix} \qquad \text{and} \qquad A_2 = \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} .$$

Since $|A_1| = -16$ and $|A_2| = -4$, we know that $|M| = \pm 64$. The eigenvalues of A_1 are ± 4 and the eigenvalues of A_2 are ± 2, thus M has the four eigenvalues ± 4, ± 2.

The block lower-triangular form also simplifies the process of calculating the inverse of a matrix. If M can be written in the form

$$\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix} ,$$

then the inverse of M can be found by computing the inverses of the smaller matrices A_{11} and A_{22}. It is easily verified that

$$M^{-1} = \begin{pmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1} \cdot A_{21} \cdot A_{11}^{-1} & A_{22}^{-1} \end{pmatrix} .$$

Applying this formula to the matrix of Figure 1, we find that

$$A_{11}^{-1} = \begin{pmatrix} 0 & 1/4 \\ 1/4 & 0 \end{pmatrix} \text{ and } \quad A_{22}^{-1} = \begin{pmatrix} 1/4 & 1/4 \\ 3/4 & -1/4 \end{pmatrix} ,$$

and therefore that

$$(PMP^t)^{-1} = \begin{pmatrix} 0 & 1/4 & 0 & 0 \\ 1/4 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 1/4 \\ 1/2 & 0 & 3/4 & -1/4 \end{pmatrix} .$$

Thus, by permuting the rows and columns of PMP^t back to their original positions, we obtain

$$M^{-1} = \begin{pmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & 0 \\ 0 & 1/4 & 0 & 0 \\ 3/4 & 1/2 & 0 & -1/4 \end{pmatrix} .$$

Another graphical restatement of the definition of determinant gives a method[2] of computing the determinant $|M|$ by means of the cycles of $R(M)$. Each term in $|M|$ is of the form

$$\operatorname{sgn}\sigma \prod_{i=1}^{n} m_{i\,\sigma i} \quad ;$$

for definiteness, we consider the term $-m_{12}m_{25}m_{34}m_{43}m_{51}$ which may be written as the product $-m_{12}m_{25}m_{51}\cdot m_{34}m_{43}$ of two cycles. If the term is nonzero, then $v_1v_2v_5v_1$ and $v_3v_4v_3$ are cycles of $R(M)$. In general, all of the nonzero terms in the expansion of the determinant correspond to cycles of $R(M)$.

Theorem 1. A term in the expansion of the determinant of a matrix M is nonzero if and only if the arcs of $R(M)$ corresponding to the entries in this term constitute a spanning collection of directed cycles.

A spanning collection of disjoint directed cycles is called a linear subgraph, since each point v has $id(v) = od(v) = 1$. We note that $\operatorname{sgn}(\sigma)$ is equal to the number of even cycles of σ, since each odd cycle contributes an even number of transpositions to the factorization of the permutation into transpositions. The task of finding the determinant is now reduced to the following problem.

---

[2] F. Harary, The determinant of the adjacency matrix of a graph, SIAM Rev. 4(1962), 202-210.

Theorem 2. Let n be the number of distinct linear subgraphs D_i of D, p_i the product of the numbers m_{ij} corresponding to the arcs in D_i, and e_i the number of even cycles in D_i. Then

$$|M| = \prod_{i=1}^{n} (-1)^{e_i} p_i \quad .$$

Corollary 2a. If M is a (0,1)-matrix, then

$$|M| = \prod_{i=1}^{n} (-1)^{e_i} \quad .$$

Referring once more to the matrix in Figure 1, we note that there are two linear subgraphs, the first consisting of the two 2-cycles $v_1v_4v_1$ and $v_2v_3v_2$, the second consisting of the 2-cycle $v_2v_3v_2$ together with the loops at v_1 and v_4. Thus $|M| = (-1)^2 \cdot 1 \cdot 3 \cdot 4 \cdot 4 + (-1)^2(-1)(4 \cdot 4) = 64$.

The relationship of the characteristic equation and eigenvalues of a (0,1)-matrix A to the properties of the relagraph associated with the matrix have been examined in detail by Collatz and Sinogowitz.[3] They first observed the following correspondence between the coefficients of the characteristic polynomial of A and the number of certain distinguished subgraphs of R(A).

Theorem 3. Let A be a (0,1)-matrix and write

$$|xI - A| = \sum_{i=0}^{n} a_i x^{n-i} \quad .$$

When A is the adjacency matrix of a relagraph R, then $-a_1$ is the number of loops of R, and further, if R is the graph G, $a_1 = 0$, $-a_2$ is the number of lines, and $-(1/2)a_3$ is the

---

[3] L. Collatz and U. Sinogowitz, Spektren endlichen Graphen, Abh. Math. Sem. Univ. Hamburg 21(1957), 63-77.

number of triangles of G.

To justify the claims of Theorem 3, we first note that the coefficient a_1 is the negative of the trace of A and thus $-a_1$ is the number of loops of R. Now considering only graphs, the coefficient a_2 is equal to

$$- \sum_{j<i} \begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$$

and $a_{ii} = a_{jj} = 0$. The product $a_{ij}a_{ji}$ is 1 if the line (v_i, v_j) is in G and 0 otherwise, whence $-a_2$ is the number of lines of G. Similarly, the coefficient a_3 is obtained by adding the negatives of the determinants of the 3 x 3 principal submatrices of A. By Theorems 1 and 2 these determinants are nonzero if and only if the induced graph is a triangle, and in such cases the determinant is 2. Hence $-(1/2)a_3$ is the number of triangles in the graph. Theorems 1 and 2 can be applied further to obtain geometrical interpretations for a_4, a_5, etc., but the interpretations are of lesser interest. For example, the coefficient a_4 is the sum of the 4 x 4 determinants on the main diagonal. Terms in such determinants are positive if the induced graph has two disjoint lines and negative if the induced graph has a 4-cycle. Thus $-a_4 = f_{22} - 2s_4$, where f_{22} is the number of pairs of disjoint lines and s_4 the number of 4-cycles. Some examples are shown in Figure 2; the first four lead to the observation that the characteristic polynomial of K_p is $[x - (p-1)](x+1)^{p-1}$.

Collatz and Sinogowitz also found bounds for the set

of eigenvalues in terms of graphical properties. Theorem 4 summarizes these results, and the eigenvalues of the graphs in Figure 2 illustrate them.

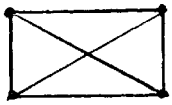
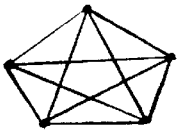
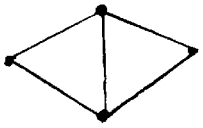
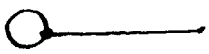
	characteristic polynomial	eigenvalues
	$x^2 - 1$	1, -1
	$x^3 - 3x - 2$	-1, -1, 2
	x^4-6x^2-8x-3	3, -1
	$x^5-10x^3-20x^2-15x-4$	4, -1
	$x^3 - 2x$	0, $\pm\sqrt{2}$
	x^4-5x^2-4x	0, -1, $\frac{1\pm\sqrt{17}}{2}$
	$x^2 - x - 1$	

Figure 2. Characteristic polynomials and eigenvalues of graphs

Theorem 4. If r is the largest eigenvalue of the connected graph G, then all of the eigenvalues of G fall in the interval $[-r,+r]$. Furthermore, letting d_{max} denote the maximum degree of the graph, d_{min} the minimum and $\approx{d}$ the average degree, then $d_{min} \leq \approx{d} \leq r \leq d_{max}$ with $\approx{d} = r$ if and only if G is regular.

The Cayley-Hamilton theorem states that any nonsingular matrix satisfies its characteristic equation $|A - xI| = 0$.

Hoffman[4] has examined the problem of finding all graphs whose adjacency matrices satisfy the equation $P(A) = J$, for a given polynomial P. Since the entries of A^m give the number of walks of length m from one point to another, the problem is to find, given a positive integer n and rational coefficients $a_0, a_1, \ldots, a_k$, all graphs with n points satisfying the equation

$$a_0 p_0(i,j) + a_1 p_1(i,j) + \cdots + a_k p_k(i,j) = 1$$

for all pairs of points v_i, v_j, where $p_m(i,j)$ is the number of walks of length m from v_i to v_j. Theorem 5 limits the range of admissible graphs to those which are regular and connected.[5]

Theorem 5. Let A be a symmetric $(0,1)$-matrix with zeros on its diagonal. Then there exists a polynomial P such that $J = P(A)$ if and only if A is the adjacency matrix of a regular, connected graph G.

It is easy to show that G must be regular, for if $J = P(A)$, then A commutes with J. Therefore, the (i,j)-entry of AJ, which is the degree of v_i, equals the (i,j)-entry of JA, which is the degree of v_j. Furthermore, since $J = P(A)$, the (i,j)-entry of some power of A must be nonzero, and thus G must be connected. The construction of the polynomial P for a connected graph G which is regular of

---

[4] A.J.Hoffman, On the polynomial of a graph, Amer. Math. Monthly 70(1963), 30-36.

[5] Hoffman and M.H.McAndrew have also characterized the digraphs which satisfy such an equation in The polynomial of a directed graph, Proc. Amer. Math. Soc. 10(1965), 303-309.

degree d makes use of the eigenvalues of A(G). Theorem 4 states that d is one eigenvalue; letting $d, e_1, \ldots, e_k$ be the distinct eigenvalues of A, the polynomial P is defined to be

$$P(x) = \frac{n \prod_{i=1}^{k} (x-e_i)}{\prod_{i=1}^{k} (d-e_i)} .$$

For the graph shown in Figure 3, the adjacency matrix A is

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

the eigenvalues are 0, 2, -1, and the polynomial P is $\frac{x(x+2)}{2}$.

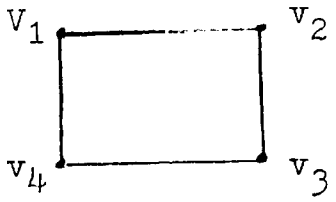


Figure 3. Graph with polynomial $\frac{x(x+2)}{2}$

The polynomial P associated with a graph G can be used to determine whether G is bipartite. If the points of a bipartite graph G are labeled in such a way that $v_1, v_2, \ldots, v_r$ are in one color set and $v_{r+1}, \ldots, v_p$ are in the other color set, then

$$A(G) = \begin{pmatrix} 0 & B \\ B^t & 0 \end{pmatrix} .$$

If G is regular, then the matrix B is square. It is clear that the vector (u;-u) where every component of u is 1, is

an eigenvector of A corresponding to the eigenvalue -d. Using this fact, we arrive at the following characterization of regular bipartite graphs.

Theorem 6. If G is a regular connected graph of degree d, and P(x) is the polynomial of G, then G is bipartite if and only if P(-d) = 0.

To complete this chapter we show how graph theoretical methods have been applied to simplify considerably the proof of a significant matrix theorem. The theorem, first proved by Amitsur and Levitski[6], is given below.

Theorem 7. If $A_1, A_2, \ldots, A_k$ are any n x n matrices with entries in a commutative ring, and $[A_1, \ldots, A_k]$ is defined to be $\sum_{\sigma \in S_k} \operatorname{sgn}(\sigma) A_{\sigma_1} \cdots A_{\sigma_k}$, then 2n is the smallest value of k for which $[A_1, \ldots, A_k]$ is always zero.

Amitsur and Levitski's proof is an involved induction argument on the number of idempotent A_1. A later proof by Kostant[7] makes use of a theorem on the primitive cohomology classes of Lie algebras. In 1963, however, Swan[8] published an elementary proof based on the following theorem. An eulerian walk of a directed graph D is a spanning walk (not necessarily closed) which contains every arc of D exactly once. If the arcs of D are labeled $x_1, x_2, \ldots, x_q$, and σ is

---

[6] A.S.Amitsur and J.Levitski, Minimal identities for algebras, Proc. Amer. Math. Soc. 1(1950), 449-463.

[7] B.Kostant, A theorem of Frobenius, a theorem of Amitsur and Levitski, and cohomology theory, J. Math. Mech. 7(1958), 237-264.

[8] R.G.Swan, An application of graph theory to algebra, Proc. Amer. Math. Soc. 14(1963), 367-373.

the permutation in S_q given by the successive indices of an eulerian walk W, we define $\epsilon(W)$ to be sgn σ.

Theorem 8. Let D be a digraph (possibly with loops or multiple arcs) with $q \leq 2p$, whose arcs have been labeled $x_1, x_2, \ldots, x_q$, and let u and v be any fixed points of D (not necessarily distinct). Then the number of eulerian walks W beginning at u and ending at v with $\epsilon(W) = 1$ equals the number with $\epsilon(W) = -1$.

To show that Theorem 8 implies Theorem 7, we let $e_{i,j}$ denote the n x n matrix with 1 in the (i,j)th position and 0 elsewhere, and we call $e_{i,j}$ a matrix unit. Then $e_{i,j} \cdot e_{k,\ell} = 0$ if $j \neq k$ and $e_{i,\ell}$ if $j = k$. Clearly any n x n matrix is a linear combination of such matrix units, and since $[A_1, \ldots, A_{2n}]$ is multilinear, it is sufficient to prove Theorem 7 for the case where all the A_i are matrix units. We now associate a digraph D (possibly with loops or multiple arcs) with the matrices $A_1, \ldots, A_{2n}$. D has n points $v_1, \ldots, v_n$, and the arc $x_i = (v_j, v_k)$ wherever $A_i = e_{j,k}$; clearly $q = 2n$. Hence the product $A_1 \cdots A_{2n}$ has a nonzero (i,j)-entry if and only if the corresponding sequence $e_{\sigma_1} \cdots e_{\sigma_{2n}}$ is an eulerian walk. Theorem 8 guarantees that these eulerian walks occur in pairs which cancel each other out in the sum $\Sigma \operatorname{sgn}(\sigma) A_{\sigma_1} \cdots A_{\sigma_{2n}}$. To show that 2n is the smallest value of k for which $[A_1, \ldots, A_k]$ is always zero, we need only consider the digraph D_1 of Figure 4, in which $q < 2p$. This digraph is associated with the matrices $A_1 = e_{1,2}$, $A_2 = e_{2,1}$, and $A_3 = e_{2,2}$, and

it is easy to verify that $[A_1,A_2,A_3] \neq 0$. The digraph D_2 shows how D_1 can be extended to the general case.

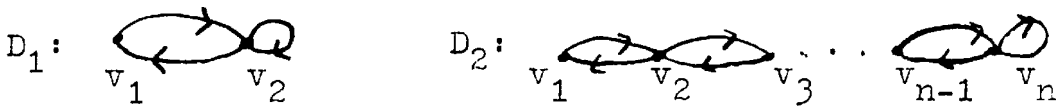


Figure 4. Directed graph with $[A_1,A_2,A_3]\neq 0$

We now justify our claim that Swan's proof is an elementary one, and not just a complicated graphical replacement of the matrix argument, by sketching a proof of Theorem 8. We first show that, without loss of generality, we may consider only connected digraphs in which $q = 2p$ and such t that every point v has $id(v) = od(v)$. First, if D is disconnected, it has no eulerian walks. Second, if $q > 2p$, then attaching a path with $k = q - 2p$ new points and arcs to the digraph D as in Figure 5, we obtain a digraph D' in which $q = 2p$, and the eulerian walks from u in D are in 1-1 correspondence with the eulerian walks from u' in D'.

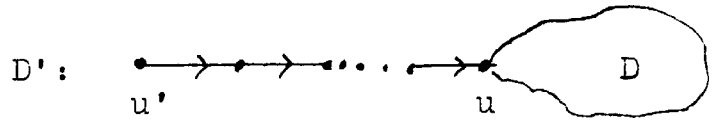


Figure 5. Construction of a digraph in which $q = 2p$

If D contains two arcs x and y from u to v, then the transposition (x,y) converts a walk W with $\epsilon(W) = 1$ into one with $\epsilon(W) = -1$. Therefore, any digraph with multiple arcs satisfies the theorem. Finally, we show that it is sufficient to consider only those digraphs in which $id(v) = od(v)$ for all points v. In order for an eulerian walk from

u to v to exist, either u = v and every point has id = od, or u ≠ v and od(u) - id(u) = 1 and id(v) - od(v) = 1. In the latter case we adjoin the point r, together with the arcs (v,r) and (r,u) to the digraph D and obtain a digraph D' in which id = od for all points, and the eulerian walks from u to v in D correspond to those from r to r in D'.

Now assuming that q = 2p and id(v) = od(v) for each v in D, we proceed by induction on p. If p = 1, the D is a digraph with a single point and two loops, and we have already shown that the theorem is true for digraphs with multiple arcs. For p > 1, the argument breaks down into three cases.

Case 1. D contains the subgraph shown in Figure 6.

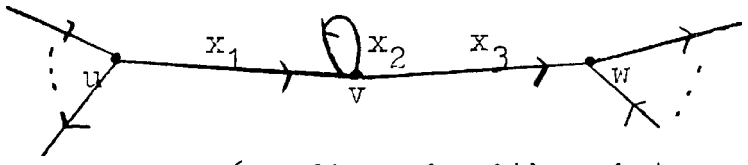


Figure 6. Digraph with point v with three incident arcs

If the initial point of the eulerian walk is v, then every walk must either begin or end with x_2, and there is a 1-1 correspondence between the two types of walks, one with $\epsilon = 1$ and the other with $\epsilon = -1$. If the initial point is not v, then any eulerian walk must contain the arcs x_1, x_2, x_3 in that order. We then suppress the point v, and replace the three arcs x_1, x_2, x_3 with the single arc (u,w); this preserves $\epsilon(W)$, and hence the induction hypothesis applies.

Case 2. D has a point v with id(v) = od(v) = 1. Define the order of a point u to be id(u) + od(u). Since

$q > p$, there must be at least one point with order > 2. Let u and v be adjacent points with order (u) > 2, and order (v) $= 2$, as in Figure 7.

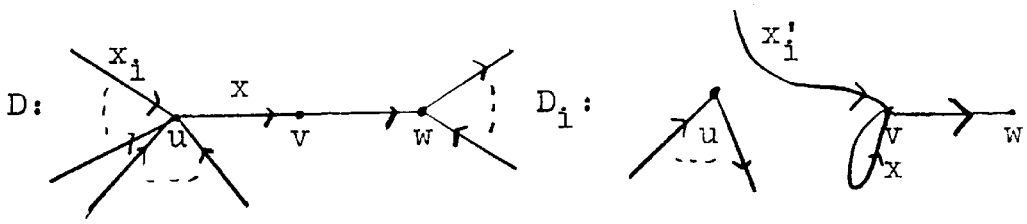


Figure 7. Construction of digraph D_i

For each arc x_i entering u we create a new digraph D_i by letting x_i go to v without passing through u, and by converting x into a loop at v, as shown in Figure 7. On any eulerian walk of D, one of the x_i must precede x, and consequently, every eulerian walk on D_i is eulerian on D, and hence Case 2 reduces to Case 1.

Case 3. All points have order ≥ 2. The average order of the points is $\frac{2q}{p} = 4$, since each arc has two endpoints. Therefore all points must have order 4 and D contains the subgraph shown in Figure 8. We then proceed to construct the digraph D_i as in Case 2. Again every eulerian walk in D is eulerian on exactly one of D_1 and D_2.

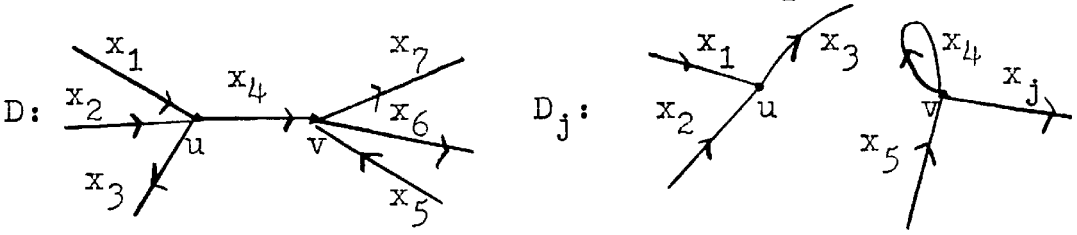


Figure 8. Construction of digraph D_j

However, some of the eulerian walks in D_i which enter v by x_i' then leave v without going around the loop, and these are not eulerian on D. However they all contain a subwalk

$x_5x_4x_j$ where $j = 6$ or 7. Therefore, they are exactly those walks which are eulerian on one of the digraphs D_j, shown in Figure 8. The eulerian walks on the D_i, $i = 1,2$ are exactly those eulerian on D or D_j and since Case 2 applies to D_i and Case 1 to D_j, the theorem holds for D.

V. GROUPS AND GRAPHS

A group, being a set with a binary operation, lends itself easily to a graphical representation; conversely, a graph, as a geometrical configuration, has a group of automorphisms. In this chapter we discuss the interaction between graphs and groups. We begin by defining the Cayley diagram of a group and examining the graphical interpretation of certain group properties. Then a graphical question on groups--which groups have a planar Cayley diagram?--is raised and answered. Secondly, we describe the automorphism group of a graph and study the question of characterizing those groups which are isomorphic to the automorphism group of some graph.

1. Cayley diagrams

Cayley[1] developed a graphical representation for a group of order n which consisted of a complete symmetric digraph with n points. The points correspond to the group elements $g_1, g_2, \ldots, g_n$, and the arc (g_i, g_j) is labeled $g_i^{-1}g_j$. A simpler version of this representation includes only those arcs corresponding to generators of the group. A Cayley

[1] A. Cayley, The theory of groups: a graphical representation, Amer. J. Math. 1(1879), 174-176.

diagram[2] of a group Γ of order n with generators $g_1, g_2, \ldots, g_s$ is a directed graph with n points, $g_1, g_2, \ldots, g_n$, corresponding to the group elements. At each point there is an arc corresponding to each generator and to the inverse of each generator; that is, for each g in Γ and each generator g_i, the arcs (g, gg_i) and (gg_i^{-1}, g) are in the Cayley diagram $D(\Gamma)$. In Figure 1 the Cayley diagrams of S_3, with generators (12) and (123), and A_4, with generators (12)(34) and (123), are shown. When a generator is involutory, a single undirected line is used. Note that the point representing I may be arbitrarily designated. The arcs are colored, in keeping with Cayley's original representation, to indicate which generator is being applied. This is not to be confused with a line-coloring, as defined in Chapter II.

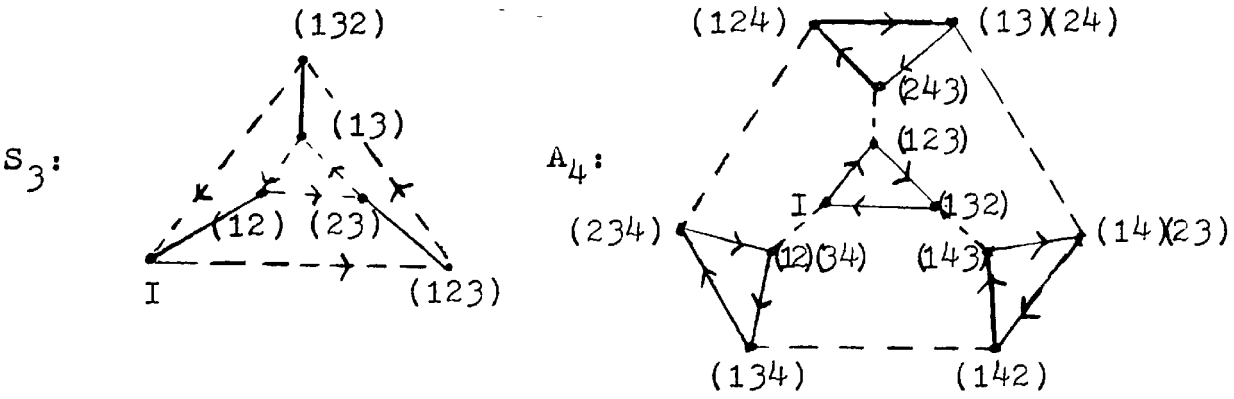


Figure 1. Cayley diagrams of S_3 and A_4

A product of elements in Γ corresponds to a walk in D, the elements in the product designating the sequence of arcs to be taken from a point v. Furthermore, if a walk from I to v consists of the sequence $a_1, a_2, \ldots, a_k$, where the

---

[2]Also called "color group" or "Dehnsche Gruppenbild".

a_i are generators or inverses of generators, then v represents the element $a_1 a_2 \cdots a_k$. A product which equals the identity element corresponds to a closed walk, starting from any point. Whether or not a given group is commutative can be determined by examining the walks of length 2 from I; if the walks g_i, g_j and g_j, g_i lead to the same point for each pair (i,j), then Γ is commutative. The solvability of the equation $rx = s$ for all r and s in a group Γ implies that $D(\Gamma)$ is strongly connected, since the solution x represents a walk from r to s. This fact can also be verified by observing that every element can be written as a product of the generators and their inverses and thus lies on some directed cycle containing the identity element.

An element of a generating set is <u>redundant</u> if it can be written as a product of the other generators; that is, if the remaining generators form a generating set for the group. A graphical criterion for redundancy can also be given.

<u>Theorem 1</u>. A generator g is redundant if and only if the deletion of all arcs labeled g in the Cayley diagram leaves a strongly connected digraph.

If the arcs corresponding to one generator, say g_1, are removed, the result is either a strong digraph or a collection of isomorphic disjoint subgraphs, each representing the subgroup of Γ with generators $g_2, \ldots, g_s$. In particular, if the involutory generators are deleted from the Cayley diagrams in Figure 1, the remaining subgroups are

all isomorphic to $H = \langle(123)\rangle$, which is normal in S_3 but not normal in A_4. A graphical criterion for the normality of such a subgroup is given in Theorem 2.

Theorem 2. Let Γ be a group with generators $g_1, g_2, \ldots, g_s$, and H a subgroup of Γ with generators $g_2, g_3, \ldots, g_s$. Let $D(\Gamma)$ be the Cayley diagram of Γ and $C_1, C_2, \ldots, C_k$, the weak components of the digraph $D_{g_1}(\Gamma)$ obtained from $D(\Gamma)$ by deleting the arcs corresponding to g_1. Then H is normal in Γ if the deleted arcs from any given component C_i all lead to a single other component C_j.

We may verify this assertion by considering any element r of Γ and any element g in the component containing the identity, say C_1, and showing that rgr^{-1} is in C_1. The element r may be written $r = a_1^{e_1} a_2^{e_2} \cdots a_m^{e_m}$ where a_i is a generator of Γ and $e_i = \pm 1$. If g_1 occurs in r exactly w times with $e_1 = +1$ and v times with $e_1 = -1$, then the path corresponding to r leads from the identity (in C_1) through w+v other components, and ends in $C_{1+w=v}$. The path corresponding to g leads to another point in $C_{1+w=v}$ and the path corresponding to r^{-1} leads back to a point in C_1. Therefore if g is in C_1, then rgr^{-1} is in C_1, and H is normal.

With this graphical description of the normal subgroup H, the Cayley diagram of the factor group Γ/H is easy to describe; the elements of the factor group are the components of $D_{g_1}(\Gamma)$. After these components have each been shrunk to a point, the arcs corresponding to the generator g_1 can be restored unambiguously by Theorem 2 and the resulting

configuration is the Cayley diagram of Γ/H. In Figure 2 the Cayley diagram of S_4, generated by (12), (123), (12)(34), and (13)(24), is shown. We assign the colors red (heavy) to (12), green (dashed) to (123), blue (light) to (12)(34), and brown (dotted) to (13)(24). The existence of a red-brown-red-brown path from I to (12)(34) shows that the generator (12)(34) is redundant, (12)(34)

$$(12)(34) = (12)\cdot(13)(24)\cdot(12)\cdot(13)(24).$$

If the arcs corresponding to (123) are deleted, then the conditions of Theorem 2 are satisfied, and hence the subgroup H generated by (12), (12)(34), and (13)(24) is normal in S_4. Taking S_4/H, we obtain the Cayley diagram for C_3, shown in Figure 3.

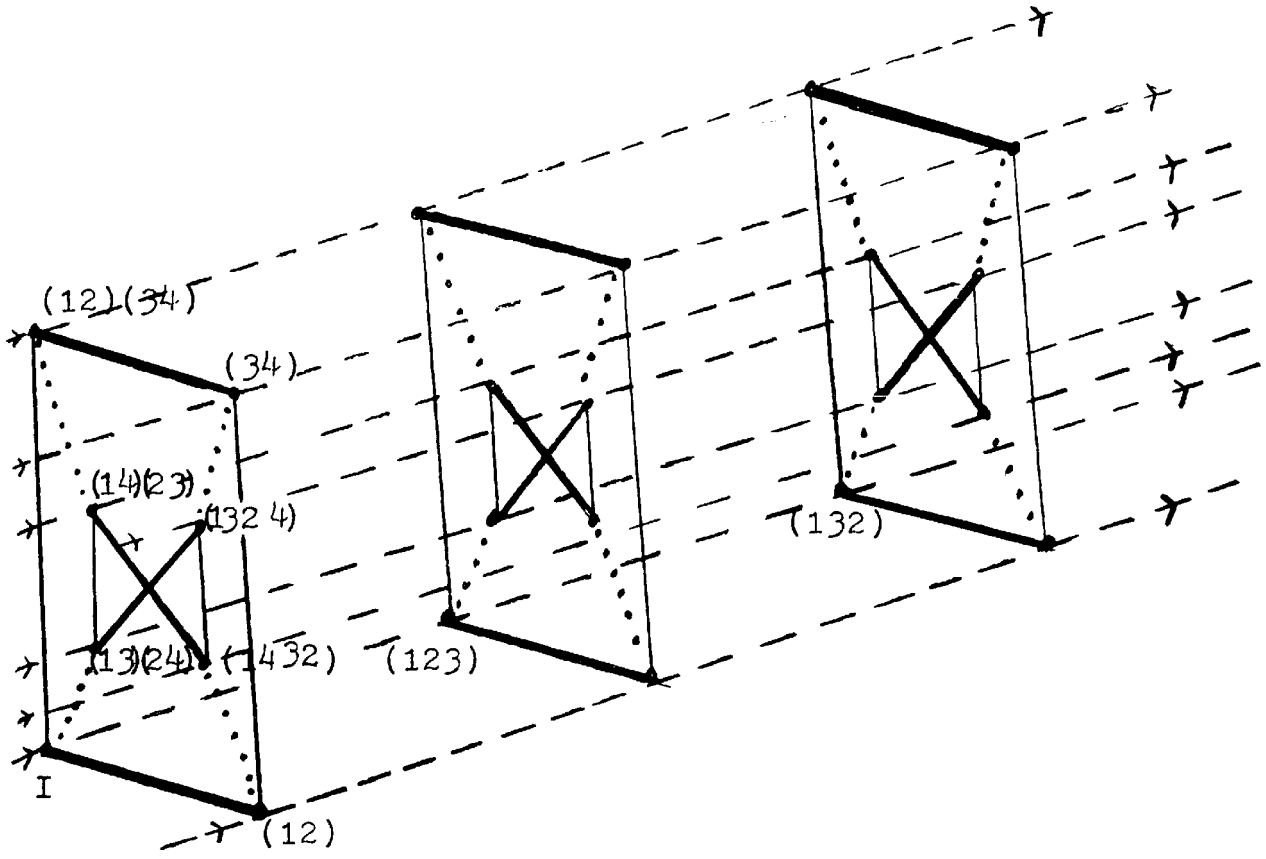


Figure 2. Cayley diagram of S_4



Figure 3. Cayley diagram for C_3

Turning now to a graphical question about groups, we note that the Cayley diagram of S_3 is planar, while that of S_4 is not. The problem of characterizing planar Cayley diagrams is a problem which was studied and settled in 1896 by Maschke[3], who related the planar Cayley diagrams to rotation groups of regular polyhedra. Considering first the one-color Cayley diagrams, we observe that if a group Γ has only one generator, then Γ is cyclic, and $D(\Gamma)$ is a directed cycle of a single color. Thus all one-color Cayley diagrams are planar.

At each point in a two-color Cayley diagram there is an outgoing arc and an incoming arc corresponding to each generator of order > 2; a generator of order 2 is represented by a single line at each point. Furthermore, if $w=a_1a_2\cdots a_k$ describes, in terms of its arcs, a closed walk from v to v, then w describes a similar closed walk from any point u to itself. Two-color plane Cayley diagrams also satisfy the properties in the next two lemmas.

Lemma 1. In a plane two-color Cayley diagram, the arcs of one color at a point do not separate the arcs of the other color.

---

[3]W. Maschke, The representation of finite groups, Amer. J. Math. 16(1896), 156-194.

To show this, we assume the two colors are red and green and that at some point u the red arcs, representing the generator r, are separated by the green arcs, representing the generator g. Then the cycles $r^i = I$, $g^j = I$, at u must intersect at a second point v. (See Figure 4.) A third cycle at u, going from u to v on the green cycle and back to u on the red cycle, is given by the relation $g^m r^n = I$. If u_1 is the first point following u on the green cycle and v_1 the first point following v on the green cycle, then g^m describes a walk from u_1 to v_1. Furthermore, r^n must then describe a path from v_1 back to u_1. This path cannot intersect the original red cycle, because there can be no additional red arcs to the points on the original red cycle. Applying the Jordan Curve Theorem, one of u_1 and v_1 is inside the original red cycle and the other is outside it; a contradiction is thus obtained and the lemma is proved. (If d(u,v) or d(v,u) = 1, in which case $u_1 = v$ or $v_1 = u$, a similar proof can be devised.)

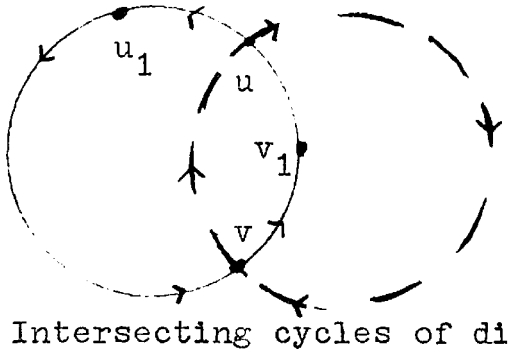


Figure 4. Intersecting cycles of different colors

A similar application of the Jordan Curve Theorem yields the second lemma.

Lemma 2. Any green cycle or red cycle in a plane two-color Cayley diagram D contains either all the points of D or none of them in its interior.

This lemma implies that every one-color cycle bounds a face of the Cayley diagram (including the exterior region as a face). In addition to these cycles there are also cycles of two colors, called intermediate cycles, which bound faces. The arcs of an intermediate cycle are alternately red and green, for if there were two consecutive red arcs at a point u, then there would be three cycles at u, one red, one green, one intermediate. The first and last of these would share two common arcs at u. This is impossible since none of these cycles can contain any points of the other two in its interior. We have thus proved part of the following characterization of intermediate cycles.

Lemma 3. The arcs of an intermediate cycle are alternately red and green, and thus every intermediate cycle has an even number of arcs; the intermediate cycle represents the relation $(rg)^k = I$. Furthermore, every cycle whose arcs are alternately red and green is an intermediate cycle.

Applying Euler's formula $V - E + F = 2$ to a plane two-color Cayley diagram with N points and 2N arcs, we obtain $F = N + 2$. If the red cycles have length i, the green cycles length j, and the intermediate cycles length 2k, there are then N/i, N/j, and N/k of each, and hence the equation

$$N + 2 = N/i + N/j + N/k$$

must be satisfied. The possible solutions to this equation are given in Table 1.

	N	i	j	k
1	2j	2	2	j
2	6	3	2	2
3	12	3	2	3
4	12	3	3	2
5	24	4	2	3
6	24	3	2	4
7	24	4	3	2
8	60	5	2	3
9	60	3	2	5
10	60	5	3	2

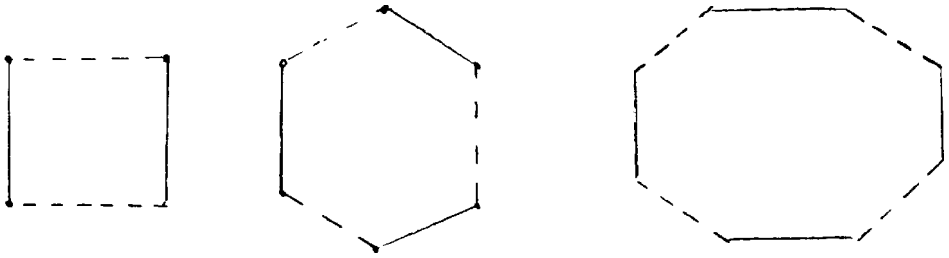
Table 1. Cycle lengths of two-color Cayley diagrams with N points

The Cayley diagrams are shown in Figures 5 - 8. Note that all but the first represent the rotation groups of the regular polyhedra. The second corresponds to the triangle, the third and fourth to the tetrahedron (under two different sets of generators), the fifth, sixth, and seventh to the cube and octahedron, and the eighth, ninth, and tenth to the dodecahedron and icosahedron.

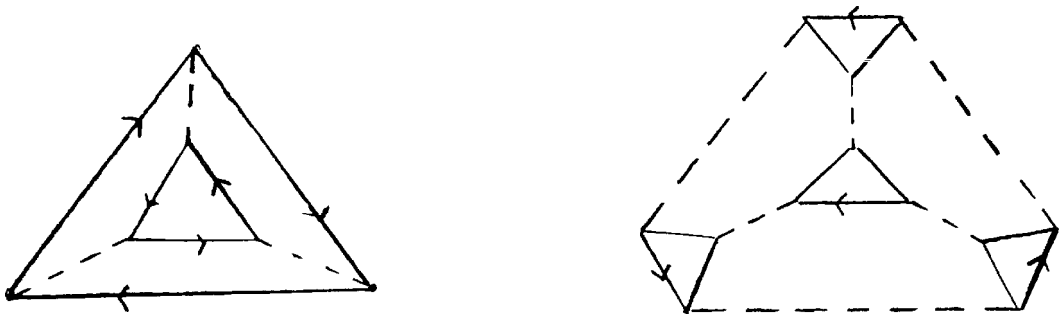
Plane Cayley diagrams with three independent generators have an additional property which simplifies the characterization problem.

Lemma 4. Each of the generators of a plane three-color Cayley diagram has order 2.

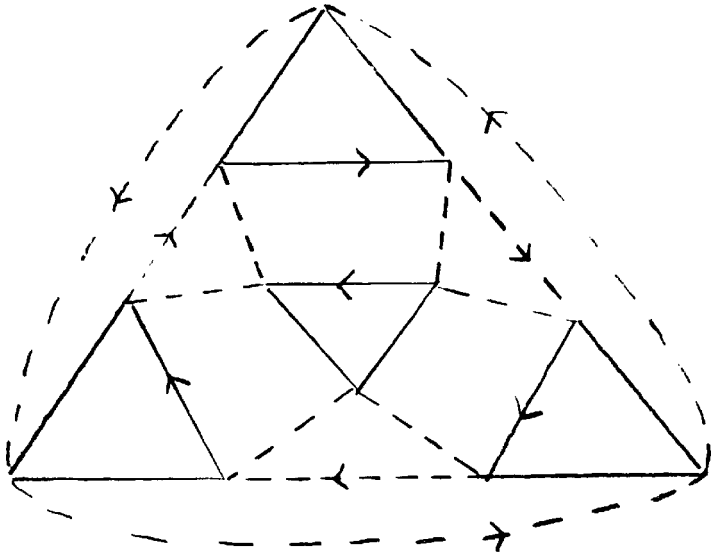
To show this, we note that if the arcs corresponding to one generator, say the blue arcs, are removed, then each component of the resulting diagram is a two-color plane



$r^2=s^2=(rs)^2=I$ $(rs)^3=I$ $(rs)^4=I$

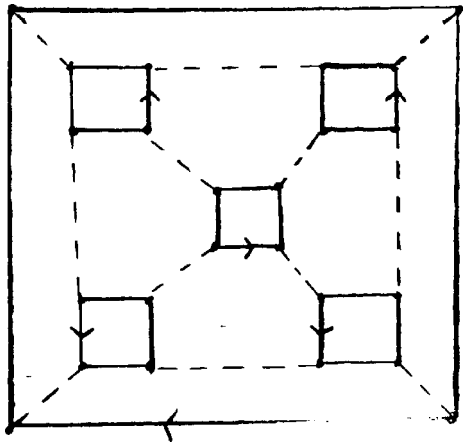


$r^2=s^3=(rs)^2=I$ $r^2=s^3=(rs)^3=I$

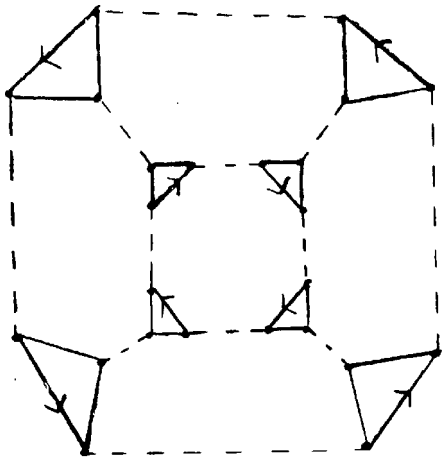


$s^3=r^3=(rs)^2=I$

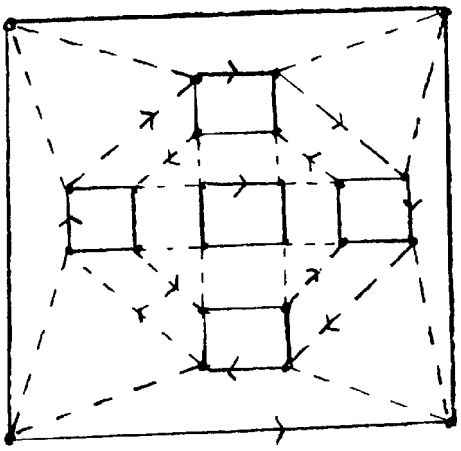
Figure 5. Plane two-color Cayley diagrams with 2j, 6, and 12 points



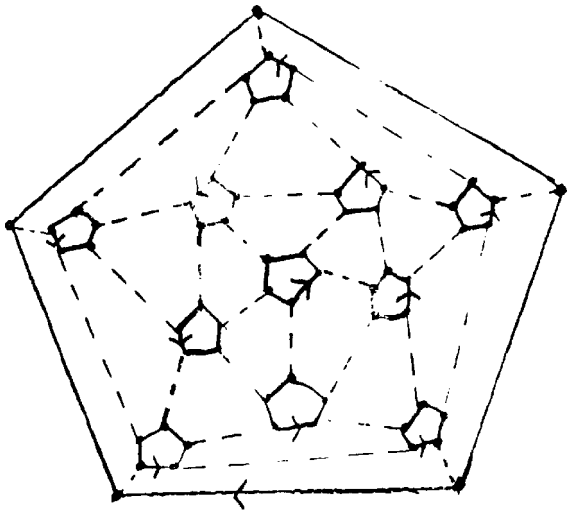
$s^4=r^2=(rs)^3=I$



$s^3=r^2=(rs)^4=I$

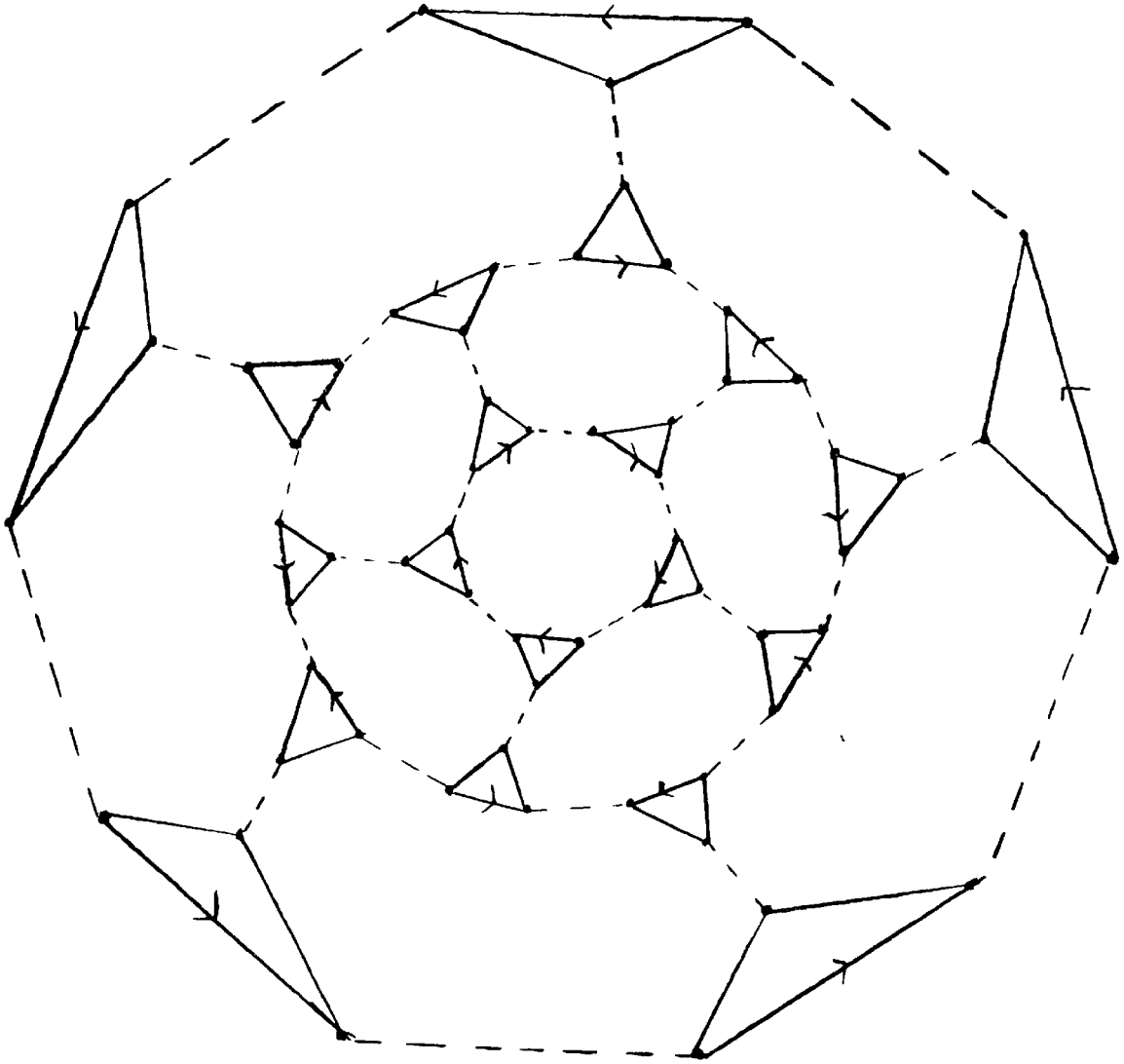


$s^4=r^3=(rs)^2=I$



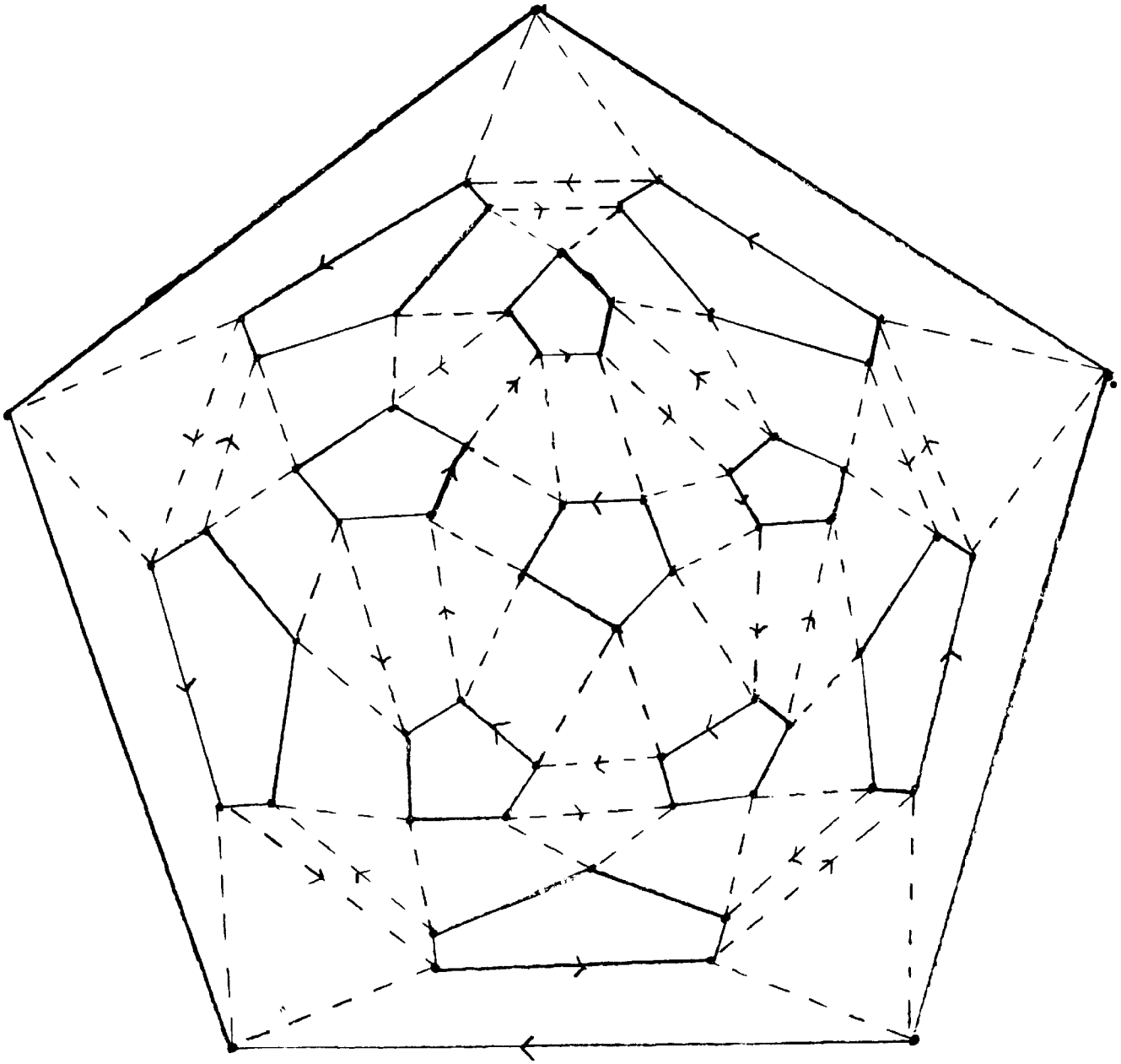
$s^5=r^2=(rs)^3=I$

Figure 6. Plane two-color Cayley diagrams with 24 and 60 points



$$s^3 = r^2 = (rs)^5 = I$$

Figure 7. Plane two-color Cayley diagram with N = 60, i = 3, j = 2, k = 5



$$s^5 = r^3 = (rs)^2 = I$$

Figure 8. Plane two-color Cayley diagram with N = 60, i = 5, j = 3, k = 2

diagram. Letting these components be, for example, the first tetrahedral diagram in Figure 5, we now attempt to restore the blue arcs. The blue arcs from the points of the center triangle must lead to another component C_2 (or the generator g_3 would be redundant), and in order to preserve planarity, the component C_2 must lie in the interior of this triangle. The component C_2 also has a center triangle, and the blue arcs restored to the points of this center triangle must lead to a component C_3 in its interior. Since there are a finite number of components, it is impossible to restore all of the blue arcs in a planar fashion; thus no three-color Cayley diagram can have the tetrahedral diagram for its components when the arcs corresponding to one generator are removed. The same argument applies to components isomorphic to any of the plane two-color Cayley diagrams except the first, in which the generators all have degree 2.

The intermediate cycles of a three-color plane Cayley diagram have arcs of two alternating colors, as in the two-color diagrams. Restating Lemma 3 for three-color diagrams, we have the following.

Lemma 5. All of the intermediate cycles of a three-color plane Cayley diagram represent relations of the form $(rs)^n = I$.

Now applying Euler's formula to a three-color plane Cayley diagram with N points (and therefore 3N/2 lines), we obtain $F = 2 + N/2$. Since the generators are of order two, there are no one-color cycles, and Lemma 5 implies that

there are no three-color intermediate cycles. The three types of intermediate cycles represent the relations $(rb)^i = (rg)^j = (bg)^k = I$, and there are N/2i, N/2j, and N/2k of them, respectively. Therefore, the three-color plane Cayley diagrams correspond to the solutions of

$$2 + N/2 = N/2i + N/2j + N/2k \quad ,$$

which are given in Table 2. The corresponding diagrams are shown in Figures 9 and 10.

N	i	j	k
4k	2	2	k
24	2	3	3
48	2	3	4
120	2	3	5

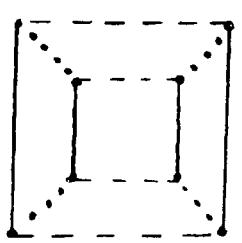
Table 2. Cycle lengths of plane three-color Cayley diagrams with N points

In attempting to describe the four color plane Cayley diagrams, we proceed as before by deleting all the arcs corresponding to one generator. The components of the resulting diagram must be three-color Cayley diagrams. Referring to Figures 9 and 10, however, we note that every three-color plane Cayley diagram has a cycle in its interior. The argument used in Lemma 4 then proves that there is no four-color plane Cayley diagram, and hence no plane Cayley diagram with more than three colors.

A study of the Cayley diagrams of binary relations in general has been done by Artzy[4], and his results give

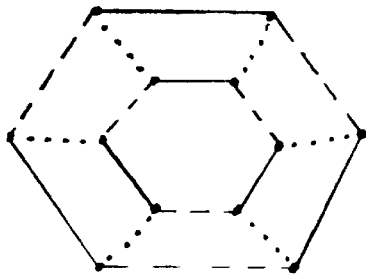
---

[4] R. Artzy, Cayley diagrams of binary systems, Duke Math. J. 28(1961), 491-495.



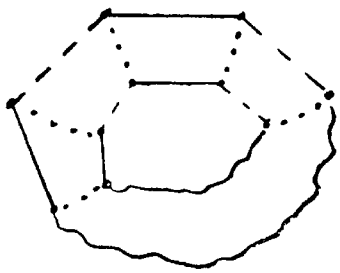
$b^2 = r^2 = g^2 = I$

$(rb)^2=(rg)^2=(bg)^2=I$



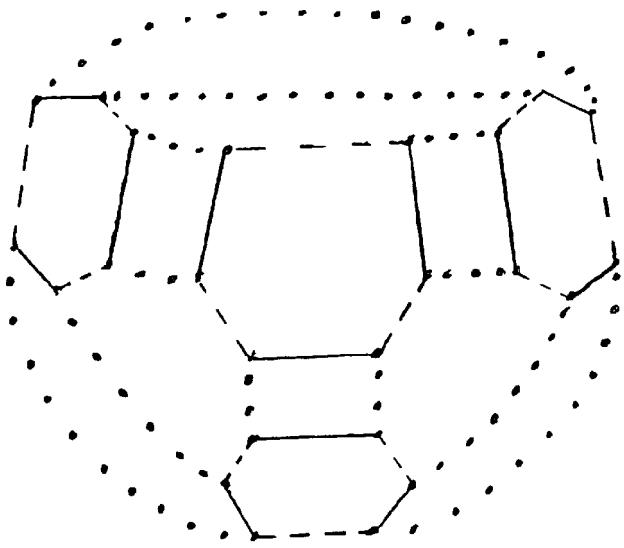
$b^2=r^2=g^2=I$

$(rb)^3=(rg)^2=(bg)^2=I$



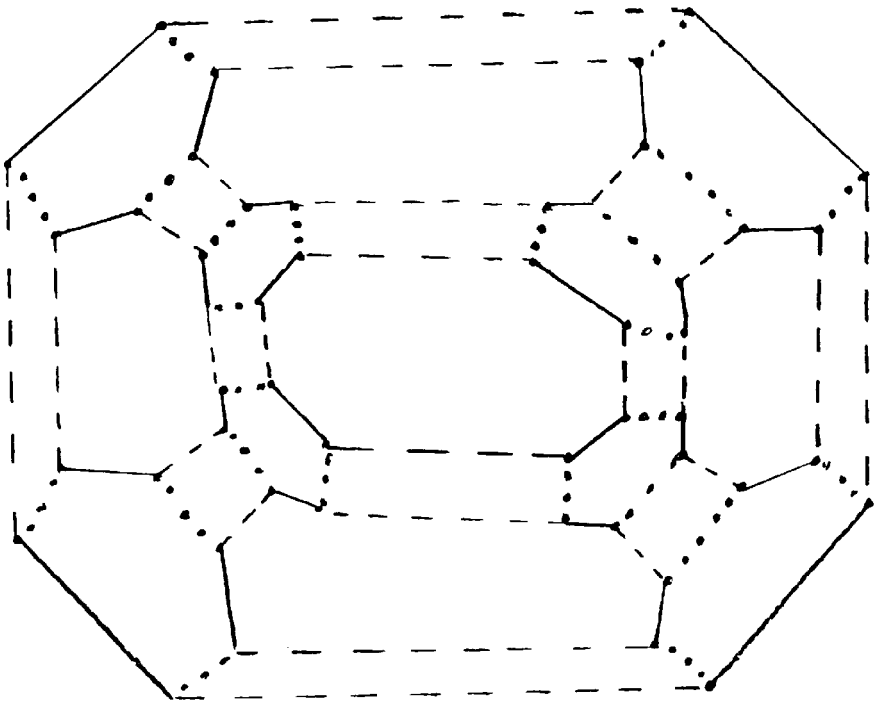
$b^2=r^2=g^2=I$

$(rb)^n=(rg)^2=(bg)^2=I$

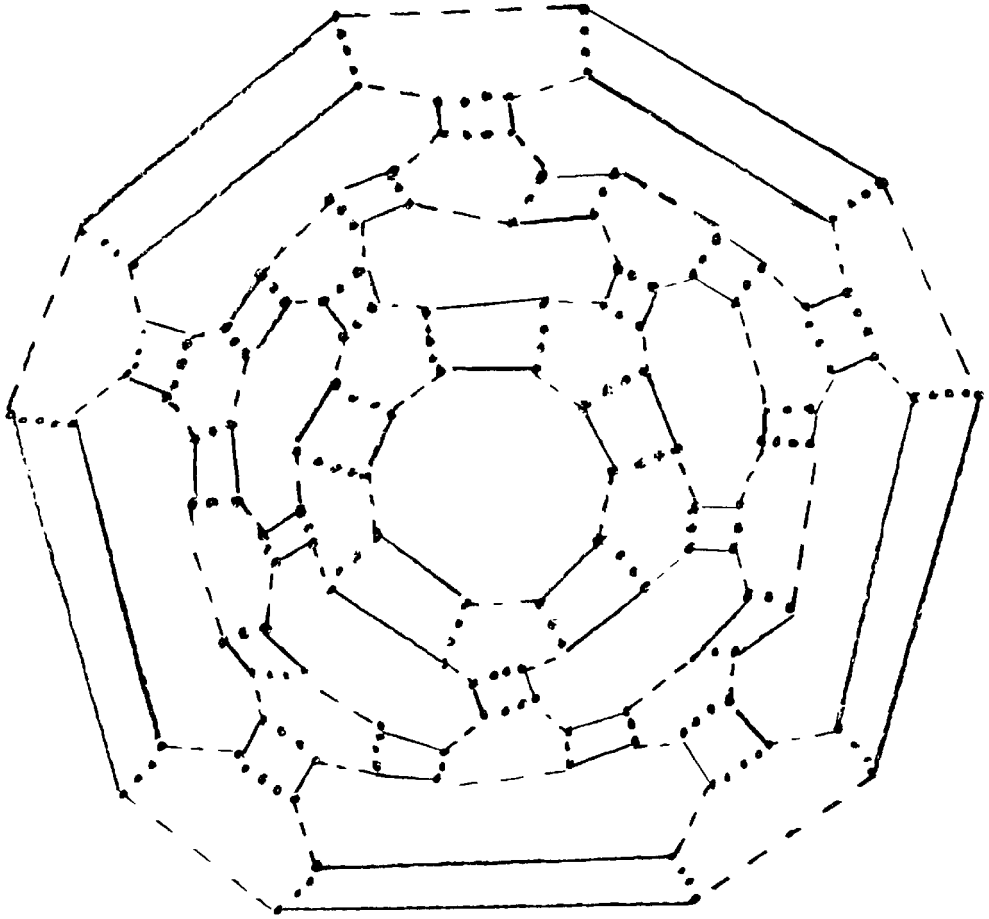


$r^2=g^2=b^2=(rg)^2=(rb)^3=(bg)^3=I$

Figure 9. Plane three-color Cayley diagrams with 4k and 24 points



$r^2=b^2=g^2=(bg)^2=(rg)^3=(br)^4=I$



$r^2=b^2=g^2=(rg)^5=(bg)^3=(br)^5=I$

Figure 10. Plane three-color Cayley digarams with 48 and 120 points

insight into the relationship between the graphical properties and the axiomatic structure of a group. We begin with a half-groupoid and add group properties one by one, describing the graphical effect at each stage, until a group is obtained. A half-groupoid Γ is defined as a non-empty set V with a binary relation "$\cdot$" on certain pairs of elements. The Cayley diagram D of Γ is the digraph with n points corresponding to the elements of V and with an arc labeled g_2 from g_1 to g_3 whenever $g_1 \cdot g_2 = g_3$ in Γ. The half-groupoid is cancellative if its Cayley diagram satisfies the following two properties.

Lemma 6. The half-groupoid Γ is cancellative if and only if there are no two distinct arcs in $D(\Gamma)$ from one point to another and there are no two arcs with the same label which terminate at the same point.

A multiplicative set[5] is a half-groupoid in which multiplication is defined between every pair of elements. This property of multiplicative sets is demonstrated in its Cayley diagram as follows.

Lemma 7. Γ is a multiplicative set if and only if in $D(\Gamma)$ there is an arc of every color originating at every point.

A multiplicative set is associative if from every point, the two paths $g_i g_j$ followed by g_k and g_i followed by

---

[5]This structure was formerly called a "groupoid"; see, for example, R.H.Bruck, A Survey of Binary Systems, Springer-Verlag, Berlin, 1956, 1. The term "groupoid" is currently used with a different meaning, and so we use the term "multiplicative set" as in E.S.Ljapin, Semigroups, Translations of Mathematical Monographs, 3, Amer. Math. Soc., Providence, 1963.

$g_j g_k$ both lead to the same point, for all i, j, and k.

A quasi-group is a multiplicative set in which the equations $ax = b$ and $ya = b$ have solutions for all a and b in Γ. Lemma 8 describes the Cayley diagram of a quasi-group.

Lemma 8. Γ is a quasi-group if and only if $D(\Gamma)$ satisfies the following properties:

1) Arcs of all colors originate from each point.

2) Exactly one arc of each color terminates in each point.

3) Exactly one arc leads from any given point to any other point.

To construct a group we now further require that the quasi-group is associative. This adds the graphical condition that if a given sequence of arcs is a cycle from one point, then it is a cycle from every point.

Lemma 9. Γ is a group if and only if its Cayley diagram $D(\Gamma)$ satisfies the three properties of Lemma 8 and, in addition,

4) If the sequence $a_1 a_2 \cdots a_s$ of arcs describes a cycle at one point of D, then it describes a cycle at every point of D.

2. The group of a graph

We have discussed the Cayley diagram above as a pictorial representation of a group in which certain group properties can readily be visualized. We now concentrate our attention on graphs, and examine the groups of

automorphisms which arise. An automorphism of a graph or digraph is an adjacency-preserving map of the set of points of the graph into itself; the group of automorphisms of a graph is thus a permutation group acting on the points of the graph. A fundamental question concerning the automorphism group of a graph, posed by König[6] in 1936, asks whether all groups Γ are isomorphic to the (automorphism) group of some graph or digraph. We remark that the most natural place to look for an affirmative answer is to the Cayley diagrams; here, however, Γ is isomorphic to the group of adjacency- and color-preserving maps of $D(\Gamma)$.

To answer König's problem, it is convenient to examine first graphs and then digraphs; in each case, two types of permutation group isomorphism must be considered. Two permutation groups A and B, acting on the sets X and Y respectively, are abstractly isomorphic, $A \cong B$, if they are isomorphic as groups, i.e., if there is a 1-1 mapping $f: A \to B$ such that for all a_1, a_2 in A, $f(a_1a_2) = f(a_1)f(a_2)$. The groups A and B are permutationally isomorphic, $A = B$, if $A \cong B$ and there is a 1-1 correspondence between X and Y. For example, the groups $H_1 = \{(12), (12)^2 = I\}$ and $H_2 = \{(12)(34), [(12)(34)]^2 = I\}$ are abstractly, but not permutationally, isomorphic; in graph-theoretical terms, the graphs G_1 and G_2 in Figure 11, with groups H_1 and H_2 respectively, do not have the same number of points.

---

[6] König, Theorie der Endlichen und Unendlichen Graphen, 5.

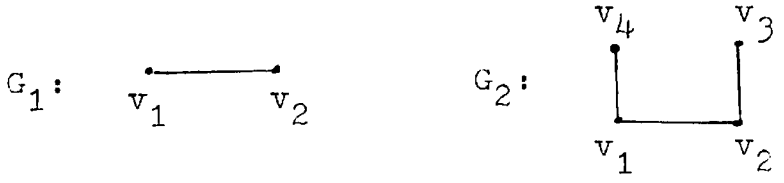


Figure 11. Graphs with abstractly, but not permutationally isomorphic groups

Before discussing König's question further, we shall examine the groups of some of the more simple graphs and how they are affected by the operations of union, cartesian product, and composition on these graphs. It is clear that the group of the complete graph, $\Gamma(K_n)$, is the symmetric group S_n, and also $\Gamma(\overline{K_n}) = S_n$. If C_n is a cycle of length n, then $\Gamma(C_n)$ is the dihedral group D_n; if G is $K_4 - x$, the third graph in Figure 12, then the automorphisms of G are I, (13)(2)(4), (24)(1)(3), and (13)(24), thus $\Gamma(G)$ is the Klein 4-group.

It will be convenient to use certain operations on permutation groups to classify the groups of graphs. The sum A + B of the permutation groups A and B acting on the disjoint sets X and Y respectively, is the permutation group acting on $X \cup Y$ and having as elements the ordered pairs $\alpha\beta$. If z is an element of $X \cup Y$, then $\alpha\beta(z)$ is defined to be $\alpha(z)$ if z is in X and $\beta(z)$ if z is in Y. Thus the group of $K_4 - x$ may be written as $S_2 + S_2$, where X is {1,3} and Y = {2,4}. In the case where A = B, we write A + B = 2A and define nA inductively. The product, A x B, of the permutation groups A and B acts on the set X x Y; its elements are of the form $\alpha \times \beta$ where $\alpha \times \beta\,(x,y)$ is defined as $(\alpha x, \beta y)$. The composition A[B] of A and B

also acts on the set X x Y, but its elements are of the form $(\alpha; \beta_1, \ldots, \beta_m)$ where $|X| = m$. Labeling the elements of X by $x_1, \ldots, x_m$ and the elements of Y by $y_1 \ldots, y_n$, the permutation $(\alpha; \beta_1, \ldots, \beta_m)$ is defined by

$$(\alpha; \beta_1, \ldots, \beta_m)(x_i, y_j) = (\alpha x_i, \beta_i y_j).$$

Several examples are shown in Figure 12; they illustrate the fact that several nonisomorphic graphs may have (permutationally or abstractly) isomorphic groups.

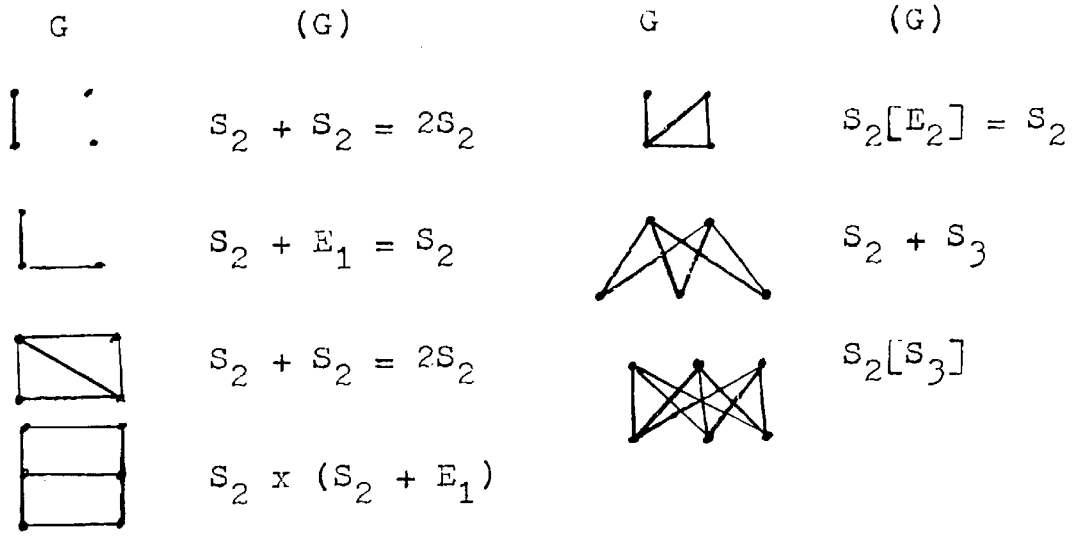


Figure 12. Graphs and their groups

We now list some of the more obvious properties of the groups of graphs.

1) $\Gamma(G) = \Gamma(\overline{G})$.

2) $\Gamma(K_n) = S_n$.

3) $\Gamma(C_n) = D_n$.

4) $\Gamma(K_{m,n}) = \begin{cases} S_m + S_n & \text{if } m \neq n \\ S_2[S_n] & \text{if } m = n. \end{cases}$

In Chapter I certain operations--union, cartesian product, composition--on graphs were defined, and it is no accident that the names of these operations closely

correspond to those for groups; Properties 5 - 8 demonstrate this relationship.

5) If no component of G_1 is isomorphic to a component of G_2, then $\Gamma(G_1 \cup G_2) = \Gamma(G_1) + \Gamma(G_2)$.

6) If G is connected, then $\Gamma(nG) = S_n[\Gamma(G)]$.

If there are graphs $G_1, G_2 \neq K_1$ such that $G = G_1 \times G_2$, we say that G is composite; if G is not composite, then it is prime. We say the graphs G_1 and G_2 are relatively prime if no prime factor of one is isomorphic to a prime factor of the other.

7) If G_1 and G_2 are relatively prime, then

$$\Gamma(G_1 \times G_2) = \Gamma(G_1) \times \Gamma(G_2).$$

Furthermore, Sabidussi[7] has shown that

8) $\Gamma(G_1[G_2]) = \Gamma(G_1)[\Gamma(G_2)]$ if and only if

a) Whenever there are two points in G_1 with the same set of adjacent points, then G_2 is connected.

b) Whenever there are two points in G_1 such that their sets of adjacent points together with the points themselves are the same, then $\overline{G_2}$ is connected.

König's question for permutational isomorphism was answered in the negative by Kagno[8] in 1947 when he proved that there is no graph whose group is permutationally

---

[7] G. Sabidussi, The composition of graphs, Duke Math. J. 26(1959), 693-696.

[8] I.N. Kagno, Linear graphs of degree ≤ 6 and their groups, Amer. J. Math. 68(1947), 505-520. Corrections 69(1947), 872, and 77(1955), 392.

isomorphic to a cyclic group; that is, there is no graph with n points whose group is C_n. To show this, we suppose that G is a graph whose group is the cyclic group $(a_1a_2\cdots a_n)$ on n elements; then $(a_1a_2\cdots a_n)$ is an automorphism of G. This, however, implies that the "reflection"--$(a_1a_n)\cdot(a_2a_{n-1})\cdots(a_ja_{n-j+1})\cdots(a_{n/2}a_{n+2/2})$ if n is even and $(a_2a_n)(a_3a_{n-1})\cdots(a_{n+1/2}a_{n+3/2})$ is n is odd--is also an automorphism of G.[9] Kagno also calculated the groups of all graphs having at most six points, thereby settling the problem of which groups are permutationally isomorphic to automorphism groups of graphs for groups of order ≤ 6.

If only abstract isomorphism is required, Frucht[10] has shown that König's question may be answered in the affirmative.

Theorem 3. Given any finite group H, there is a graph G whose group is abstractly isomorphic to H.

We shall outline the construction of such a graph as given in a later paper by Frucht.[11] We first note that any graph with points x_i can be fully described by a quadratic form $\sum_{i<k} a_{ik}x_ix_k$, where $a_{ik} = 1$ if x_i adj x_k and 0 otherwise.

---

[9] This can be verified by checking adjacency. If a_j adj a_k, where $k = j+i$ (that is, if a_j adj a_{j+i}), then, since $(a_1a_2\cdots a_n)$ is an automorphism of G, a_{j+1} adj a_{k+1}, a_{j+2}adj a_{k+2},$\cdots$,a_{n+1-k} adj $a_{n+1-k+i} = a_{n+1-j}$.

[10] R.Frucht, Herstellung von Graphen mit vorgegebener abstrakter Gruppe, Compositio Math. 6(1938), 239-250.

[11] R.Frucht, Graphs of degree three with a given abstract group, Canad. J. Math. 1(1949), 365-378.

The group of the graph is then the group of all permutations of the x_i which leave the quadratic form unchanged. In a cubic graph, we may describe each point v by a triple (κ, λ, μ), called the type of v, defined as follows: If v_1, v_2, v_3 are the three points adjacent to v, then κ, λ, μ (labeled so that $\kappa \leq \lambda \leq \mu$) give the lengths of the shortest cycles containing (v,v_1), (v,v_2), and (v,v_3). If (v,v_i) lie on no cycle, then $\mu = \infty$. For example, the type of the point v in Figure 13 is (3,3,3), while the type of the point u is $(4,4,\infty)$. It is clear that the automorphisms of a graph must also preserve the type of each point.

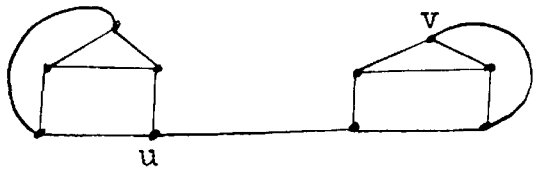


Figure 13. Cubic graph with points of type (3,3,3) and $(4,4,\infty)$

Lemma 10. A necessary condition that a point v of a cubic graph G be mapped into another point u by an automorphism of the graph is that u and v have the same type.

Corollary 10a. If v is the only point of a graph G with a given type, then v is invariant under all automorphisms of G.

We first construct a cubic graph whose group is cyclic, proving Theorem 4.

Theorem 4. If H is a cyclic group of order $h > 2$, there is a cubic graph with 6h points whose group is abstractly isomorphic to H.

Consider the quadratic form

$$\sum_{i=1}^{h} (a_i a_i + a_i e_i + a_i f_i + b_i c_i + c_i d_i + c_i f_i + e_i f_i)$$

$$+ \sum_{j=1}^{h-1} (b_j e_{j+1} + d_j d_{j+1}) + b_h e_1 + d_1 d_h \ .$$

For $h = 5$, the graph is shown in Figure 14. The types of the points in this graph are

a_i and f_i	(3,4,5)
b_i	(4,7,9)
c_i	(4,7,7)
d_i	(5,7,7)
e_i	(3,7,8) .

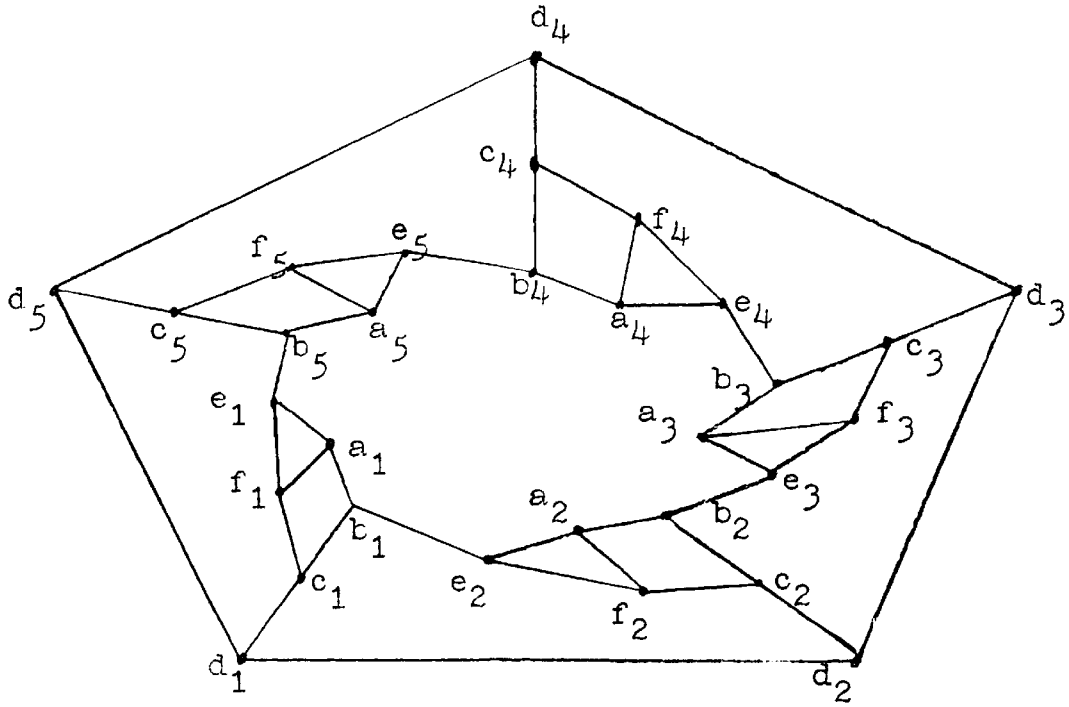


Figure 14. Graph whose group is cyclic

If $h \neq 5$, then the only types which would change would be those of the d_1, and these would become $(h,7,7)$ or $(7,7,h)$ for $h \leq 11$ and $(7,7,11)$ for $h > 11$. It is clear that a cyclic permutation ω of the indices of the points gives one automorphism of the graph. To show that this is the only automorphism, we consider first any automorphism τ which leaves the point b_1 invariant. Then since the neighbors a_1, c_1, e_2 all have different types, they too must remain invariant under τ. The new neighbors e_1, f_1 of a_1 have different types, so they remain unchanged and the third neighbor b_h of e_1 must then be invariant. The same argument now holds for all points with subscript h and further for all points in the graph.

If an automorphism σ mapped b_1 into some other point of the same type, say b_j, then this mapping taken with the inverse of the cyclic permutation ω^{j-1}, that is, $\sigma(\omega^{j-1})^{-1}$, leaves b_1 unchanged. By the above argument, then $\sigma(\omega^{j-1})^{-1}$ is the identity and $\sigma = \omega^{j-1}$. Therefore the group of the graph is the cyclic group of order h.

For noncyclic groups H it is also possible to find a cubic graph whose group is abstractly isomorphic to H.

Theorem 5. If H is any finite group of order $h \geq 3$ which is generated by n of its elements, then there is a cubic graph with $2(n+2)h$ points whose group is abstractly isomorphic to H.

When $n = 1$, the theorem is just Theorem 4. For $n > 1$, we suppose that H is generated by elements $g_1, g_2, \ldots, g_n$, with the other elements of H designated by $g_{n+1}, \ldots, g_h = I$.

Then letting

$$Q_{ij} = \sum_{k=1}^{h} x_{i,g_k} x_{j,g_k} \quad ,$$

the quadratic form

$$\begin{aligned} Q = {} & Q_{12}+Q_{14}+Q_{15}+Q_{23}+Q_{24}+Q_{35}+Q_{36}+Q_{2,2n+4}+Q_{78}+Q_{89}+\cdots \\ & +Q_{2n+3,2n+1} \\ & + \sum_{k=1}^{h} x_{5,g_k} x_{6,g_1 g_k} + x_{7,g_k} x_{8,g_2 g_k} + x_{2n+3,g_k} x_{2n+4,g_n g_k} \end{aligned}$$

describes a graph whose $2(n+2)h$ points are labeled x_{ig_k}, $i = 1,2,\ldots,2n+4$; $k = 1,2,\ldots,h$. In Figure 15 we show this graph for group whose multiplication table is

	1	h_1	h_2	h_3
1	1	h_1	h_2	h_3
h_1	h_1	1	h_3	h_2
h_2	h_2	h_3	1	h_1
h_3	h_3	h_2	h_1	1

.

To show that the group of the graph described by this quadratic form Q is actually isomorphic to H, we show first that there are h permutations $\sigma_1,\ldots,\sigma_h$ (of the points) leaving Q unaltered which constitute a group isomorphic to H. We then indicate that these are the only automorphisms of G. The permutations

$$\sigma_m = \begin{pmatrix} x_{i,h_k} \\ x_{i,h_k h_m} \end{pmatrix} \quad i = 1,2,\ldots,2n+4;\ k = 1,2,\ldots,h,$$

clearly leave each of the Q_{ij} unchanged. It is also obvious that these permutations form a group isomorphic to H. The proof that these are the only permutations is a lengthy

argument on the types of the points, similar to that for cyclic groups.

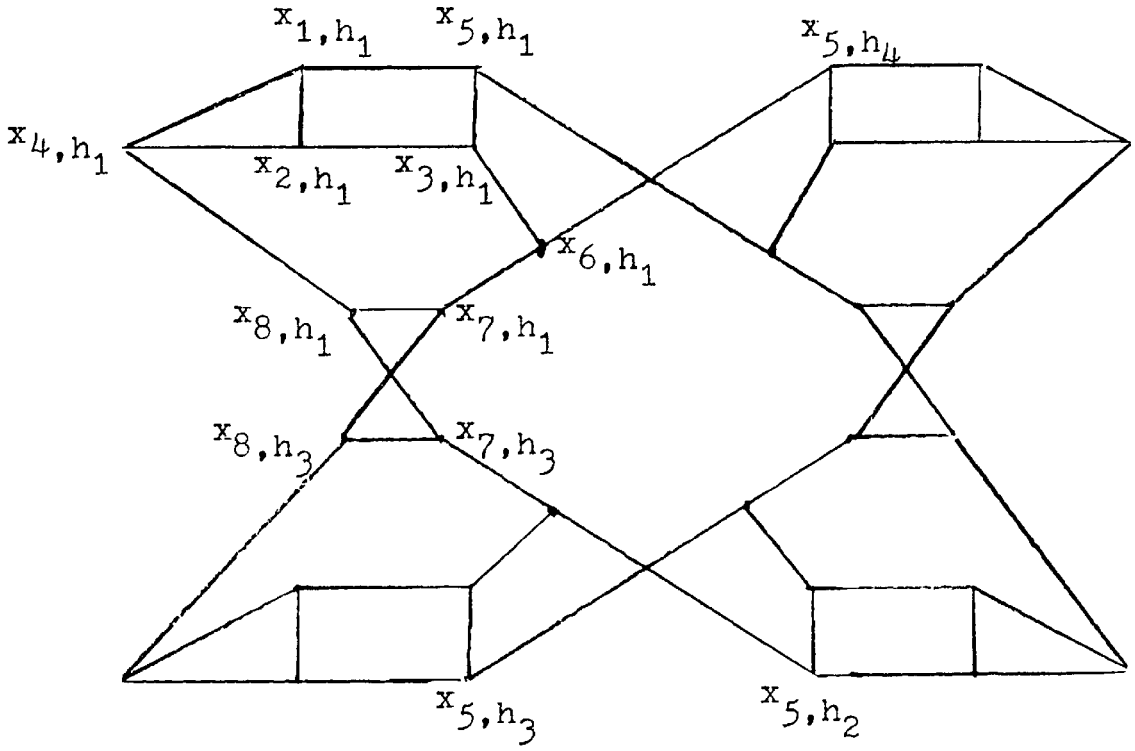


Figure 15. Graph with preassigned group

If the cubic condition is removed, satisfactory graphs with far fewer points can be found. Frucht showed that for cyclic groups of order $h > 3$, there is always a graph with $3h$ points whose group is isomorphic to H, and for noncyclic groups with n generators, there is a graph with $2nh$ points. Sabidussi [12] has shown that if n is a prime power ≥ 7, then the minimum number of points in a graph whose group is cyclic of order h is $2n$. For $n = 3,4,5$, he exhibited graphs with $3n$ points whose groups are cyclic

---

[12] G. Sabidussi, On the minimum order of graphs with given automorphism group, Monatsh. Math. 63(1959), 124-127.

of order n. Harary and Palmer[13] have also minimized the number of lines necessary in the case of n = 3,4,5, and produced the smallest graphs whose groups are C_3, C_4, C_5. We show these in Figure 16.

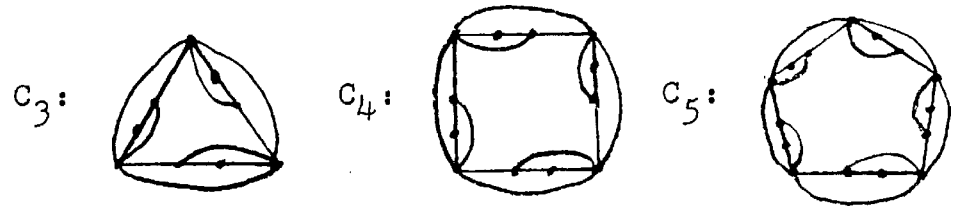


Figure 16. Smallest graphs whose groups are C_3, C_4, C_5

Many additional restrictions may be imposed on the cubic graph without destroying the property that there exists a graph of this type whose group of automorphisms is abstractly isomorphic to any given group. Sabidussi[14] and Izbicki[15] have shown that arbitrary regularity and chromatic numbers can be simultaneously specified, and in certain cases an arbitrary connectivity number can also be assigned.

Extending the discussion to include digraphs, we remark that Hemminger[16] has classified every group Γ of order ≤ 6, according to whether or not there exists a graph

---

[13] F. Harary and E. Palmer, The smallest graph whose group is cyclic, Czech. Math. J. 16(91)(1966), 70-71.

[14] G. Sabidussi, Graphs with given group and given graph-theoretical properties, Canad. J. Math. 63(1959), 124-127.

[15] H. Izbicki, Reguläre Graphen 3.4. und 5 Grades mit vorgegebenen abstrakten Automorphismengruppen, Farbenzahlen, und Zusammenhängen, Monatsh. Math. 61(1957), 42-50.

[16] R. L. Hemminger, On the group of a directed graph, Canad. J. Math. 18(1966), 211-220.

or digraph isomorphic with Γ and hence has extended Kagno's results. We refer the reader to his paper for this classification.

VI. FINITE PROJECTIVE PLANES AND RELATED SYSTEMS

In this chapter we discuss several topics from combinatorial mathematics; orthogonal Latin squares, finite projective planes, Hadamard matrices, symmetric block designs, and difference sets, which are very closely related to each other and to a class of bipartite graphs. These topics have, to a large extent, been studied separately, but they share many existence theorems and unsolved problems. It is our purpose to give brief descriptions of the topics and to analyze the relationship between them. Several unsolved problems are presented, together with a conjecture made by Euler which was not resolved until 1959.

A Latin square of order n is an n x n array of numbers from the set $S = \{0,1,\ldots,n-1\}$ with the property that each element of S occurs exactly once in each row and column. Two Latin squares are <u>weakly equivalent</u> if one may be obtained from the other by permuting the rows, permuting the columns, or relabeling the entries uniformly. We say that a Latin square is in <u>normalized</u> form if the elements in its first row and column occur in increasing order, and it is clear that every Latin square is weakly equivalent to one in normalized form. All of the normalized Latin squares of order 4 are shown in Figure 1; they fall into two weak

equivalence classes, since the Latin squares L_2 and L_3 may be obtained from L_1 by interchanging the labels 0 and 1 or 0 and 2, respectively.

$$L_1 = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 2 & 3 & 0 \\ 2 & 3 & 0 & 1 \\ 3 & 0 & 1 & 2 \end{bmatrix} \qquad L_2 = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 3 & 0 & 2 \\ 2 & 0 & 3 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$$

$$L_3 = \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 1 & 0 \\ 3 & 2 & 0 & 1 \end{bmatrix} \qquad L_4 \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}$$

Figure 1. Latin squares of order 4

We have given the classical definition of a Latin square, but there are many equivalent definitions. Some of these lead to natural generalizations or emphasize the characteristics which Latin squares share with other mathematical configurations. Two, in particular, are of value here.[1]

Theorem 1. A Latin square of order n is a set of triples (a,b,c) of elements from the set S with the property that each ordered pair of elements of S occurs exactly once in each pair of coordinates.

In this representation, the first coordinate represents the row, the second coordinate the column, and the third coordinate the digit of the classical Latin square. A third representation of a Latin square is given in

---

[1]We remark that the multiplication table of any group of order n gives rise to a Latin square; the converse, however, is not true.

Theorem 2.

Theorem 2. A Latin square of order n is equivalent to a collection of three families of n parallel lines on n^2 points such that each line contains n points, and through each point passes exactly one line from each family.

In Theorem 2 the families of lines represent the rows, columns, and digits of a classical Latin square, and in this setting, two Latin squares are weakly equivalent if one may be obtained from the other by permutations of lines within families. A second type of equivalence, strong equivalence, permits also the interchange of families.

It is clear that there is at least one Latin square of every order n, since the elements of S can be written in increasing order in the first row of the array and permuted cyclically to give the remaining rows. A lower bound for the number of Latin squares of order n can be obtained from Hall's theorem on systems of distinct representatives. An r x n Latin rectangle is an r x n array such that each element of S occurs exactly once in each row and at most once in each column. The sets S_i, containing the elements of S which do not appear in the i-th column of a Latin rectangle, satisfy Hall's condition and therefore have at least (n-r)! systems of distinct representatives. This proves the following theorem.

Theorem 3. An r x n Latin rectangle can be extended to an (r+1) x n Latin rectangle in at least (n-r)! ways.

Corollary 3a. There are at least $n!(n-1)!\ldots1!$ Latin squares of order n.

The determination of the exact number of Latin squares of each order n is a difficult and, in general, unsolved problem. Table 3 gives the numbers for small n, where the symbol $L_n(n)$ denotes the number of normalized Latin squares of order n, $L_s(n)$ the number of strong equivalence classes, $L_w(n)$ the number of weak equivalence classes, and L(n) the total number of Latin squares of order n.

n	2	3	4	5	6	7	8
$L_w(n)$	1	1	2	2	17	563	
$L_s(n)$	1	1	2	2	12	147	
$L_n(n)$	1	1	4	56	9408	16,942,080	535,281,401,856

$L(n) = n!(n-1)!L_n(n)$

Table 3. The number of Latin squares of small order[2]

Two Latin squares $L_1 = ((a_{ij}^{(1)}))$ and $L_2 = ((a_{ij}^{(2)}))$ are called orthogonal if the n^2 pairs $(a_{ij}^{(1)}, a_{ij}^{(2)})$ are distinct; a pair of orthogonal Latin squares is often called an Euler square, or a Graeco-Latin square. Euler squares of orders 3 and 10 are shown in Figure 2. An Euler square may be alternately defined in a manner corresponding to the alternate definitions of a Latin square.

---

[2]For $n \leq 7$ these values come from R.A.Fisher and F. Yates, Statistical Tables for Biological, Agricultural, and Medical Research, Hafner Publishing Co., 1963. For n = 8, see M.B.Wells, The number of Latin squares of order eight, J.Combinatorial Theory 3(1967), 98-99.

Theorem 4. An Euler square of order n may equivalently be defined as a set of 4-tuples of elements from S such that each ordered pair (u,v) of elements of S occurs exactly once in each pair of coordinates.

Theorem 5. An Euler square of order n is equivalent to a collection of four families of n parallel lines on n^2 points such that each line contains n points, and through each point passes exactly one line from each family.

$$L_1 = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{bmatrix} \qquad L_2 = \begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}$$

$$L_3 = \begin{bmatrix} 0 & 6 & 5 & 4 & 9 & 8 & 7 & 1 & 2 & 3 \\ 7 & 1 & 0 & 6 & 5 & 9 & 8 & 2 & 3 & 4 \\ 8 & 7 & 2 & 1 & 0 & 6 & 9 & 3 & 4 & 5 \\ 9 & 8 & 7 & 3 & 2 & 1 & 0 & 4 & 5 & 6 \\ 1 & 9 & 8 & 7 & 4 & 3 & 2 & 5 & 6 & 0 \\ 3 & 2 & 9 & 8 & 7 & 5 & 4 & 6 & 0 & 1 \\ 5 & 4 & 3 & 9 & 8 & 7 & 6 & 0 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 & 0 & 1 & 7 & 8 & 9 \\ 4 & 5 & 6 & 0 & 1 & 2 & 3 & 8 & 9 & 7 \\ 6 & 0 & 1 & 2 & 3 & 4 & 5 & 9 & 7 & 8 \end{bmatrix} \qquad L_4 = \begin{bmatrix} 0 & 7 & 8 & 9 & 1 & 3 & 5 & 2 & 4 & 6 \\ 6 & 1 & 7 & 8 & 9 & 2 & 4 & 3 & 5 & 0 \\ 5 & 0 & 2 & 7 & 8 & 9 & 3 & 4 & 6 & 1 \\ 4 & 6 & 1 & 3 & 7 & 8 & 9 & 5 & 0 & 2 \\ 9 & 5 & 0 & 2 & 4 & 7 & 8 & 6 & 1 & 3 \\ 8 & 9 & 6 & 1 & 3 & 5 & 7 & 0 & 2 & 4 \\ 7 & 8 & 9 & 0 & 2 & 4 & 6 & 1 & 3 & 5 \\ 1 & 2 & 3 & 4 & 5 & 6 & 0 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 6 & 0 & 1 & 9 & 7 & 8 \\ 3 & 4 & 5 & 6 & 0 & 1 & 2 & 8 & 9 & 7 \end{bmatrix}$$

Figure 2. Euler squares of orders 3 and 10

It is clear that if two Latin squares are orthogonal, that a uniform relabeling of the elements of one or both does not destroy the orthogonal property. Therefore, in considering a set of mutually (i.e., pairwise) orthogonal Latin squares, we may assume that the first row of each Latin square contains the elements of S in increasing order. The orthogonality condition then implies that the first element of the second row must be different in each Latin square in the set, thus n-1 is an upper bound on the number of mutually orthogonal Latin squares of order n.

Theorem 6. If $L_1, L_2, \ldots, L_t$ are t orthogonal Latin squares of order n, then $t \leq n-1$.

A set of n-1 orthogonal Latin squares of order n is called complete.

It is convenient to represent a set of t orthogonal Latin squares of order n by a $n^2 \times (t+2)$ array. This array is obtained by letting the first two columns contain the n^2 pairs of elements of S in the order (0,0), (0,1),...,(0,n), ...,(n,n), and each remaining column contain the elements of one of the Latin squares, taken by rows. Thus the Latin squares of order 3 in Figure 2 can be represented by the 9 x 4 matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 2 & 0 \\ 2 & 1 & 0 & 1 \\ 2 & 2 & 1 & 2 \end{bmatrix}.$$

We note that this representation amounts to a listing of the 4-tuples of the Euler square, and, in the general case, a listing of the n^2 (t+2)-tuples equivalent equivalent to a set of t mutually orthogonal Latin squares. Theorem 7 summarizes this correspondence.

Theorem 7. A set of $t \geq 2$ orthogonal Latin squares of order $n \geq 3$ is equivalent to an $n^2 \times (t+2)$ array of elements of S in which each $n^2 \times 2$ subarray gives the n^2 ordered pairs of elements of S.

We may assume that the first two columns of this $n^2 \times (t+2)$ array contain the n^2 2-samples of S in the order listed above. Then each of the remaining t columns lists the elements of a Latin square by its rows. The arrangement of elements in the first column of the array insures that no row of a Latin square represented by a subsequent column has two equal entries, and the arrangement of elements in the second column of the array does the same for the columns of the Latin squares.

Using this representation of a set of orthogonal Latin squares, it is not difficult to establish the following existence theorem.

Theorem 8. If there is a set of t mutually orthogonal Latin squares of order n and a set of t mutually orthogonal Latin squares of order n', then there is a set of t orthogonal Latin squares of order nn'.

It is an easy task to construct pairs of orthogonal Latin squares of orders 3, 4, and 5. Euler conjectured that not only is there no Euler square of order 6, but also that there is none of any order $n \equiv 2 \pmod 4$. The conjecture for $n = 6$ was verified by Tarry[3] in 1900 by a systematic examination of all Latin squares of that order.

Such a method, however, was impractical for larger n, even with the aid of computers. In 1959 Bose and

---

[3] G. Tarry, Le Problème de 36 officieurs, Compte Rendu de l'Association Francaise pour l'Avancement de Science Naturel 1(1900), 122-123, and 2(1901), 170-203.

Shrikhande[4] found a counterexample for $n = 22$, and Parker[5] published two orthogonal Latin squares of order 10, in fact, those of Figure 2, thereby disproving Euler's conjecture. Moreover, in 1960 Bose, Shrikhande, and Parker[6] completed a lengthy process of disproving the conjecture for every $n > 6$. Theorem 9 summarizes their results.

Theorem 9. If $n \equiv 2 \pmod 4$ and $n > 6$, then there is an Euler square of order n.

The existence of complete sets of orthogonal Latin squares is a much more difficult question which has been answered in the affirmative only for the case in which n is a prime power.[7]

Theorem 10. If $n = p^{\alpha}$, and $n \geq 3$, then there is a complete set of orthogonal Latin squares of order n.

The usual construction of a complete set of orthogonal Latin squares of order p^{α} makes use of the field of order p^{α} with elements $0, a_1, \ldots, a_{n-1}$. The $n-1$ Latin squares L_k are defined by $L_k = ((a_{ij}^{(k)}))$ where $a_{ij}^{(k)} = a_k a_i + a_j$. Using the field properties, one may show that each L_k is a Latin square and that any pair is orthogonal.

Theorem 10 combines with Theorem 8 to give the

---

[4] R.C.Bose and S.S.Shrikhande, On the falsity of Euler's conjecture about the non-existence of two orthogonal Latin squares of order 4t+2, Proc. Nat. Acad. Sci. U.S.A. 45(1959), 734-737.

[5] E.T.Parker, Orthogonal Latin squares, Proc. Nat. Acad. Sci. U.S.A. 45(1959), 859-862.

[6] R.C.Bose, S.S.Shrikhande, and E.T.Parker, Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture, Canad. J. Math. 12(1960), 189-203.

following corollary.

Corollary 10a. If $n = \prod_{i=1}^{r} p_i^{\alpha_i}$ is the prime power decomposition of n, and if $t = \min_{1 \le i \le r} (p_i^{\alpha_i} - 1) > 1$, then there is a set of t orthogonal Latin squares of order n.

Complete sets of orthogonal Latin squares of orders 3,4,5,7,8, and 9 are guaranteed by Theorem 10; the truth of Euler's conjecture for n = 6 implies that there is no complete set of order 6; and it is unknown whether a complete set for n = 10 exists. The importance of complete sets of orthogonal Latin squares is illustrated by their relationship to finite projective planes, which we now discuss.

A finite projective plane π is a set of objects called points and a set of collections of points, called lines, satisfying the following incidence axioms.

A1. Two distinct points of π are on exactly one line of π.

A2. Two distinct lines of π have exactly one point of π in common.

A3. There are four distinct points of π, no three of which lie on the same line.

The third axiom eliminates certain degenerate cases such as a structure with one line and one point or one with three lines and three points. From these axioms, the following observations may be made.

Theorem 11. The total number of points on any line of π is equal to the total number of lines through any point. Calling this number n+1, the total number of points of π is n^2+n+1 and the total number of lines is n^2+n+1.

The number n is called the order of the projective plane π, and axiom A1 can be replaced by the following axiom.

A1'. Every line passes through n+1 points and every point lies on n+1 lines.

A fundamental property of finite projective planes is their equivalence to complete sets of orthogonal Latin squares.

Theorem 12. For $n \geq 3$, there is a finite projective plane of order n if and only if there is a complete set of orthogonal Latin squares of order n.

We give the construction. Given a complete set of orthogonal Latin squares, we represent it in an $n^2 \times (n+1)$ array A and construct a finite projective plane π_n from it. Let the points of π_n be $Q_0, Q_1, \ldots, Q_{n^2-1}$, corresponding to the rows of A, together with the additional points P_0, $P_1, \ldots, P_n$. A line L_{ij}, $0 \leq i \leq n-1$ and $0 \leq j \leq n$, will contain the points P_j and all Q_k having an i in the j-th column. This defines $n^2 + n$ lines; the remaining line L contains the points $P_0, P_1, \ldots, P_n$. These points and lines clearly satisfy axiom A1'. Furthermore the four points corresponding to the rows with (1,1), (1,2), (2,1), and (2,2) in the first two columns of A satisfy axiom A3. It remains

to show that two distinct lines contain exactly one common point. If L_{ij}, $L_{i',k}$ are two lines with $j \neq k$, then they share the unique point corresponding to the row with i in column j and i' in column k. If L_{ij} and $L_{i',j}$ are lines with $i \neq i'$, then they both pass through P_j and no other point. The lines L_{ij} and L share the point P_j. Therefore the points and lines described comprise a projective plane π_n.

Conversely, if π_n is a projective plane of order n, an n^2 x (n+1) array A can be constructed as follows. Label the points on any line L by $P_0, P_1, \ldots, P_n$, and let the points not on L be labeled arbitrarily $Q_0, Q_1, \ldots, Q_{n^2-1}$. Arbitrarily label the lines other than L through each P_j by $L_0, L_1, \ldots, L_{n-1}$. Then $A = ((a_{ij}))$ for $i = 0,1,\ldots,n^2-1$ and $j = 0,1,\ldots,n$, where a_{ij} is the label of the line Q_iP_j. The rows of each n^2 x 2 subarray are the n^2 ordered pairs of S, for if $a_{ij} = a_{i',j}$ and $a_{ik} = a_{i',k}$ (where $i \neq i'$, $j = k$), then $Q_iP_j = Q_{i'}P_j$ and $Q_iP_k = Q_{i'}P_k$. But this implies that the line $Q_iQ_{i'}$ passes through both P_j and P_k, and therefore by axiom A1, the line $Q_iQ_{i'}$ coincides with L. However, by definition Q_i and $Q_{i'}$ are not on L.

As an example, we construct π_3 from the complete set of orthogonal Latin squares of order 3 given by the matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 \\ 1 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ 2 & 0 & 2 & 1 \\ 2 & 1 & 0 & 2 \\ 2 & 2 & 1 & 0 \end{bmatrix}.$$

The points of π_3 are $P_0, P_1, P_2, P_3, Q_0, Q_1, \ldots, Q_8$ and the lines, given by their points, are

$$L = (P_0, P_1, P_2, P_3), \qquad L_{00} = (P_0, Q_0, Q_1, Q_2),$$

$$L_{10} = (P_0, Q_3, Q_4, Q_5), \qquad L_{20} = (P_0, Q_6, Q_7, Q_8),$$

$$L_{01} = (P_1, Q_0, Q_3, Q_6), \qquad L_{11} = (P_1, Q_1, Q_4, Q_7),$$

$$L_{21} = (P_1, Q_2, Q_5, Q_8), \qquad L_{02} = (P_2, Q_0, Q_5, Q_7),$$

$$L_{12} = (P_2, Q_1, Q_3, Q_8), \qquad L_{22} = (P_2, Q_2, Q_4, Q_0),$$

$$L_{03} = (P_3, Q_0, Q_4, Q_8), \qquad L_{13} = (P_3, Q_1, Q_5, Q_6),$$

$$L_{23} = (P_3, Q_2, Q_3, Q_7).$$

The question of the existence of finite projective planes is thus the same question posed for complete sets of orthogonal Latin squares. There are projective planes of order $n = p^{\alpha}$ where p is a prime and $n > 2$; there is no projective plane of order 6; and the first undecided case is $n = 10$. Axiom A3 implies that there is no projective plane of order 1. Moreover, it has been shown that the projective planes of orders 2,3,4,5,7, and 8 are unique[7], whereas[8] there is more than one plane of order 9. The exact number of projective planes of order 9 is not known.

---

[7]The verification of uniqueness for $n = 2,3,4,5$ is straightforward. The proof for $n = 7$ is given in M.Hall, Uniqueness of the projective plane with 57 points, Proc. Amer. Math. Soc. 4(1953), 912-916, with corrections in 5(1954), 994-997, together with W.A.Pierce, The impossibility of Fano's configuration in a projective plane with eight points per line, Proc. Amer. Math. Soc. 4(1953), 908-912. The uniqueness of π_8 is proved in M.Hall, J.D.Swift, and R.J.Walker, Uniqueness of the projective plane of order eight, Math. Tables Aids Comput, 10(1956), 186-194.

[8]O.Veblen and J.H.M.Wedderburn, Non-desarguesian and non-Pascalian geometries, Trans.Amer.Math.Soc.8(1907),370-388.

A fundamental theorem on the nonexistence of finite projective planes was first proved by Bruck and Ryser[9] in 1943. (We will indicate its proof at a later time when we have developed the necessary prerequisites.[10])

Theorem 13. If $n \equiv 1$ or $2 \pmod 4$ and if there is a prime $p \equiv 3 \pmod 4$ which divides the squarefree part of n, then there is no projective plane of order n.

The projective plane of order 2 has seven points $P_0, P_1, \ldots, P_6$, and seven lines which may be given by the following triples of points: (P_0,P_1,P_3), (P_1,P_2,P_4), (P_2,P_3,P_5), (P_3,P_4,P_6), (P_4,P_5,P_0), (P_5,P_6,P_1), (P_6,P_0,P_2). We note that π_2 could have been defined recursively by taking $L_0 = (P_0,P_1,P_3)$ and constructing L_i from L_{i-1} by raising the index of each point of L_{i-1} by 1 (mod 7). A finite projective plane which can be described in this manner is called cyclic. Noncyclic planes are known to exist; there is one of order 9, for example.[11] We show below that there are cyclic planes of every order p^α, however, so that π_2, π_3, π_4, π_5, π_7, and π_8 are all cyclic. Two conjectures about cyclic projective planes are now given.

Conjecture 1. The order of every cyclic projective plane is p^α.

Conjecture 2. There is a cyclic projective plane of every order for which there is a projective plane.

---

[9] R.H.Bruck and H.J.Ryser, The nonexistence of certain finite projective planes, Canad. J. Math. 1(1949), 88-93.

[10] See p. 158.

[11] O.Veblen and J.H.M.Wedderburn, 379-388.

The projective plane π_2 can also be viewed as an example of another mathematical system closely related to projective planes. A Steiner triple system is a set of triples of elements of a set S such that each pair of elements of S occurs in exactly one triple; the Steiner triple system on three elements is $\{(1,2,3)\}$. The order of a Steiner triple system in $|S|$ and therefore the triple system of order 7 is equivalent to the projective plane of order 2.

Steiner triple systems are part of a large class of mathematical configurations called balanced incomplete block designs. Let $S = \{s_1, s_2, \ldots, s_v\}$ be a set of v elements and let $A_1, A_2, \ldots, A_b$ be arbitrary nonempty distinct subsets of S. The collection $\{A_i | i = 1,2,\ldots,b\}$ is a (v,b,r,k,λ)-balanced incomplete block design whenever

1) $|A_i| = k$ for $i = 1,2,\ldots,b$.
2) Each element of S is contained in exactly r of the sets A_i.
3) Each (unordered) pair (s_m, s_n) of elements of S occurs in exactly λ of the sets A_i.
4) $0 < \lambda < k < v-1$.

The numbers v,b,r,k,λ in a balanced incomplete block design are not independent, but satisfy the following identities.

Lemma 1. In a (v,b,r,k,λ)-block design,

1) $bk = vr$.
2) $r(k-1) = \lambda(v-1)$.

These identities are easy to verify since the left hand side of (1) adds the cardinalities of the A_i while the right hand side counts the occurrences of each s_m in the sets A_i. The second identity counts, on the right, the number of occurrences of the pairs (s_m, s_n) for some fixed m, and on the left, the number of s_n, where $m \neq n$, in all sets A_i in which s_m appears. We remark that the second identity shows that condition 2 for a balanced incomplete block design is a necessary consequence of conditions 1 and 3.

The Steiner triple system of order v is a (v,b,r,3,1)-block design. Moreover, the identities of Lemma 1 imply that $r = \frac{v-1}{2}$ and $b = \frac{v(v-1)}{6}$, and therefore a Steiner triple system always has $v \equiv 1$ or 3 (mod 6) and $v \geq 3$. These conditions on v, b, and r are also sufficient for the existence of a Steiner triple system.

A b x v incidence matrix $A = ((a_{ij}))$ is associated with a block design in a natural way, the (i,j)-entry being 1 if s_j is in A_i and 0 otherwise. Theorem 14 describes some of the properties of this matrix.

Theorem 14. If A is the matrix of a (v,b,r,k,λ)-block design, then

1) $AJ = kJ'$,

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

3) $|A^tA| = (r-\lambda)^{v-1}(v\lambda - \lambda + r)$,

where J is the v x v matrix of ones, and J' is the b x v matrix of ones.

The first property follows immediately from the fact

that each A_i has exactly k elements. The second property is the result of the fact that each element s_i is in r subsets A_m and each pair (s_i, s_j) is in λ subsets A_m. The third property is a direct computational consequence of the second.

Property 2 is also a sufficient condition for a (0,1)-matrix A to be the incidence matrix of a block design, and given such a matrix we may (by interchanging the 0's and 1's of A) construct a new block design, called the complementary block design.

Theorem 15. If the 0's and 1's of the incidence matrix of a (v,b,r,k,λ)-block design are interchanged, the resulting matrix is the incidence matrix of a $(v,b,v-r,v-k,b-2r+\lambda)$-block design.

The relationship of finite projective planes with block designs becomes evident when we consider a particular class of block designs. The subsets $A_1, A_2, \ldots, A_v$ of a set S of v elements form a (v,k,λ)-symmetric block design if

1) $|A_i| = k$.

2) $|A_i \cap A_j| = \lambda$ for $i \neq j$.

3) The integers v,k,λ satisfy $1 \leq \lambda < k < v-1$.

Of course, it must be shown that this configuration is actually a block design; this is accomplished by establishing the properties of Theorem 14 for the incidence matrix of a symmetric block design.

Lemma 2. If A is the incidence matrix of a (v,k,λ)-symmetric block design, then

1) $AJ = kJ'$.

2) $AA^t = (k-\lambda)I + \lambda J$, where J is the vxv matrix of 1's.

Furthermore, if A is a (0,1)-matrix satisfying (2), then A is the incidence matrix of a symmetric block design. An important new property of the incidence matrix of a symmetric block design is normality. The verification of this property uses an involved matrix argument.[12]

Theorem 16. If A is the incidence matrix of a symmetric block design, then $AA^t = A^tA$.

Combining Theorem 16 with Lemma 2, we note that A satisfies property 2 of block designs and therefore A is also the incidence matrix of a block design.

Corollary 16a. A (v,k,λ)-symmetric block design is equivalent to a (v,b,r,k,λ)-block design with $v=b$ and $k=r$.

Corollary 16b. In a (v,k,λ)-symmetric block design, $k(k-1) = \lambda(v-1)$.

Now viewing the elements of S as points and the sets A_i as lines, we obtain a correspondence between symmetric block designs and finite projective planes. The point-line incidence of the plane is equivalent to the element-set inclusion of the symmetric block design, and thus property 1 of a symmetric block design guarantees n+1 points on a line, and property 2 says that every pair of lines intersect in exactly one point.

Corollary 16c. A projective plane of order n is equivalent to a (v,k,λ)-symmetric block design with $v = n^2 + n + 1$, $k = n+1$, and $\lambda = 1$.

---

[12] A proof is given in H.J.Ryser, Combinatorial Mathematics, Carus Mathematical Monographs, 14, Mathematical Association of America, Buffalo, 1963, 103-104.

Necessary conditions for the existence of a symmetric block design have been established by Chowla and Ryser.[13] Other than the condition $k(k-1) = \lambda(v-1)$, these are the only known necessary conditions.

Theorem 17. If a (v,k,λ)-symmetric block design is given and n is defined as $k - \lambda$, then

1) whenever v is even, n is a square,

2) whenever v is odd, the Diophantine equation

$$z^2 = nx^2 + (-1)^{(v-1)/2}\lambda y^2$$

has a nontrivial solution.

The proof is immediate when v is even, since $|A|^2 = |AA^t| = k^2(k-\lambda)^{v-1}$. Then $k^2(k-\lambda)^{v-1}$ must be a perfect square, and since v-1 is odd, $k-\lambda = n$ must be a square. The proof for v odd involves a surprisingly deep theorem--the Hasse-Minkowski theorem--on the representation of natural numbers in the theory of binary quadratic forms. Except for the case $\lambda = 1$, no elementary proof is known. An alternate statement of the condition for v odd is given in the first corollary.

Corollary 17a. If the g.c.d. of k and λ is 1 and if n' and λ' denote the squarefree parts of n and λ respectively, then a necessary condition for the existence of a (v,k,λ)-symmetric block design is

1) if v is even, then n is a square;

2) if v is odd, then $(-1)^{(v-1)/2}\lambda'$ is a square (mod n').

---

[13] S. Chowla and H.J. Ryser, Combinatorial Problems, Canad. J. Math. 2(1950), 93-99.

The Bruck-Ryser theorem can now be proved as a direct corollary to the Chowla-Ryser theorem. We state it as the second corollary.

Corollary 17b. If $n \equiv 1$ or $2 \pmod 4$ and if there is a prime $p \equiv 3 \pmod 4$ which divides the squarefree part n' of n, then there is no projective plane of order n.

To establish this corollary, we note that if $n \equiv 1$ or $2 \pmod 4$, then $v = n^2 + n + 1$ is odd, and $\frac{v-1}{2} = \frac{n(n+1)}{2}$ is also odd. The Diophantine equation of Theorem 17 is then

$$z^2 = nx^2 - y^2 = n'x_1^2 - y^2.$$

If this is to have a solution, $\left(\frac{-1}{p}\right)$ must be 1; however, $\left(\frac{-1}{p}\right) = -1$ for all $p \equiv 3 \pmod 4$.

It has been conjectured that the Chowla-Ryser conditions are also sufficient for the existence of a (v,k,λ)-symmetric block design. The verification of such a conjecture would have many far-reaching consequences. Since $(v,k,\lambda) = (111,11,1)$ satisfies the conditions, the conditions, the conjecture implies the existence of a projective plane of order 10 which is not a prime power. Furthermore, it is known that there is no cyclic projective plane of this order.[14]

A partial converse of the Chowla-Ryser theorem has been proved[15], and serves to enhance the possibility that a complete converse exists.

---

[14] See p.165 below.

[15] Stated in Chowla and Ryser, 93-99. The proof is given in detail in M.Hall, Combinatorial Theory, Blaisdell Publishing Co., Waltham, Mass., 1967, 114-119.

Theorem 18. If the integers v,k,λ with $0 < \lambda < k < v-1$ satisfy the conditions

1) $n = k - \lambda$ is a square when v is even,

2) the Diophantine equation

$$z^2 = nx^2 + (-1)^{(v-1)/2}\lambda y^2$$

has a nontrivial solution when v is odd,

then there is a $v \times v$ matrix A with rational entries which satisfies $AA^t = A^tA = (k - \lambda)I + \lambda J$.

There is a particular subclass of symmetric block designs which have been studied in detail within the context of matrix theory. A Hadamard matrix H_n of order n is an $n \times n$ matrix whose elements are $+1$ or -1 such that $HH^t = nI$. For example, $H_1 = (1)$ or (-1),

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

and

$$H_4 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Since $HH^t = nI$, then $|HH^t| = n^n$ and $||H|| = n^{n/2}$, and since $n^{n/2} \neq 0$, H^{-1} exists and $H^{-1} = n^{-1}H^t$. Solving for H^t and multiplying on the right by H, we obtain $HH^t = nI = H^tH$, thus proving the following lemma.

Lemma 3. Every Hadamard matrix H is normal.

The inner product of any two rows of a Hadamard matrix is zero, and therefore it is clear that for $n > 1$, all Hadamard matrices H_n have even order. Moreover, we now show that for $n > 2$, the existence of H_n implies $n \equiv 0 \pmod 4$.

Lemma 4. If H_n is a Hadamard matrix with $n > 2$, then $n \equiv 0 \pmod 4$.

To show this, we let $H = ((a_{ij}))$ and define $B = ((b_{ij}))$ by $b_{ij} = (1-a_{ij})/2$, that is, $b_{ij} = 0$ if $a_{ij} = 1$ and $b_{ij} = 1$ if $a_{ij} = -1$. Taking the inner product of the first two rows of H,

$$0 = \sum_j a_{1j}a_{2j} = \sum_j (1 - 2b_{1j})(1 - 2b_{2j})$$

$$= n - 2\sum_j (b_{1j} + b_{2j}) + 4\sum_j b_{1j}b_{2j} \quad .$$

Similarly, taking the inner products $\sum_j a_{1j}a_{3j}$ and $\sum_j a_{2j}a_{3j}$, we obtain

$$0 = n - 2\sum_j (b_{1j} + b_{3j}) + 4\sum_j b_{1j}b_{3j}$$

and

$$0 = n - 2\sum_j (b_{2j} + b_{3j}) + 4\sum_j b_{2j}b_{3j} \quad .$$

Adding these three equations,

$$0 = 3n - 4\sum_j (b_{1j} + b_{2j} + b_{3j}) + 4\sum_j (b_{1j}b_{2j}+b_{1j}b_{3j}+b_{2j}b_{3j})$$

which implies that $n \equiv 0 \pmod 4$.

Paley[16] conjectured that Hadamard matrices exist for all $n \equiv 0 \pmod 4$; his conjecture has been verified[17] for

---

[16] R.E.A.C.Paley, On orthogonal matrices, Jour. Math. and Phys. 12(1933), 311-320.

[17] Paley proved it with the exception of n = 92, 116, 172, 184, and 188. The cases n = 92 and 184 were solved by L.Baumert, S.W.Golomb, and M.Hall in Discovery of an Hadamard matrix of order 92, Bull. Amer. Math. Soc. 68(1962), 237-238. The cases n = 116 and 232 were solved by L. Baumert in Hadamard matrices of orders 116 and 232, Bull. Amer. Math. Soc. 72(1966), 237. J.Williamson solved the case n = 172 in Hadamard's determinant theorem and the sum of four squares, Duke Math. J. 11(1944), 65-81.

A classification of all known results, together with a new class of Hadamard matrices, which leaves only the

all such $n \leq 200$ except 156 and 188. We now mention briefly some general existence theorems.

Theorem 19. If there is a Hadamard matrix of order n, then there is a Hadamard matrix of order 2n.

Theorem 20. If there are Hadamard matrices of orders n and m, then there is a Hadamard matrix of order mn.

These two theorems are easily proved by taking

$$H_{2n} = \begin{pmatrix} H_n & H_n \\ H_n & -H_n \end{pmatrix}$$

and $H_{mn} = ((a_{ij}H_m))$ where $H_n = ((a_{ij}))$.

An impressive existence theorem, published by Paley[18] in 1933, constructs a Hadamard matrix of order p+1 for every prime $p \equiv 3 \pmod 4$.

Theorem 21. If $p \equiv 3 \pmod 4$, then there is a Hadamard matrix of order p+1.

It is clear that two rows or two columns of a Hadamard matrix may be interchanged and any row or column multiplied by -1 without destroying the Hadamard property. A Hadamard matrix is said to be normalized if all of the entries of the first row and first column are +1. We now show that the normalized Hadamard matrices are isomorphic with a subclass of symmetric block designs[19], as asserted above.

---

values 188, 236, 260, 268, 292, 356, 372, and 376 less than 400 to be decided, is given in E.Spence, A new class of Hadamard matrices, Glasgow Math. J. 8(1967), 59-62.

[18] Paley, 311-320.

[19] J.A.Todd, A combinatorial problem, Jour. Math. and Phys. 12(1933), 321-333.

Theorem 22. A normalized Hadamard matrix of order $n = 4t \geq 8$ is equivalent to a $(4t-1, 2t-1, t-1)$ symmetric block design.

To prove this, we let H be a Hadamard matrix of order $4t$, and construct a (0,1)-matrix of order $4t-1$ by deleting the first row and column of H and changing all -1's to 0. Since H is Hadamard and normalized, every row of A will have $2t$ zeros and $2t-1$ ones. Therefore, the diagonal entries of AA^t are $2t-1$. A similar consideration of the Hadamard property shows that the other entries of AA^t are $t-1$. Therefore $AA^t = tI + (t-1)J$ and A is the incidence matrix of a $(4t-1, 2t-1, t-1)$-symmetric block design.

The converse construction produces a normalized Hadamard matrix of order $4t$ from a $(4t-1, 2t-1, t-1)$-symmetric block design.

The normalized Hadamard matrix of order 8 is of special interest since it is equivalent to a (7,3,1)-symmetric block design and therefore to the projective plane π_2. (We note that this is the only Hadamard matrix equivalent to a projective plane.)

$$H_8 = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 \end{pmatrix}$$

The Hadamard matrix

$$H = \begin{pmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$$

has the property that the j-th entry of the i-th row is equal to the (j-1)-entry of the preceding row; such a matrix is called a <u>circulant</u> and has the properties that

$HH^t = nI$, since H is Hadamard,

$HJ = H^t$, since H is a circulant,

$= kJ$ where k is the column sum of H.

Therefore, $HH^tJ = kHJ = k^2J = nIJ = nJ$, and $n = k^2$, proving the following lemma.

<u>Lemma 5</u>. If H_n is a circulant, then n is a square.

It is clear that no normalized Hadamard matrix of order > 1 is a circulant, and it is conjectured that no Hadamard matrix of order $n > 4$ is a circulant. We remark that the definition of a cyclic projective plane implies that the points and lines can be labeled in such a way that the incidence matrix of the plane is a circulant.

The cyclic projective planes are closely related to a second subclass of symmetric block designs. A subset $D = \{d_1, d_2, \ldots, d_k\}$ of elements of Z_v is called a <u>(v,k,λ)-perfect difference set</u> if, for each $g \not\equiv 0$ in Z_v, the equation $d_i - d_j = g$ has exactly λ solutions from D. (We again assume that $0 < \lambda < k < v-1$.) Some examples are

$v=7,\ k=3,\ \lambda=1,\ D = \{0,1,3\}$,

$v=13,\ k=4,\ \lambda=1,\ D = \{0,1,3,9\}$,

$v=7,\ k=4,\ \lambda=2,\ D = \{0,1,3,6\}$.

Lemma 6. If D is a (v,k,λ)-perfect difference set, then $k(k-1) = \lambda(v-1)$.

Lemma 6 furnishes the basis of an equivalence between perfect difference sets and certain symmetric block designs.

Theorem 23. A (v,k,λ)-perfect difference set is equivalent to a (v,k,λ)-symmetric block design whose incidence matrix is a circulant.

To complete the equivalence, we define the e^{th} translate D_e of a difference set $D = \{d_1,d_2,\ldots,d_k\}$ to be the set $\{d_1+e,\ d_2+e,\ldots,d_k+e \pmod v\}$; clearly D_e is also a difference set. Then the sets A_i are defined to be the translates of D, that is, $A_v = D$, $A_1 = D_1,\ldots,A_{v-1} = D_{v-1}$, satisfy

1) $0 < \lambda < k < v-1$,

2) $|A_i| = k$,

3) Each element of Z_v occurs in exactly k of the sets A_i. Moreover, $|A_i \cap A_j| = \lambda$, since if d_m+i in A_i and d_n+j in A_j satisfy $d_m+i \equiv d_n+j \pmod v$, then $d_m-d_n \equiv j-i \pmod v$, which has λ solutions. Therefore the A_i form a (v,k,λ)-symmetric block design, whose incidence matrix is a circulant.

The converse construction produces a perfect difference set from a symmetric block design whose incidence matrix is a circulant.

Corollary 23a. A $(v,k,1)$-perfect difference set is equivalent to a cyclic projective plane of order k-1.

Another direct consequence of Theorem 23 is that the

Chowla-Ryser conditions are necessary for a perfect difference set. They are definitely not sufficient, however, since the triple (111,11,1) satisfies these conditions, but there is no (111,11,1)-perfect difference set, as we now show. An integer t is a multiplier of the difference set $D = \{d_1, d_2, \ldots, d_k\}$ if the set $\{td_1, td_2, \ldots, td_k\}$ is itself a difference set. Since the congruences $td_i = td_j \equiv 1 \pmod{v}$ must have a solution, a multiplier t must be relatively prime to v. The following theorem, known as Hall's multiplier theorem, gives conditions for a prime number p to be a multiplier.[20]

Theorem 24. Let D be a (v,k,λ)-difference set and let p be a prime divisor of $k - \lambda$ with $(p,v) = 1$ and $p > \lambda$. Then p is a multiplier of D.

If, in addition, t-1 is relatively prime to v, then there is a difference set D* fixed by t, that is, $tD^* = D^*$. To construct it, we note that t maps $D_u = \{d_1+u, \ldots, d_k+u\}$ onto $D_{s+tu} = \{d_1+s+tu, \ldots, d_k+s+tu\}$. Selecting u so that $(t-1)u \equiv -s \pmod{v}$, we obtain $D_u = tD_u$.

Now suppose that there is a (111,11,1)-difference set D. Then $k - \lambda = 10 = 2 \cdot 5$, and 2 is a multiplier of D. Further, (111,1) = 1, so there is a difference set D* fixed by 2. This means that if x is in D*, then $2^i x$ must also be in D*. Since there are more than 11 distinct powers of 2 (mod 111), D* must have more than 11 elements, which contradicts the fact that it is a (111,11,1)-difference set.

---

[20] Ryser, Combinatorial Mathematics, 139.

We have illustrated the (7,3,1) configuration as a complete set of orthogonal Latin squares, a cyclic projective plane, a Steiner triple system, a symmetric block design, and a Hadamard matrix. It is also the perfect difference set $D = \{0,1,3\}$ over Z_7.

The first actual construction of an infinite family of cyclic projective planes was published by Singer[21] in 1938.

Theorem 25. If $n = p^{\alpha}$, then there is a $(v,k,1)$-perfect difference set with $v = n^2+n+1$ and $k = n+1$.

Corollary 25a. There is a cyclic projective plane of every order $n = p^{\alpha}$.

We outline a proof of this thoerem, which involves the construction of a cubic equation which is irreducible over a field.[22] Let $n = p^{\alpha}$ and denote by G_3 the field $GF(p^{3\alpha})$. Let x be a generator of the multiplicative group G_3^* of G_3. Since G_3 is a cubic extension of $G_1=GF(p^{\alpha})$, then x is the root of an irreducible monic cubic equation over G_1. Every element $g \neq 0$ in G_3 can be written both as $g = ax^2 + bx + c$, where a,b,c are in G_1 and not all zero, and also as $g = x^d$ since G_3^* is cyclic. The set consisting of 1 and all elements of G_3 of the form $g_i = x + c_i$ for c_i in G_1, has $p^{\alpha}+1 = n + 1 - k$ elements, no two of which

---

[21] J.Singer, A theorem in projective geometry and some applications to number theory, Trans. Amer. Math. Soc. 43 (1938), 377-385.

[22] For more detail, see T.Storer, Cyclotomy and Difference Sets, Markham Publishing Co., Chicago, 1967, 16-18.

are linearly dependent, and each element can also be written as $g_i = x^{d_i}$ where $1 = x^0$. Then $d = \{o, d_1, \ldots, d_k\}$ is a perfect difference set (mod v).

We mention here briefly two other infinite families of difference sets.[23] In 1933 Paley[24] proved that for all primes $p \equiv 3 \pmod 4$ there is a difference set with $v = p$.

Theorem 26. If $p \equiv 3 \pmod 4$, then the set R of quadratic residues of Z_p is a $(p, \frac{p-1}{2}, \frac{p-3}{4})$-perfect difference set.

The following proof of Paley's theorem leads directly to a generalization of the theorem. We denote by $\underline{N}(R+1=R)$ the number of times a (quadratic) residue in Z_p is followed by a residue, and by $\underline{N}(N+1=N)$, $\underline{N}(R+1=N)$, $\underline{N}(N+1=R)$ the corresponding quantities. Since $p \equiv 3 \pmod 4$, -1 is a quadratic nonresidue (mod p), and on multiplying $\underline{N}(R+1=R)$ by -1, we obtain

(i) $$\underline{N}(R+1=R) = \underline{N}(-R-1=-R) = \underline{N}(N-1=N) = \underline{N}(N+1=n).$$

Since there are $\frac{p-1}{2}$ quadratic residues (mod p), and since 0 follows a nonresidue,

(ii) $$\underline{N}(R+1=R) + \underline{N}(R+1=N) = \frac{p-1}{2}.$$

Similarly,

(iii) $$\underline{N}(N+1=R) + \underline{N}(N+1=N) = \frac{p-1}{2} - 1 = \frac{p-3}{2}.$$

Now letting $\overline{R}$ be a residue such that $R\overline{R} = 1 \pmod p$ for some R,

---

[23] For a complete discussion of these families, we refer the reader again to the book by T.Storer.

[24] R.E.A.C.Paley, 311-320.

(iv) $\underline{N}(R+1=R) = \underline{N}(R-1=R) = \underline{N}(1-\bar{R}=R) = \underline{N}(N+1=R)$.

Combining (i) - (iv), we obtain

$$\underline{N}(R+1=R) = \underline{N}(N+1=R) = \underline{N}(N+1=N) = \frac{p-3}{4}$$

and

$$\underline{N}(R+1=N) = \frac{p+1}{4} .$$

Applying this information to the theorem, with d being any nonzero element in Z_p,

$$\begin{aligned} \underline{N}(R-R=d) &= \underline{N}(R\bar{d}-R\bar{d}=1) = \underline{N}(R\bar{d}+1=R\bar{d}) \\ &= \underline{N}(R+1=R) \quad \text{if d is in R} \\ &= \underline{N}(N+1=N) \quad \text{if d is not in R} \\ &= \frac{p-3}{4} . \end{aligned}$$

Thus R is a $(p, \frac{p-1}{2}, \frac{p-3}{4})$-perfect difference set.

The set of quadratic residues, denoted by C_0, and the set C_1 of quadratic nonresidues are called cyclotomic classes in Z_p; the numbers $\underline{N}(R+1=R)$, $\underline{N}(R+1=N)$, etc., are called cyclotomic numbers, and are denoted by (i,j) for $i,j = 0,1$, where (i,j) is the number of times an element in C_i is followed by an element in C_j. Within this framework, Lehmer[25] generalized Paley's theorem to provide necessary and sufficient conditions for a set of e^{th} residues (mod p) to be a perfect difference set, where $p = ef+1$. The cyclotomic classes relative to the e^{th} residues are given by $C_i = \{g^{es+i}, s=0,1,\ldots,f+1\}$, where g is a primitive root of p.

---

[25] E. Lehmer, On residue difference sets, Canad. J. Math. 5(1953), 425-432.

Theorem 27. The set R_e of e^{th} residues (mod p) is a $(p,\frac{p-1}{e},\frac{p-1-e}{e^2})$-perfect difference set if and only if $(i,0) = \frac{p-1-e}{e^2}$ for $i = 0,1,\ldots,e-1$.

Using Lehmer's theorem, one can verify without difficulty that the fourth powers in Z_{37}, $R_4=\{1,7,10,12,16,26,33,34\}$ form a (37,9,2)-perfect difference set. This theorem in itself does not, however, guarantee the existence of any difference sets. Paley's theorem determines when Lehmer's conditions are satisfied for e = 2. Chowla[26] has characterized the difference sets arising from the case e = 4; Lehmer[27] has shown that there are no difference sets of the Lehmer type when e = 6 and has given conditions for e = 8 to yield a perfect difference set. No perfect difference set of this type is known for e > 8.

A third family of difference sets can be constructed when p and q = p+2 are a pair of twin primes and g is a common primitive root. Then the sets $\{g^i,\ i=1,2,\ldots,\frac{(p-1)(q-1)}{2}\}$ and $\{0,q,2q,\ldots,(p-1)q\}$ together form a $(pq,\frac{pq-1}{2},\frac{pq-3}{4})$-perfect difference set. This is a special case of a theorem by Stanton and Sprott.[28]

---

[26] S.Chowla, A property of biquadratic residues, Proc. Nat. Acad. Sci. India 14(1944), 45-46.

[27] Lehmer, 425-235.

[28] R.G.Stanton and D.A.Sprott, A family of difference sets, Canad. J. Math. 10(1958), 73-77.

Theorem 28. If p and q are primes such that $p = q+2$ and g is a common primitive root of p and q, then the union of the sets $C_0 = \{g^i,\ i=1,\ldots,\frac{(p-1)(q-1)}{2}\}$ and $A = \{0,q,2q,\ldots,(p-1)q\}$ forms a $(pq,\frac{pq-1}{2},\frac{pq-1-e}{4})$-perfect difference set.

As an example, consider the twin primes p=3 and q=5. Then since 2 is a common primitive root, $C_0 = \{2^1,2^2,2^3,2^4 \equiv 1\}$ and $A = \{0,5,10\}$ combine to give the (15,7,3)-perfect difference set $D = \{0,1,2,4,5,8,10\}$.

Stanton and Sprott's theorem has been extended by Whiteman[29] to apply to all primes p and q.

Theorem 29. Let p and q be two primes with a common primitive root g, let $e = g.c.d.(p-1,q-1)$, and let $C_0 = \{g^i,\ i=1,2,\ldots,\frac{(p-1)(q-1)}{2}\}$, $A = \{0,q,2q,\ldots,(p-1)q\}$. Then $D = C_0 \cup A$ is a $(pq, \frac{pq-1}{e}, \frac{pq-1-e}{e^2})$-perfect difference set if and only if the following two conditions hold:

1) $q = (e-1)p + 2$,

2) $(i,0) = (e-1)(\frac{p-1}{e})^2$, for $i=0,1,\ldots,e-1$.

The existence of difference sets of the Whiteman type is in general unknown for $e > 8$. It is known[30] that there are none of this type when $e = 6$ or when $e = 8$.

A finite projective plane, being a set of points and a set of lines, has a natural "graphical" representation.

---

[29] A.L.Whiteman, A family of difference sets, Illinois J. Math. 6(1962), 107-121.

[30] J.W.Bergquist, Difference sets and congruences modulo a product of primes, Ph.D.dissertation, University of Southern California, 1963.

The "graphs" of π_2 and π_3 are shown in Figure 3. However, the "lines" in these graphs are (n+1)-tuples and do not correspond to the ordered pair concept of a line in graph theory, and, in general, these graphs have not proved useful in studying the properties of projective planes. A second graphical representation focuses attention on incidence, with the point set of the graph including both the lines and points of the plane and the lines of the graph indicating incidence between points and lines of the plane. This graph $G(\pi)$ is bipartite and regular of degree n+1, with $p = 2(n^2+n+1)$ and $q = (n+1)(n^2+n+1)$. Furthermore, its adjacency matrix A can be written in the form

$$\begin{pmatrix} 0 & B \\ B^t & 0 \end{pmatrix},$$

where B is the point-line incidence matrix of π. The bipartite graph associated with π_2 is also shown in Figure 3.

The projective planes of a given order have been characterized by using this bipartite representation together with a related graphical structure. The <u>line graph</u> L(G) of a graph G has as its points the lines of G. The line (u,v) is in L(G) when the lines u and v in G have a common endpoint. An example is shown in Figure 4. We denote the line graph of $G(\pi)$ by $L(\pi)$, and note that it has $(n+1)(n^2+n+1)$ points and is regular of degree 2n. In the first characterization of the line graph of a projective

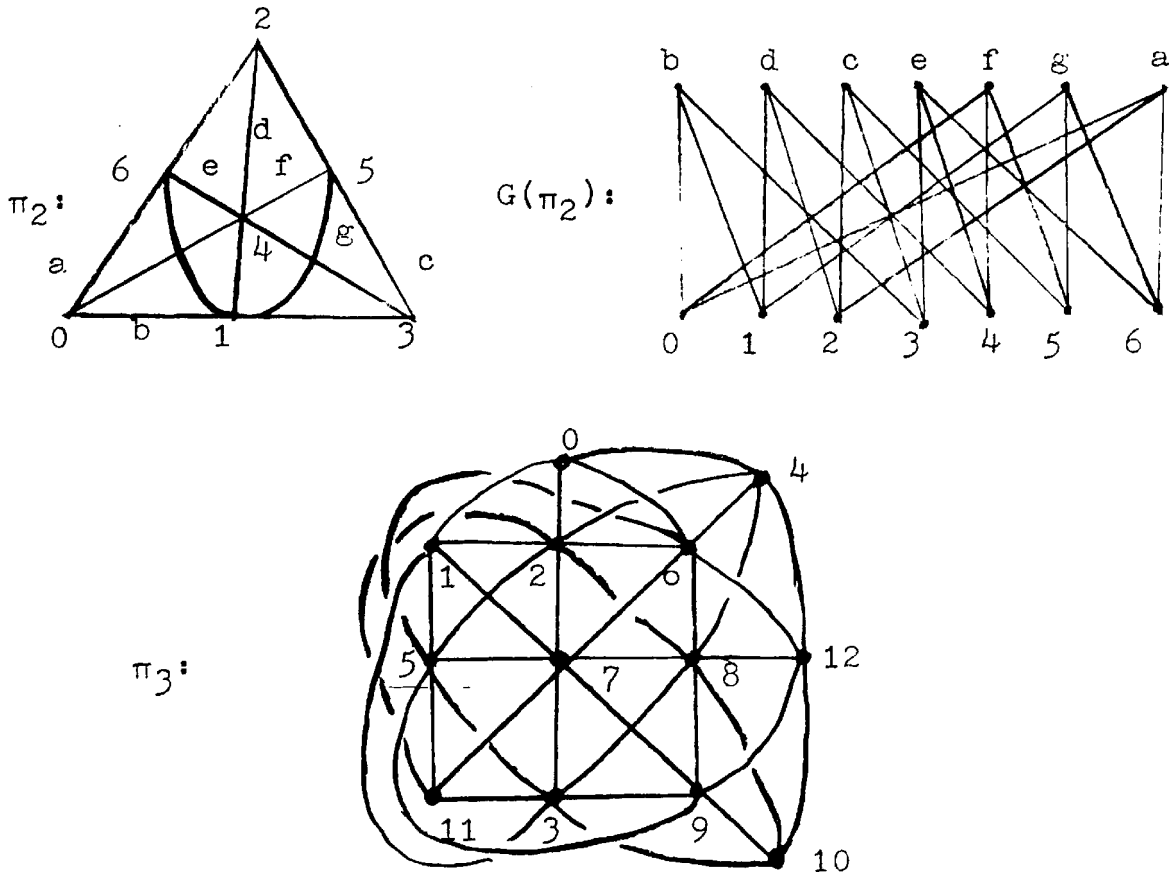


Figure 3. Graphs of π_2 and π_3

plane π_n of order n, Hoffman[31] showed that the set of distinct eigenvalues of the adjacency matrix of $G(\pi)$ completely determines the class of projective planes of order n, but does not distinguish between projective planes of the same order. The same theorem also holds when the eigenvalues of the adjacency matrix of $L(\pi)$ are considered.

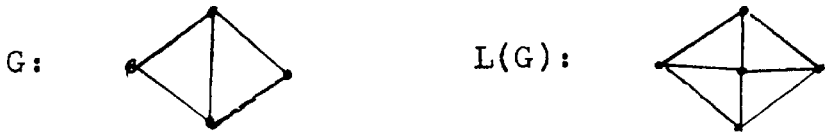


Figure 4. Graph and its line graph

---

[31]A.J.Hoffman, On the line graph of a projective plane, Proc. Amer. Math. Soc. 16(1965), 297-302.

Theorem 30. A connected graph, regular of degree $n+1$, with $2(n^2+n+1)$ points has the distinct eigenvalues $(n+1)$, $-(n+1)$, $\sqrt{n}$, $-\sqrt{n}$ if and only if G is the graph of a projective plane of order n.

A proof of this theorem relies heavily on classic results in matrix theory which we cannot present here. Briefly, the argument runs as follows. If π is a projective plane of order n, then the adjacency matrix A of its graph can be written in the form

$$A = \begin{pmatrix} 0 & B \\ B^t & 0 \end{pmatrix},$$

where B is the point-line incidence matrix of π. Therefore, the eigenvalues of A are the eigenvalues of B and their negatives. It is known that the distinct eigenvalues of B are $n+1$ and $\sqrt{n}$.

Another matrix argument yields the eigenvalues of the adjacency matrix of the line graph of π.

Theorem 31. The distinct eigenvalues of $A(L(\pi))$ are $2n$, -2, $n-1\pm\sqrt{n}$.

The proof that these eigenvalues characterize the line graph of a projective plane of order n involves first a consideration of the maximal complete subgraphs, called cliques.

Theorem 32. If G is a connected graph, regular of degree $2n$, with $(n+1)(n^2+n+1)$ points which has $2n$, -2, $n-1\pm\sqrt{n}$ as its distinct eigenvalues, contains $2(n^2+n+1)$ cliques $C_1,\ldots,C_{2(n^2+n+1)}$ with the property that

1) Each C_i contains exactly n+1 points.

2) Each point of G is in exactly two C_i.

3) Each pair of adjacent points is in exactly one C_i.

We now form a new graph $\tilde{G}$ whose points are the cliques of G and say that two points are adjacent in $\tilde{G}$ if their corresponding cliques in G have a common point. This graph is connected, regular of degree n+1, and has the property that $L(\tilde{G}) = G$. A matrix argument again shows that $\tilde{G}=G(\pi)$ where π is a projective plane of order n.

Theorem 33. If G is a connected graph, regular of degree 2n, with $(n+1)(n^2+n+1)$ points, then A(G) has 2n, -2, $n-1\pm\sqrt{n}$ as its distinct eigenvalues if and only if G is the line graph of some projective plane of order n.

Another characterization of the line graph of a projective plane in terms of graphical properties has been found by Laskar and Dowling.[32] We define $\Delta(u,v)$ to be the number of points of a graph adjacent to both u and v. If u is adjacent to v, $\Delta(u,v)$ is called the line degree of the line (u,v).

Theorem 34. The graph G is the line graph of a projective plane of order n if and only if it has the following properties:

1) $p = (n+1)(n^2+n+1)$.

2) G is connected and regular of degree 2n.

3) G is line-regular of degree n-1.

---

[32] R.Laskar and T.A.Dowling, A geometric characterization of the line graph of a projective plane, Institute of Statistics Mimeo Series No. 516, University of North Carolina at Chapel Hill, 1967.

4) $\Delta(u,v) = 1$ if $d(u,v) = 2$.

5) If $d(u,v) = 2$, then the number of points w such that $d(u,w) = 1$ and $d(v,w) = 2$ is at most n-1.

The proof of the necessity of these properties is straightforward. Properties 1 and 2 are immediate. For property 3, if u adj v, the $u = (P, \ell)$, $v = (P_1, \ell_1)$ in $G(\pi)$, where P, P_1 are points of π, and ℓ, ℓ_1 are lines of π, and either $P = P_1$ or $\ell = \ell_1$. If $P = P_1$, then the n-1 points $w = (P_1, \ell_j)$ are adjacent to both u and v and no other point is. Thus $\Delta(u,v) = n-1$. If $\ell = \ell_1$, then $\Delta(u,v) = n-1$ by duality.

To prove property 4, we note that if $d(u,v) = 2$, then $u = (P_1, \ell_1)$, $v = (P_2, \ell_2)$ in $G(\pi)$, and $P_1 \neq P_2$, $\ell_1 \neq \ell_2$. The only points of $L(\pi)$ which could share an endpoint with both u and v in $G(\pi)$ are (P_1, ℓ_2) and (P_2, ℓ_1). At least one of these is in $G(\pi)$ since $d(u,v) = 2$. However, both cannot be in $G(\pi)$ since (P_1, ℓ_2) and (P_2, ℓ_1) imply that P_1 is on ℓ_2 and P_2 is on ℓ_1. But P_1 is on ℓ_1, P_2 on ℓ_2 would then imply that $\ell_1 = \ell_2$.

For property 5, suppose that $d(u,v) = 2$. Then again $u = (P_1, \ell_1)$, $v = (P_2, \ell_2)$, $P_1 \neq P_2$, $\ell_1 \neq \ell_2$, and either P_1 is on ℓ_2 or P_2 is on ℓ_1. If P_2 is on ℓ_1, then the n-1 points of the form (P, ℓ_1), where $P \neq P_1, P_2$, are adjacent to (P_1, ℓ_1) is of the form (P_1, ℓ), $\ell \neq \ell_1$. If it were at a distance 2 from (P_2, ℓ_2) then P_2 would be on ℓ, implying that $\ell = \ell_1$ or P_1 would be on ℓ_2 which is again a contradiction. If P_1 is on ℓ_2, then the n-1 points (P_1, ℓ)

$\ell \neq \ell_1, \ell_2$ are adjacent to (P_1, ℓ_1) and at a distance 2 from (P_2, ℓ_2). Any other point adjacent to (P_1, ℓ_1) would be of the form (P, ℓ_1), $P \neq P_1, P_2$. If it were at a distance 2 from (P_2, ℓ_2), then P would be on ℓ_2 (since P_2 is not on ℓ_1). But P_1 on ℓ_2 implies that $\ell_1 = \ell_2$, which is a contradiction.

The proof of the converse of Theorem 34 is considerably more complicated. The general outline of the proof is that of the Hoffman proof in that it characterizes the cliques of a graph and considers the clique-condensation graph. This is accomplished, however, without the use of matrix methods, but rather by graphical techniques.

VII. MISCELLANEOUS MATHEMATICAL APPLICATIONS

Graph theory can be and has been applied fruitfully to many mathematical topics other than the ones which we have already described. It is not feasible, however, to attempt to discuss each of these within this thesis. Instead, we have chosen a few additional topics which we find important, striking, or curious, as examples of how graphical concepts are applicable. These topics cover a wide range of interests, including probability, logic, number theory, topology, and mathematical amusement.

1. Markov chains

Relagraphs are in common usage in probability theory where they provide the diagrammatic description of a Markov chain. In Figure 1, the relagraph of one of the most familiar types of Markov chain, sometimes called the "gambler's ruin", shows the possibilities available to either of two boys who are matching pennies, each starting with two cents. The probability of going from one position to another is written on each arc, and the loops at +2 and -2, having probability 1, indicate that one of the boys is bankrupt and the game is ended.

Formally, a Markov chain is a set S of n states s_1, $s_2,\ldots,s_n$, together with an initial probability distribution,

that is, the probabilities of starting in each state, and an n x n transition matrix $P = ((p_{ij}))$, where p_{ij} is the probability of going from s_i to s_j in one step. The transition matrix of the Markov chain shown in Figure 1 is

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Clearly P is the "weighted" adjacency matrix of the relagraph. Applying Theorem I.2, we easily obtain the following fundamental theorem on Markov chains.

Theorem 1. The probability of going from state i to state j in n steps in a Markov chain M with transition matrix P is the (i,j)-entry of P^n.

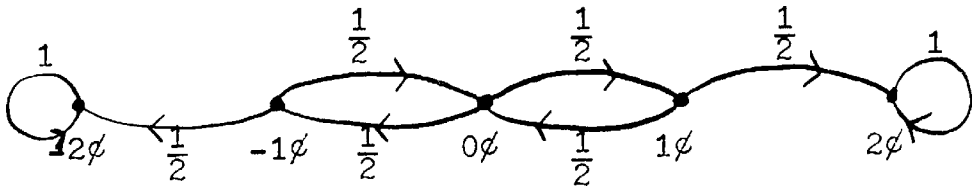


Figure 1. Relagraph of a Markov chain

Several special types of Markov chains --regular, ergodic, absorbing, irreducible--are easily visualized when their defining properties are given graphical interpretations.[1] A Markov chain is ergodic if it is possible

[1]The definitions of Markov chain properties have been taken from J.G.Kemeny and J.L.Snell, Finite Markov Chains, Van Nostrand, New York, 1960, and we refer the reader to this book for further information about Markov chains.

to go from every state to every other state. In graphical terms, this means that the relagraph of a Markov chain is strongly connected. A Markov chain is regular if some power t of its transition matrix has all positive entries; thus not only is the relagraph strongly connected, but for some t there is a path of length t between every pair of points. The relagraph R_1 of a Markov chain which is regular and R_2 which is ergodic (but not regular) are shown in Figure 2.

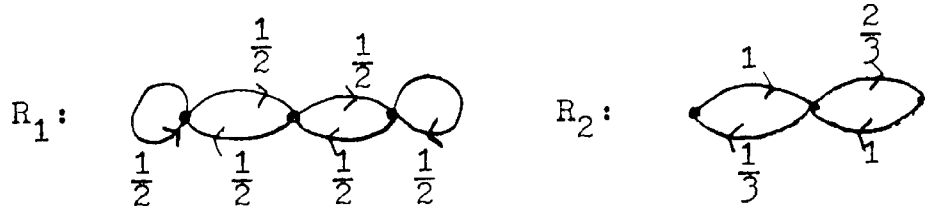


Figure 2. Relagraph R_1 of a regular Markov chain and relagraph R_2 of an ergodic Markov chain

In the gambler's ruin example, once a participant is bankrupt, it is impossible for him to continue to play. In any Markov chain a state is absorbing if the probability of leaving it is 0. The point which represents this state has only one outgoing arc, and this is a loop. A Markov chain itself is called absorbing if it has at least one absorbing state and it is possible to go from any state to an absorbing state. This implies that every unilateral component of the relagraph must contain an absorbing state. The relagraph in Figure 3, with arcs $x_1, x_2, \ldots, x_6$ in one unilateral component and arcs x_7 and x_8 in the other, represents an absorbing Markov chain.

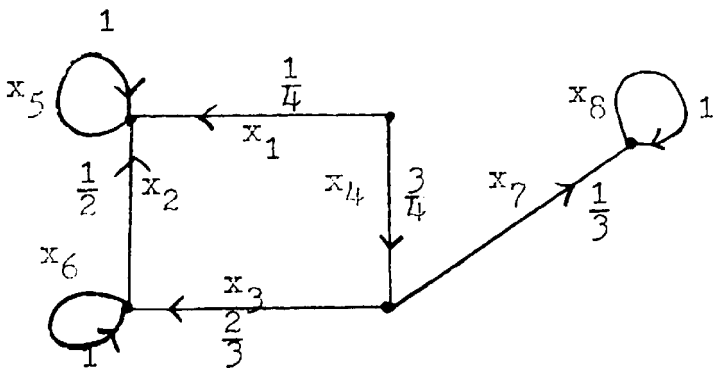


Figure 3. Relagraph of an absorbing Markov chain

A set T of states in a Markov chain is closed if no state outside T can be reached from any state in T. The smallest closed set containing any subset U of S is called the closure of U. Clearly, if a single state is closed, it is absorbing. In general, the closure of a single state is the set of states reachable from it. If there is no closed set other than the set S of all states, then the Markov chain is irreducible. In graph-theoretical terms this simply means that every pair of states is mutually reachable, that is, the relagraph of an irrdeucible Markov chain is strongly connected.

With this intimate relationship between graph theory and the diagrams and definitions of Markov chain theory, it would seem that the theorems about Markov chains should have some graphical significance. With the exception of Theorem 1, however, this is not the case; the majority of questions raised have concerned the probabilistic nature of Markov chains--properties of transition possibilities, rather than underlying structure. Perhaps a future

investigation will find the graphical framework in which Markov chains may be viewed to be of greater value.

2. Ramsey numbers

In the search for Ramsey numbers, graph theory has been applied with definite success. New Ramsey numbers have been found and better bounds have been established for others as the result of graphical constructions. Ramsey numbers first arose in logic in a question involving a set S, its set of r-subsets (subsets containing r elements, r a natural number) called $P_r(S)$, and an arbitrary partition of $P_r(S)$ into t components $A_1,\ldots,A_t$, some of which may be empty. Letting $q_1,\ldots,q_t$ be t numbers each greater than or equal to r, we say that there is a (q_i,A_i)-subset of S if S has a subset all of whose r-subsets are contained in A_i. In a theorem[2] on the existence of (q_i,A_i)-subsets, Ramsey numbers were introduced.

Theorem 2. There exists a minimal positive integer $N(q_1,\ldots,q_t,r)$ such that every set S with cardinality greater than or equal to N contains a (q_i,A_i)-subset for some $i = 1,2,\ldots,t$.

The number $N(q_1,\ldots,q_t,r)$ is called a Ramsey number, and its existence is assured by Theorem 2. However, with the exception of the following general formulas, very few of their exact values are actually known.

---

[2]T.P.Ramsey, On a problem of formal logic, Proc. London Math. Soc. 30(1929), 264-286.

$$N(q_1,q_2,1) = q_1 + q_2 - 1$$

$$N(q_1,r,r) = q_1$$

$$N(r,q_2,r) = q_2$$

The only other known[3] Ramsey numbers are $N(3,3,3,2) = 17$ and those given[4] in Table 1 in which $t = 2$ and $r = 2$.

q_2 =	q_1 = 1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	2	3	4	5	6	7
3	1	3	6	9	14		
4	1	4	9	18			
5	1	5	14				
6	1	6					
7	1	7					

Table 1. Ramsey numbers with t=2 and r=2

The case t=2 and r=2 can be immediately interpreted in terms of graph theory, by viewing the set S as a point set of a graph G and by letting the partition $A_1 \cup A_2$ of the 2-subsets of S determine the lines of G; here A_1 is the set of adjacent pairs of points and A_2 the set of non-adjacent pairs of points. Ramsey's theorem for the case t=2 is restated in a graphical context in Theorem 3.

Theorem 3. There is a minimal positive integer $N(q_1,q_2,2)$ such that every graph with at least N points has either a K_{q_1} or a $\overline{K}_{q_2}$ subgraph.

---

[3] J.G.Kalbfleisch mentions in On an unknown Ramsey number, Michigan Math. J 13(1966), 385-392, that J.E.Graver and J.Yackel claim to have shown that $N(3,7,2) = 24$.

[4] R.E.Greenwood and A.M.Gleason, Combinatorial relations and chromatic graphs, Canad.J.Math. 7(1955),4.

Several graph theorists--Erdos,[5] Szekeres,[6] Greenwood and Gleason,[7] Kalbfleisch,[8] Behzad and Radjavi[9]--have studied this version of Ramsey's theorem, establishing successively better bounds for the Ramsey numbers. A listing of these bounds would be both uninteresting and unilluminating, and therefore we mention only the following one, proved by Behzad and Radjavi,[10] which will enable us to establish the value of three Ramsey numbers.

Theorem 4. $N(3,s,2) \leq \frac{s^2+3}{2}$.

Using this bound, we now employ graphical methods to prove that $N(3,3,2) = 6$, $N(3,4,2) = 9$, and $N(3,5,2) = 14$. We note that this also establishes that $N(4,3,2) = 9$ and $N(5,3,2) = 14$, the duality being obvious in the graphical context by taking the complements of graphs involved. In Figure 4 we show a graph with 5 points which has no K_3 or $\overline{K_3}$ subgraph, a graph with 8 points having no K_3 or $\overline{K_4}$

---

[5]P.Erdos has written several articles on the subject. A few are Some remarks on the theory of graphs, Bull. Amer. Math. Soc. 53(1947), 292-294; Remarks on a theorem of Ramsey, Bull.Res. Council Israel 7(1957), 21-24; Graph theory and probability I, Canad. J. Math. 11(1959), 34-38; and Graph theory and probability II, Canad. Math. J. (1961), 346-352.

[6]P.Erdos and G.Szekeres, A combinatorial problem in geometry, Compositio Math. 2(1935), 463-470.

[7]Greenwood and Gleason, 1-7.

[8]J.G.Kalbfleisch, Construction of special edge-chromatic graphs, Canad. Math. Bull. 8(1965), 575-584, and On an unknown Ramsey number, 385-392.

[9]M.Behzad and H.Radjavi, The line analog of Ramsey numbers, Israel J. Math. 5(1967), 93-96, and On a problem of Erdos, unpublished manuscript, n.d.

[10]Behzad and Radjavi, On a problem of Erdos, 2.

subgraph, and a graph with 13 points having no K_3 or $\overline{K_5}$ subgraph. Coupling these graphs with Theorem 4, we have shown that $5 < N(3,3,2) \leq 6$, $8 < N(3,4,2) \leq 9$, and $13 < N(3,5,2) \leq 14$, and thus the Ramsey numbers are 6, 9, and 14.

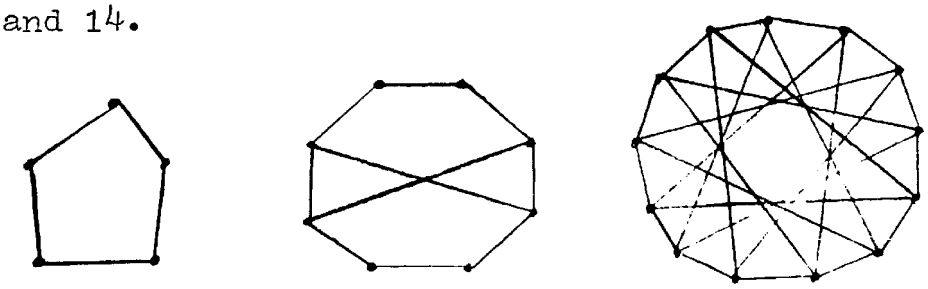


Figure 4. Graphs determining the Ramsey numbers N(3,3,2), N(3,4,2), and N(3,5,2)

Kalbfleisch[11] has developed a specific graphical construction for establishing lower bounds for Ramsey numbers with which he has proved that $N(3,6,2) = 18$, $N(3,7,2) \geq 23$, $N(4,5,2) \geq 25$, and $N(3,8,2) \geq 27$. He approaches the problem in terms of a coloring of the lines of a complete graph--a form which is easily generalized to the case for r=2 with t unrestricted. A (q_1,q_2)-coloring of K_n is a coloring of the lines with red or blue so that no red K_{q_1} and no blue K_{q_2} subgraph occurs. Thus Ramsey's theorem asserts that there is an $N(q_1q_2,2)$ such that for $n \geq N$, no (q_1,q_2)-coloring of K_n exists. Kalbfleisch's method for constructing graphs with (q_1,q_2)-colorings is too lengthy to describe here, but it is noteworthy that he has programmed the method for a computer, and using it, he has constructed a (3,7)-coloring of K_{21}, a (4,5)-coloring of

---

[11]Kalbfleisch, Construction of special edge-chromatic graphs, 575-584.

K_{24}, and a (3,8)-coloring of K_{26}.

Using the idea of a (q_1,q_2)-coloring, we may now restate Ramsey's theorem in graphical terms for the case r=2 with no restriction on the value of t, by defining a $(q_1,q_2,\ldots,q_t)$-coloring of K_n to be a coloring of the lines of K_n with t colors $c_1,c_2,\ldots,c_t$, in such a way that there is no K_{q_i} subgraph with color c_i, for any $i = 1,2,\ldots,t$. A (q_i,A_i)-subset of S then corresponds to a K_{q_i} subgraph with color c_i. This form of Ramsey's theorem is given in Theorem 5. We note, however, that it seems to provide no outstanding simplification of the problem, and no new results have come from it.

Theorem 5. There is a minimal positive integer $N(q_1,q_2,\ldots,q_t,2)$ such that for $n \geq N$, every coloring of the lines of the graph K_n with colors $c_1,c_2,\ldots,c_t$ has a subgraph K_{q_i} of color c_i for some $i = 1,2,\ldots,t$.

The graphical statement of Ramsey's theorem has inspired related graphical questions which we mention briefly. Behzad and Radjavi[12] have defined the number N'(r,s) to be the smallest positive integer such that every connected graph with p points contains either r mutually adjacent lines or s mutually disjoint lines. In other words, the number N'(r,s) is the smallest number of points in a graph G which insures that the line graph of G has a subgraph of the form K_r or $\overline{K_s}$. Surprisingly enough, this number is easily found, and we give Behzad and Radjavi's

---

[12]Behzad and Radjavi, The line analog of Ramsey numbers, 93-96.

result in Theorem 6.

Theorem 6. For $r > 2$, $s > 1$, $N'(r,s) = (r-1)(s-1)+2$, and if $r = 2$, $s > 1$, $N'(2,s) = 3$.

As an illustration of this theorem, the graphs in Figure 5 show that $N'(3,3) > 5$ and $N'(4,3) > 7$.

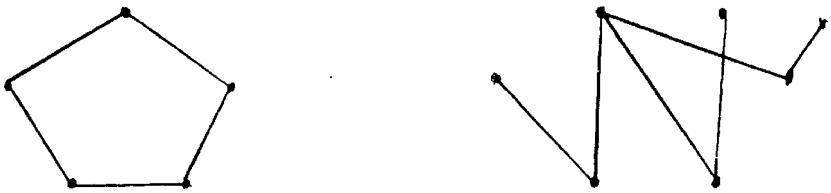


Figure 5. Graphs illustrating $N'(3,3)>5$ and $N'(4,3)>7$

In another study, Behzad and Radjavi[13] define the number $N''(r,s)$ to be the smallest positive integer such that every graph has either r mutually "joint" or s mutually disjoint elements, where an element is either a point or a line, and "joint" means "adjacent" (for two points or two lines) or "incident" (for a point and a line). This number also is known for all values of r and x, as given in Theorem 7.

Theorem 7. For all s, $N''(1,s) = 1$, $N''(2,s) = N''(3,s) = s$, and for $r \geq 4$, $N''(r,s) = 2s - 1$.

From Theorem 7 we know that $N''(4,4) = 7$, and we show by the graph in Figure 6 that necessarily $N''(4,4) \geq 7$.

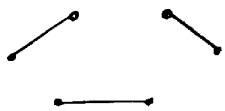


Figure 6. Graph showing $N''(4,4) \geq 7$

---

[13] M. Behzad and H. Radjavi, Another analog of Ramsey numbers, unpublished manuscript, n.d.

3. Number theory

The set of all positive integers less than a given integer m and relatively prime to m form a finite abelian group of order $\phi(m)$. It is designated by M_m. Each element in the group M_m belongs to at least one cycle, and the cycle graph of M_m is the graph of a minimal set of cycles containing all the points of M_m. Five examples are shown in Figure 7. We note that the entire multiplicative structure of the group is visible in the cycle graph.

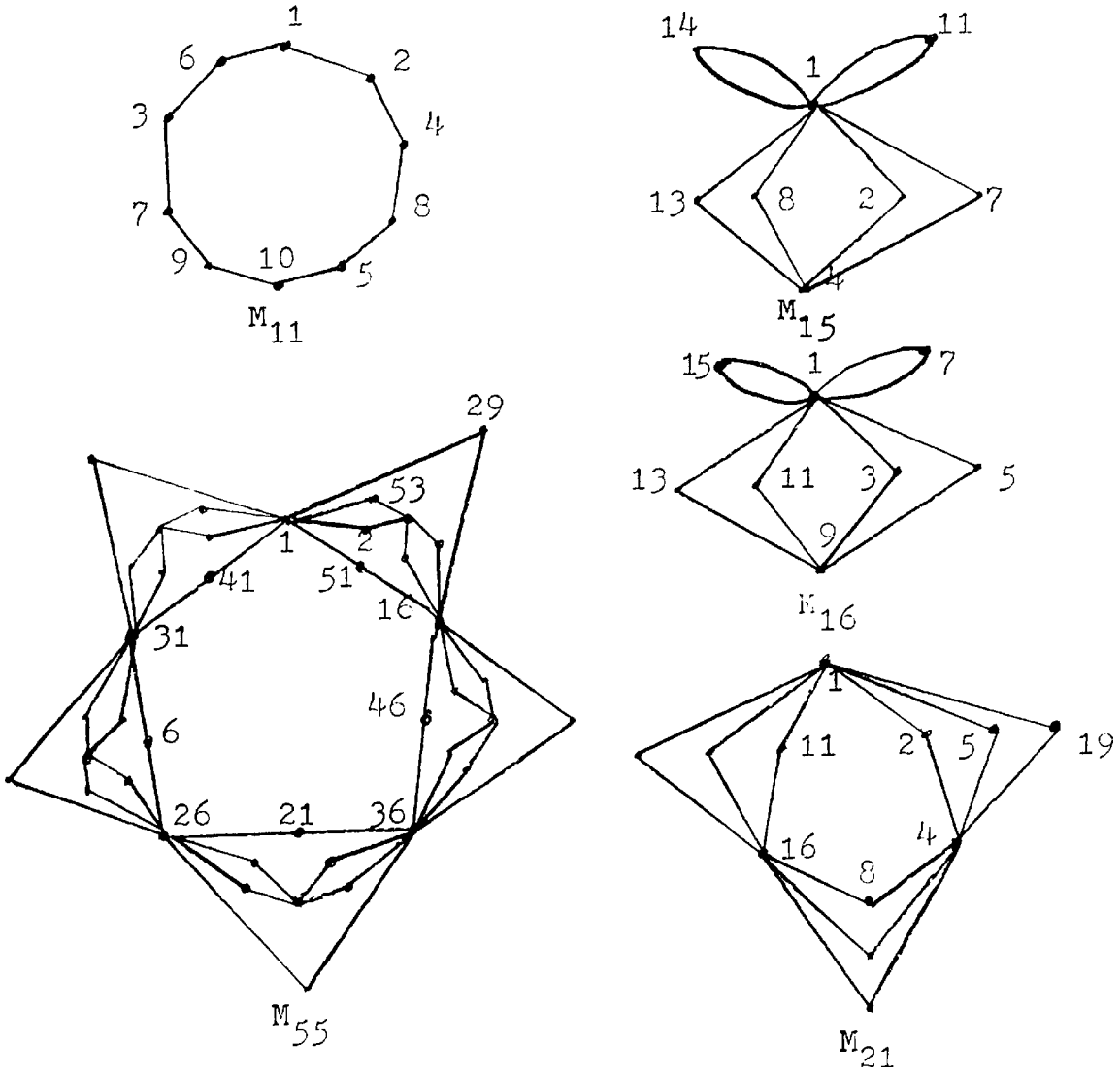


Figure 7. Cycle graphs of $M_{11}, M_{15}, M_{16}, M_{55}, M_{21}$

A complete characterization of the cycle graphs of M_m has been given[14] in terms of a factorization of m. For each odd prime p, we can write $\phi(p^i)$ as

$$\phi(p^i) = \left\langle q_1^{b_1} \right\rangle \cdot \left\langle q_2^{b_2} \right\rangle \cdot \cdots \left\langle q_s^{b_s} \right\rangle \cdot \left\langle p^{i-1} \right\rangle ,$$

where $q_1^{b_1} \cdots q_s^{b_s}$ is the prime factorization of p-1, and the symbol $\left\langle q_i^{b_i} \right\rangle$ signifies that this factor is to be written as a single number, not as a product, e.g., $\left\langle 3^2 \right\rangle$ is to be written as 9. If p=2, then we write $\phi(4)=2$, $\phi(8)=2\cdot 2$, $\phi(16)=2\cdot 4$, and, in general, $\phi(2^i)=2\cdot \left\langle 2^{i-2} \right\rangle$ for $i \geq 3$. Combining these forms we have a specific factorization (up to order) of $\phi(m)$ which is denoted by ϕ_m. For example,

$$\phi_{11} = 2\cdot 5 ,$$
$$\phi_{15} = 2\cdot 4 ,$$
$$\phi_{16} = 2\cdot 4 ,$$
$$\phi_{55} = 4\cdot 2\cdot 5 ,$$
$$\phi_{63} = 2\cdot 3\cdot 2\cdot 3 \quad .$$

Letting $\phi_m = \left\langle p_1^{1} \right\rangle \cdots \left\langle p_k^{k} \right\rangle$, we construct a second factorization Φ_m of $\phi(m)$ by first taking the largest power of each prime which occurs explicitly in ϕ_m and multiplying them together. Removing these factors from ϕ_m, we obtain the new expression $\phi_m' = \left\langle q_1^{1} \right\rangle \cdots \left\langle q_j^{j} \right\rangle$. The next factor of Φ_m is obtained from ϕ_m' by multiplying together the largest power of each prime occurring explicitly in ϕ_m'. Continuing in this fashion until all factors

---

[14] D. Shanks, Solved and Unsolved Problems in Number Theory, vol. 1, Spartan Books, Washington, D.C., 1962, 83-104.

of ϕ_m have been used, we arrive at the factorization $\Phi_m = f_1 f_2 \cdots f_r$, where $f_1 \mid f_2 \mid \cdots \mid f_r$. The f_i are called characteristic factors of M_m. For the examples above,

$$\Phi_{11} = 10 ,$$
$$\Phi_{15,16} = 2 \cdot 4 ,$$
$$\Phi_{55} = 2 \cdot 20 ,$$
$$\Phi_{63} = 6 \cdot 6 .$$

The factorization Φ_m completely determines the structure of the cycle graph.

Theorem 8. If Φ_m is the product of r characteristic factors f_i, then for each f_i there is a residue class g_i of order f_i (mod m) such that every residue class a_j in M_m can be expressed as $a_j \equiv g_1^{s_{1,j}} g_2^{s_{2,j}} \cdots g_r^{s_{r,j}} \pmod m$ with $0 \leq s_{i,j} < f_i$ in one and only one way. We say that M_m is the direct product of the r cycles of the g_i.

Before discussing the proof of the theorem, we list several corollaries which describe the groups M_m.

Corollary 8a. M_m and M_n are isomorphic if and only if Φ_m and Φ_n are identical.

Corollary 8b. M_m is cyclic if and only if $m = 2, 4, p^n$, or $2p^n$.

Corollary 8c. If f_r is the largest characteristic factor of M_m, then $a^{f_r} \equiv 1 \pmod m$ for every residue class a in M_m. Furthermore, there is an element in M_m whose order is f_r.

Corollary 8d. If $m > 2$, and M_m has r characteristic factors, then m has $\phi(m)2^{-r}$ quadratic residues and each of these has 2^r square roots.

Corollary 8e. Every finite abelian group is isomorphic to a subgroup of M_m for infinitely many different values of m.

The proof of this theorem is a lengthy number-theoretic argument. We outline it by stating the lemmas which comprise the proof. The first lemma is a well-known property of prime powers.

Lemma 1. If $m = 2, 4$, or p^n, where p is an odd prime, then m has a primitive root and M_m is cyclic.

For the higher powers of 2, we note that

$$3^2 = 1 + 8 ,$$
$$3^4 = 1 + 16 + 32t ,$$
$$3^8 = 1 + 32 + 64u ,$$

and inductively, 3 is of order $2^{n-2} \pmod{2^n}$ for $n \geq 3$. Furthermore, the negatives of the powers of 3, the numbers $a_k \equiv -3^{t_k} \pmod{2^n}$, are not congruent to any of the 2^{n-2} powers 3^{t_j} for otherwise $-1 \equiv 3^{|t_j - t_k|} \pmod{2^n}$. This would imply that $8 \mid 3^{|t_j - t_k|} + 1$, but $3^t + 1 \equiv 2$ or $4 \pmod 8$ for all t. This provides a chacterization of M_{2^n}.

Lemma 2. If $m = 2^n$ with $n \geq 3$, then $\Phi_m = 2 \cdot \langle 2^{n-2} \rangle$ and every element a_i of M_m can be written

$$a_i \equiv (-1)^{s_i} 3^{t_i} \pmod{2^n},$$

where (-1) is of order 2 and 3 is of order $2^{n-2} \pmod{2^n}$.

We now combine Lemmas 1 and 2 to devise a direct

product representation of M_m for composite m.

Lemma 3. If the $\phi(A)$ residue classes a_i in M_A are written

$$a_i \equiv g_1^{1,i} g_2^{2,i} \cdots g_n^{n,i} \pmod{A},$$

where g_j is of order m_j, and the $\phi(B)$ classes b_i in a cyclic M_B, with B prime to A, are written

$$b_i \equiv g^{i} \pmod{B},$$

then the $\phi(AB)$ classes c_i in M_{AB} can be written

$$c_i \equiv h_1^{1,i} h_2^{2,i} \cdots h_n^{n,i} h^{i} \pmod{AB},$$

where the factor generator h_j is of order m_j, and h is of order $\phi(B) \pmod{AB}$.

We illustrate Lemma 3 by constructing M_{105}. First we construct M_{15}. Since $15 = 3 \cdot 5$, we write $M_3 = \{2, 2^2=1\}$ and $M_5 = \{2, 2^2=4, 2^3=3, 2^4=1\}$. Then by Lemma 3, M_{15} can be written as a direct product of two cycles, one of order 2, the other of order 4 (mod 15), that is, $M_{15} = \{11^{\alpha} 13^{\beta} \pmod{15}\}$. To construct a direct product representation of M_{105}, we combine M_{15} with the cyclic M_7 and obtain $M_{105} = \{29^{\alpha} 13^{\beta} 19^{\gamma} \pmod{105}\}$, where 29 is of order 3, 13 of order 4, and 19 of order 6 (mod 105).

While Lemma 3 provides a constructive method for expressing M_m as a direct product, the result is not the standard form given in the theorem, in which the orders of the generators are the characteristic factors. Lemma 4 furnishes the final step in the proof of Theorem 8 by proving that the cycles of Lemma 3 can be decomposed into cycles of prime power order and then recombined into cycles whose

orders are the characteristic factors.

Lemma 4. If g is of order AB (mod m) with (A,B) = 1, the AB residue classes g^c ($0 \leq c < AB$) can be written as a direct product

$$g^c \equiv s^a t^b \pmod{m},$$

where s is of order A and t of order B (mod m). Conversely, given two residue classes s and t of orders A and B with (A,B) = 1, the AB classes $s^a t^b$ are all distinct and constitute the cycle of some g of order AB (mod m).

Applying Lemma 4 to $M_{105} = \{29^\alpha 13^\beta 19^\gamma\}$, we decompose the cycle $\{19^\gamma\}$ of order 6 into $\{34^\gamma 46^\delta\}$, where 34 is of order 2 and 46 of order 3 (mod 105). This gives the representation of M_{105} as $\{29^\alpha 13^\beta 34^\gamma 46^\delta\}$, where the orders of the generators are the factors of $\phi_{105} = 2\cdot4\cdot2\cdot3$. combining the cycles of orders 3 and 4, we finally achieve $M_{105} = \{29^\alpha 34^\beta 73^\gamma\}$, where the orders of the generators are the characteristic factors of $\Phi_{105}=2\cdot2\cdot12$.

A pictorial description of the cycle graph of M_m can also be given in terms of the factorization Φ_m. The blocks of the graph correspond to the characteristic factors which are powers of 2, and these may be drawn easily from the group structure. Several of these graphs are shown in Figure 8, where the identity element is depicted both at the top and the bottom of each of these examples, for reasons which will be clear below.

If Φ_m has at most one characteristic factor which is not a power of 2, that is, if

$$\Phi_m = \langle 2^a \rangle \langle 2^b \rangle \cdots \langle 2^y \rangle \langle 2^{z_N} \rangle ,$$

where N is an odd integer ≥ 1, then the cycle graph of M_m consists of N "lobes" each corresponding to $\langle 2^2 \rangle \langle 2^b \rangle \cdots \langle 2^y \rangle \langle 2^z \rangle$ joined cyclically. See M_{55} and M_{21} in Figure 7.

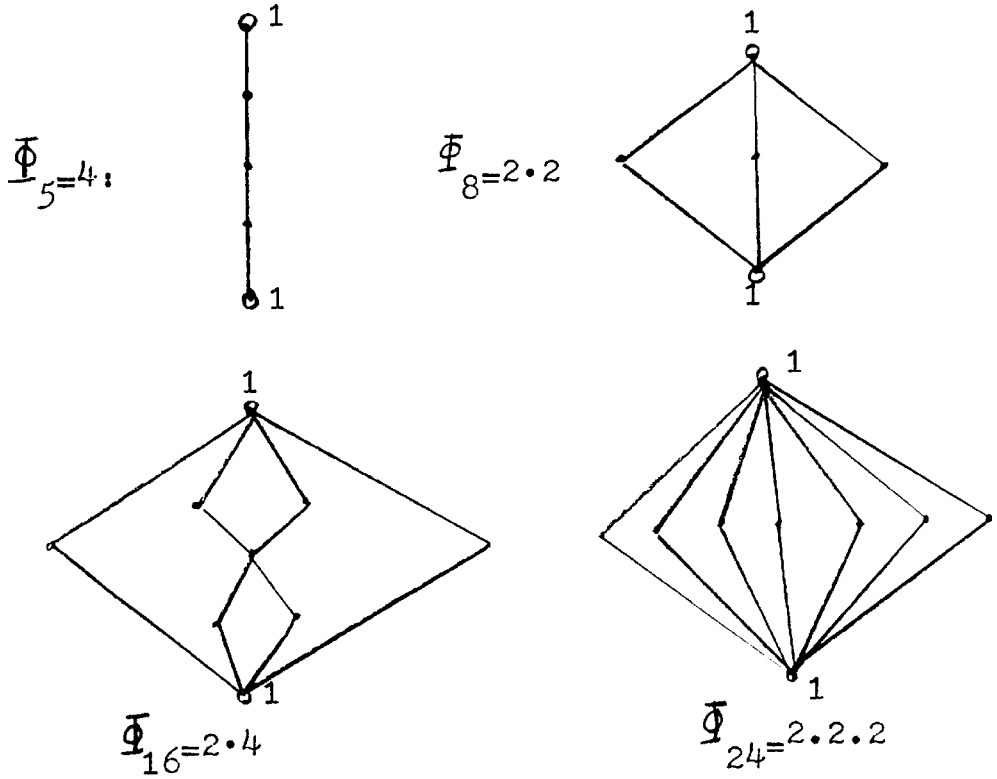


Figure 8. Cycle graphs of M_5, M_8, M_{16}, M_{24}

If Φ_m has more than one factor f_i which is not a power of 2, then the cycle graph of M_m is no longer planar. The smallest such cycle graph is M_{63} with $\Phi_m = 6 \cdot 6$. It is shown in Figure 9.

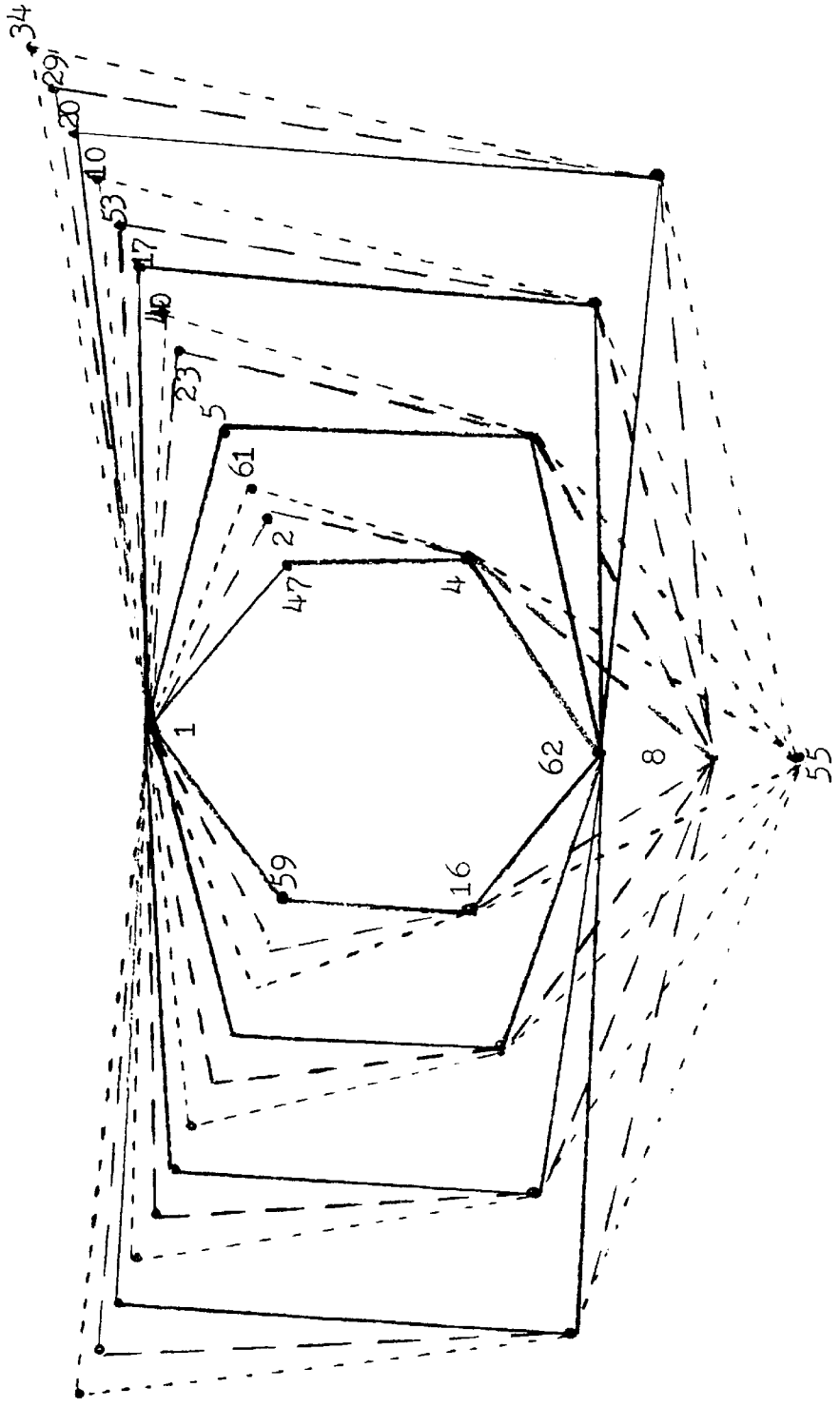


Figure 9. Cycle graph of M_{63}

4. Topology

There is an obvious connection between graph theory and topology; in fact, both trace their origins back to Euler's solution of the Königsberg bridge problem. The study of 1-complexes is closely related to the study of graphs, although for the most part, topologists have been interested in properties which have significance only for infinite sets, whereas graph theorists have restricted themselves largely to finite sets. There is another aspect of topology, however, to which graph theory can also be related. There is a direct correspondence between topologies on finite sets and transitive digraphs which has been studied by Anderson and Chartrand.[15]

A quick review of the axiom system of a topological space is necessary first. A topological space S consists of a set V together with a collection $\mathcal{O} = \{O_\alpha\}$ of subsets of V, called "open" sets, which satisfy the following axioms.

A1. The sets $\emptyset$ and V are in $\mathcal{O}$.

A2. The union of any number of sets in $\mathcal{O}$ is itself a set in $\mathcal{O}$.

A3. The intersection of any two sets in $\mathcal{O}$ is in $\mathcal{O}$.

The set $\mathcal{O}$ is called the topology on V. Any set V admits the trivial topology $\mathcal{O} = \{\emptyset, V\}$ and the discrete topology consisting of all subsets of V. A more concrete example of a topology on the finite set $V = \{a,b\}$ is given by

---

S.S.Anderson and G.Chartrand, The lattice-graph of the topology of a transitive directed graph, to appear in Math. Scand.

$\mathcal{O} = \{\emptyset, V, \{a\}\}$. We remark that when V is finite, axiom A2 may be replaced by

A2'. The union of two sets in $\mathcal{O}$ is in $\mathcal{O}$.

To describe a nontrivial topology on an infinite set, it is convenient to first introduce the concept of "basis". A collection of subsets $\mathcal{B} = \{B_\alpha\}$ of a given set V is a <u>basis for a topology</u> on V if

B1. $\bigcup B_\alpha = V$.

B2. For every point p in $B_\alpha \cap B_\beta$ there is a B_γ in such that p is in B_γ and $B_\gamma \subset B_\alpha \cap B_\beta$.

It follows from this definition that the collection of open sets in any topology is a basis. We may also induce a topology from a basis $\mathcal{B}$ by defining a subset of V to be open if it is a union of elements of $\mathcal{B}$. The collection of interiors of all circles induces a topology in the Euclidean plane, and the open intervals $(a,b) = \{x \mid a<x<b,$ a,b, in R$\}$ form a basis for a topology on the set of real numbers. The latter topological space provides a ready illustration of the fact that the intersection of an infinite number of open sets need not be open; namely,

$$\bigcap_{n=1}^{\infty} \left(\frac{-1}{n}, \frac{1}{n} \right) = \{0\}.$$

The two topologies $\mathcal{O}_a = \{\emptyset, V_1, \{a\}\}$ and $\mathcal{O}_b = \{\emptyset. V_1, \{b\}\}$ on the set $V_1 = \{a,b\}$ are clearly identical in structure; the only difference lies in the labeling of the points of V_1. Isomorphism of topological spaces, called <u>homeomorphism</u>, is defined as follows: Two spaces $S_1 = (V_1, \mathcal{O}_1)$

and $S_2 = (V_2, \mathcal{O}_2)$ are called homeomorphic if there is a 1-1 correspondence between their point sets which preserves open sets; that is, a correspondence such that an open set in T_1 corresponds to an open set in T_2, and conversely. Thus the above topological spaces $(V_1, \mathcal{O}_a)$ and $(V_1, \mathcal{O}_b)$ are homeomorphic, although their labeled topologies are different.

We now establish a correspondence between labeled topologies and labeled transitive digraphs.[16] It is clear that this can also be done for the unlabeled case simply by relating the homeomorphism classes of topologies to the isomorphism classes of digraphs.

Theorem 9. There is a 1-1 correspondence between the labeled topologies on n points and the labeled transitive digraphs.

To describe this correspondence we first associate with each topology $\mathcal{O}$ a digraph $D(\mathcal{O})$ on the same set of points with the point u adjacent to the point v ($v \neq u$) if u is in every open set containing v. Clearly $D(\mathcal{O})$ is transitive and uniquely determined by $\mathcal{O}$. According to this rule, the topology $\mathcal{O}_a$ described above corresponds to the digraph D_1 in Figure 10, and the topology $\mathcal{O}_b$ to D_2; the topology on $V_2 = \{a,b,c,d\}$ defined by $\mathcal{O}_2 = \{\emptyset, V_2, \{a\}, \{a,b\}, \{a,b,c\}\}$ corresponds to the digraph D_3.

---

[16] J.W.Evans, F.Harary, and M.S.Lynn, On the computer enumeration of finite topologies, Comm. A.C.M. 10(1967), 295-297.

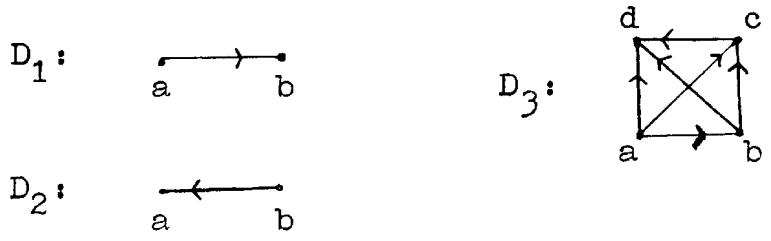


Figure 10. Digraphs of labeled topologies

Conversely, associate with each labeled transitive digraph D with n points the topology $\mathcal{O} = \mathcal{O}(D)$ induced by the basis $\mathcal{B}$ consisting of the sets $Q(u) = \{v \mid u \text{ is reachable from } v\}$ for all u in V. Since D is transitive, Q(u) is the set containing u and all points adjacent to u. The digraph in Figure 11 gives rise to the basis

$$\mathcal{B} = \{Q(a)=\{a\}, Q(b)=\{b\}, Q(c)=\{c,b\}, Q(d)=V, Q(e)=\{a,e\}\}$$

which induces the topology

$$\mathcal{O}(D)=\{\emptyset, V, \{a\}, \{b\}, \{a,b\}, \{b,c\}, \{a,b,c\}, \{a,e\}, \{a,b,e\}, \{a,b,c,e\}\}.$$

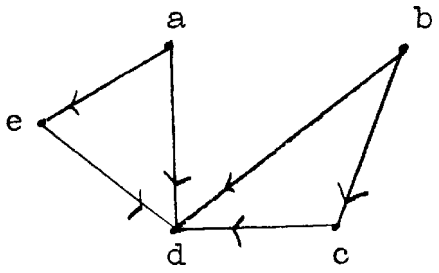


Figure 11. Transitive digraph

To show, in general, that $\mathcal{O}(D)$ is a topology, it is necessary only to show that $\mathcal{B}$ is a basis. Clearly $\bigcup B_i = V$; the set $B_i \cap B_j$ is the set of points adjacent to v_i and to v_j, and by transitivity, $Q(v_k) \subset B_i \cap B_j$ for each v_k in $B_i \cap B_j$. Therefore $\mathcal{B}$ is a basis.

We see that $D(\mathcal{O}(D)) = D$ is an immediate consequence of the defining constructions. It is clear that there is also a 1-1 correspondence between finite topologies and reflexive transitive relations (quasi-orders) by merely adding a loop at every point. This correspondence has been used[17] to enumerate the number of labeled topologies on n points, for $n \leq 7$. Table 2 shown below lists the results of this enumeration.

n	$\mathcal{O}_n$
1	1
2	4
3	29
4	355
5	6,942
6	209,527
7	9,535,241

Table 2. Number of labeled topologies on n points

Several types of topological spaces are classified according to the "separation" axioms given below; a space which satisfies axiom T_i is called a T_i-space.

T_0. Given any two points of a topological space S, at least one lies in an open set not containing the other.

T_1. Given any two points of S, each lies in an open set not containing the other.

T_2. Given any two points of S, there are disjoint open sets, each containing just one of the two points.

---

[17] Evans, Harary, and Lynn, 295-297.

We have already encountered examples of these spaces; the finite topological space S consisting of the set {a,b} and the topology {∅, {a,b},{a}} is T_0 (but not T_1) since the point a is in an open set not containing b. A space which is T_1 but not T_2 is given by the real line together with the topology of sets having finite complements; to see this, note that, for any two points, the complement of one point is an open set containing the other, but disjoint open sets containing the points cannot be obtained. The ordinary topology on the real line (whose basis is the set of open intervals) defines a T_2-space, usually called a Hausdorff space.

For a finite set S to be T_1, every subset of S must occur in the topology. This is easily proved, since any given point p lies in an open set not containing p', call it $O_{p'}$, for each p' in S. Then $\bigcap O_{p'} = \{p\}$ is open since the intersection is finite. Thus the sets containing a single point of S form a basis for the topology. This also implies that for finite sets the T_1 axiom is equivalent to the T_2axiom. Moreover, since every point constitutes an open set, we have the following theorem.

Theorem 10. A finite topological space S corresponds to a totally disconnected digraph if and only if S is a T_1-space.

In a T_0-space, given any two points, one is not in an open set containing the other; this fact implies the next theorem.

Theorem 11. A finite topological space S corresponds to an asymmetric digraph if and only if S is a T_0-space.

With each topology θ on a finite set A, we may also associate a lattice[18] by representing the open sets O_i by points v_i and drawing an arc from v_i to v_j if $O_i \subset O_j$; the lattice L is called the topological lattice of θ. The axioms A1, A2, and A3 insure that this structure is indeed a lattice, in fact, a sublattice of the lattice L_A of all subsets of the set A. Conversely, any sublattice of L_A defines a topology on A.

To describe the relationship between transitive digraphs and lattices of topologies, a few intermediate concepts are convenient. A clique in a digraph is a maximal complete symmetric subgraph. In Figure 12, the digraph D_1, which has two cliques having two and three points respectively, illustrates the fact that, in general, the cliques of a digraph do not partition the points. However, a partitioning does occur when the digraph is transitive, as illustrated by the digraph D_2 of Figure 12.

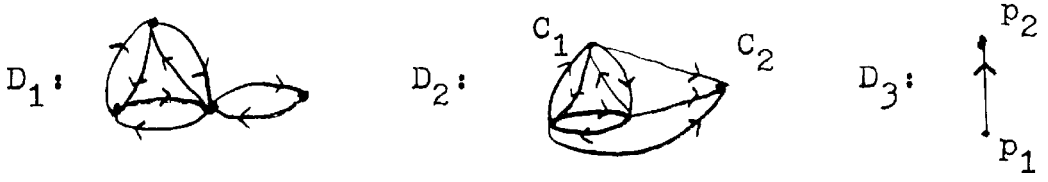


Figure 12. Cliques in digraphs

Theorem 12. The cliques of a transitive digraph partition the point set.

---

[18]See the paper by Anderson and Chartrand for more detail.

Therefore, for a transitive digraph we may define the clique-condensation digraph whose points are the cliques of D and in which there is an arc from p_1 to p_2 if there is an arc from clique C_1 to clique C_2 in the original digraph D. The digraph D_3 in Figure 12 is the clique-condensation digraph of the digraph D_2. By transitivity, if there is an arc from a point of C_1 to a point of C_2, then there must be an arc from every point of C_1 to every point of C_2. Furthermore, there can be no arcs from C_2 to C_1 since the cliques are maximal.

Theorem 13. The clique-condensation digraph of a transitive digraph is transitive and asymmetric.

Corollary 13a. A transitive digraph D is isomorphic to its clique-condensation digraph if and only if D is asymmetric.

Since there is a topology corresponding to each transitive digraph D, we may speak of the topological lattice of D. It is not difficult to find two different transitive digraphs having the same topological lattice. For example, D_1 and D_2 in Figure 13 both have the topological lattice L. Such digraphs, however, do share a common property.

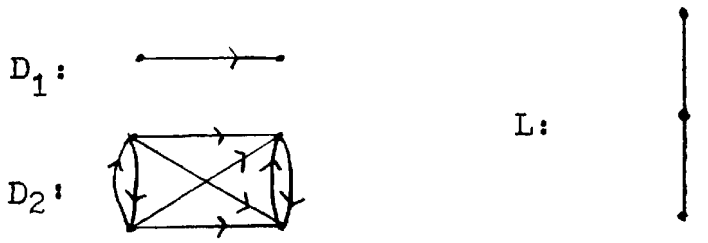


Figure 13. Two digraphs with the same topological lattice

Theorem 14. Two transitive digraphs have the same topological lattice if and only if they have isomorphic clique-condensation digraphs.

Corollary 14a. Given a lattice L, there exists a unique asymmetric transitive digraph D, whose topological lattice is isomorphic to L.

We remark also that the digraph of Corollary 14a will be a tournament if and only if L is a single path. To illustrate the corollary, consider the lattice L in Figure 14. The lowest point represents the empty set, and the highest point the entire set. The points between represent a succession of sets each included in the next. The unique asymmetric transitive digraph associated with L is also shown.

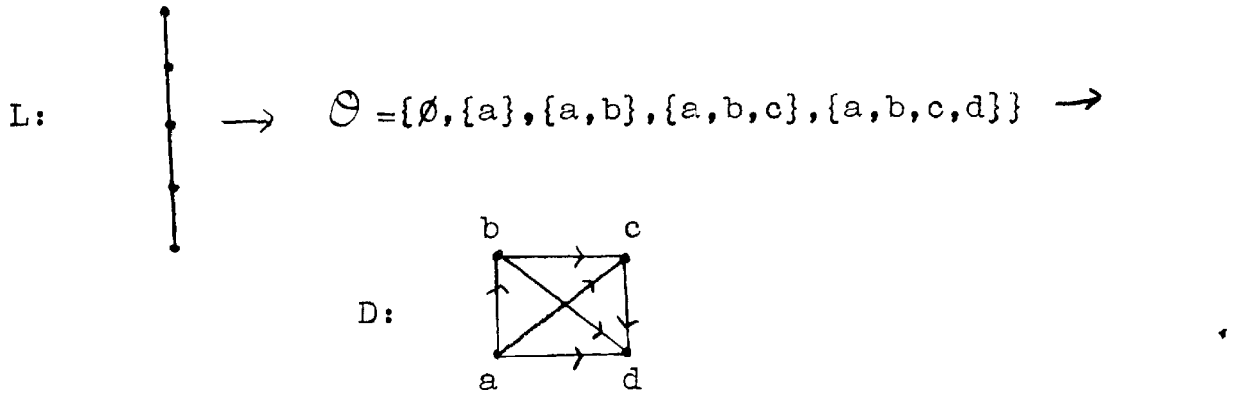


Figure 14. Tournament associated with lattice L

It follows that every transitive digraph whose topological lattice is L can be constructed by a suitable replacement of the points of D by complete symmetric digraphs.

5. Mathematical puzzles

In discussing mathematical applications of graph theory, it would be unthinkable to omit all mention of the mathematical puzzles to which graph theory owes its first discovery and, to a certain extent, its continuing popularity. As we have mentioned above, the first paper on graph theory was published by Euler[19] in 1736 when he solved the Königsberg bridge problem. This problem involves the city of Königsberg with its seven bridges connecting the islands in the Pregel River with the mainland and with each other, as shown in Figure 15. The challenge is to take a walk which crosses each bridge exactly once and ends at the starting point.

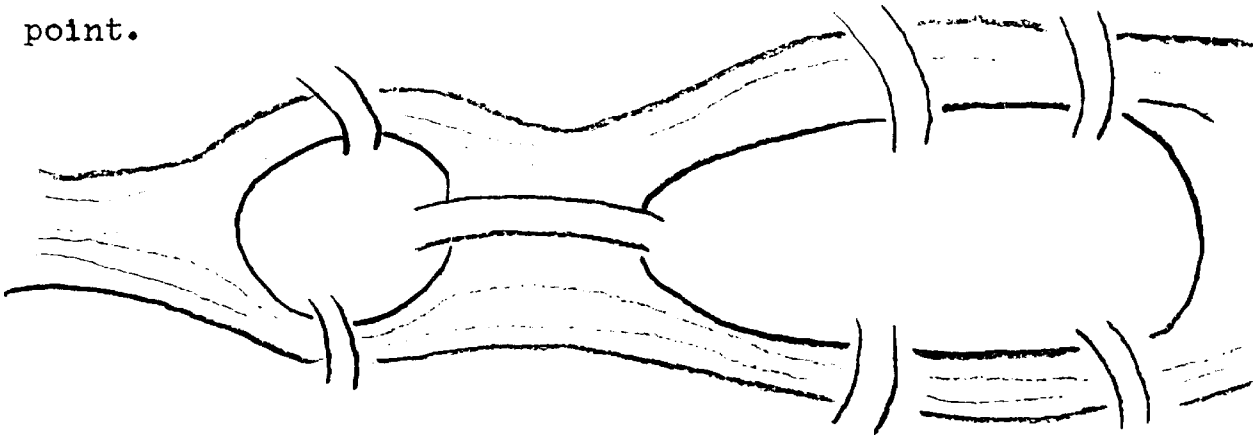


Figure 15. Map of Königsberg

We can abstract from Figure 15 the diagram (multigraph) of Figure 16, in which the points represent the land areas and the lines represent the bridges. In his paper, Euler proved, along with some generalizations of the problem,

---

[19] L. Euler, Solutio problematis ad geometriam situs pertinentis, Commentarii Academiae Petropolitanae 8(1736), 128-140. It is translated in Sci.Amer. 189(July, 1953), 66-70.

that a walk of the prescribed type is not possible, his negative solution being based on the observation that each land area must be left as often as it is entered. Since there is an odd number of bridges at some (in fact all) of these areas, such a walk is impossible.

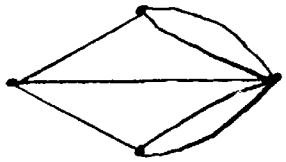


Figure 16. Graph of Königsberg

In a second puzzle, invented by Sir William Hamilton, the player is given a regular dodecahedron, each of whose vertices is labeled with the name of an important city. The object of the game is to make a trip "around the world" on the edges of the dodecahedron passing through each city exactly once and returning home; thus the player is asked to show that the graph of the dodecahedron is hamiltonian by finding a spanning cycle. The dodecahedron and its graph are shown in Figure 17; we leave it to the reader to find a spanning cycle.

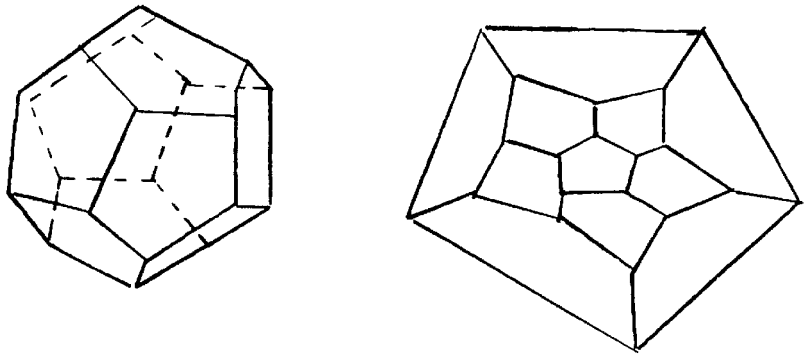


Figure 17. Dodecahedron and its graph

The graph $K_{3,3}$ is a representation of an ancient puzzle involving three feuding Arabs. There are three public wells in the region where the three Arabs live, and since any one of them may run dry during certain periods of the year, each Arab must have access to each of the wells. The player is asked to construct, from each Arab's tent to each well, paths which, in the interests of peaceful coexistence, are nonintersecting. Of course, this is impossible, since $K_{3,3}$ is nonplanar.

A fourth puzzle, called the Tower of Hanoi, still found among children's toys, consists of a board with three pegs, and, spindled on one of the three pegs, a set of n disks in increasing size, as shown in Figure 18.

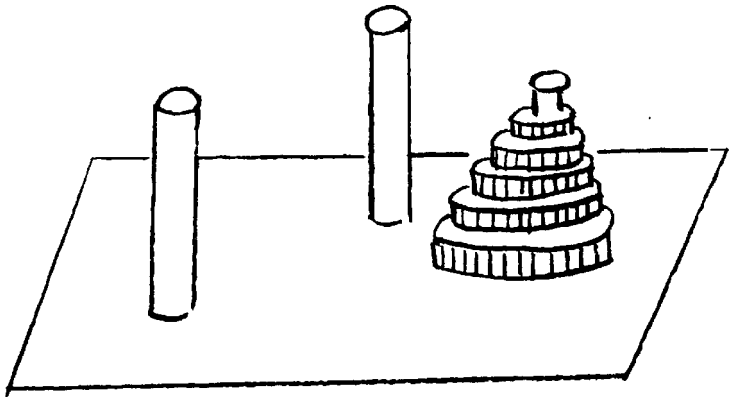


Figure 18. Tower of Hanoi puzzle

The object of the puzzle is to move the disks one at a time, never permitting any disk to be on top of a smaller disk, from the original peg to the other pegs, until the tower is rebuilt on another peg. The method usually obtained by the successful child (by trial-and-error) is to move the smallest disk on alternate moves one step clockwise

and in the other steps, to make the one possible remaining move, thus completing the puzzle in $2^n - 1$ moves. It is surprising that this puzzle has a graphical interpretation, but we now show that the solution gives a spanning cycle of the graph of the n-cube Q_n. We label the 2^n points of the n-cube by n-tuples $(e_1,\ldots,e_n)$ where $e_i = 0$ or 1, and draw a line between two points if their labels differ in exactly one coordinate. The graph of Q_3 is shown in Figure 19. A spanning cycle of Q_n is then a sequence of length $2^n - 1$ of the n-tuples in which two consecutive n-tuples (and the first and the last) differ in exactly one coordinate. Now labeling the disks of the puzzle from 1 to n in order of increasing size, we associate with each position in the game an n-tuple in which the i-th coordinate gives the number of moves completed by the i-th disk (mod 2). It is easy to see that the positions in the solution of the puzzle correspond to a walk on Q_n which starts at $(0,0,\ldots,0)$, and a simple induction argument[20] proves that each n-tuple occurs exactly once in this sequence, the final position being $(0,0,\ldots,0,1)$. Therefore, the solution of the Tower of Hanoi puzzle corresponds to a spanning cycle of Q_n. For Q_3, this spanning cycle is (0,0,0), (1,0,0), (1,1,0), (0,1,0), (0,1,1), (1,1,1), (1,0,1), (0,0,1).

---

[20] D.W.Crowe, The n-dimensional cube and the tower of Hanoi, Amer. Math. Monthly 63(1956), 29-30.

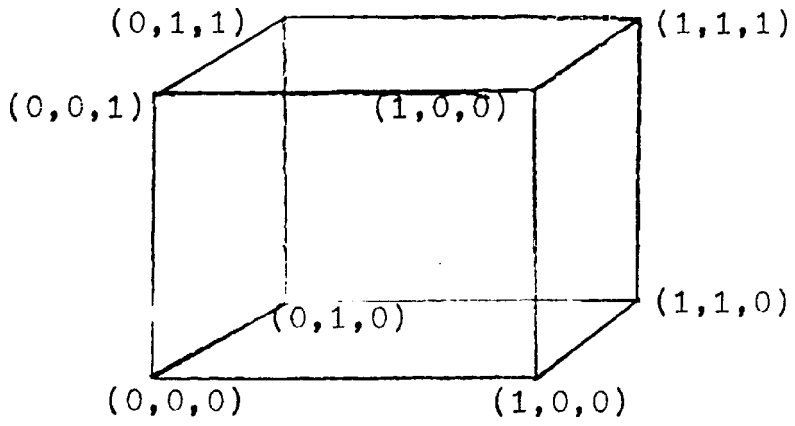


Figure 19. The 3-cube Q_3

VIII. NONMATHEMATICAL APPLICATIONS OF GRAPH THEORY

Graph theory, by providing an abstract model in which the structure of various situations can be clearly viewed and analyzed, plays a particularly important role in nonmathematical areas, and, in return, nonmathematical applications have provided an added stimulus to graphical research. It is therefore appropriate to include a chapter on the areas outside mathematics with which graph theory is closely associated. There are, however, hundreds of topics in this category--in engineering, in the physical, biological, and social sciences, and in numerous other areas--and a comprehensive up-to-date survey of them is impossible, since this is one of the most rapidly expanding areas of graph theory. Therefore, we have chosen, somewhat arbitrarily, a few of the many possible topics in order to demonstrate the possible uses of graph theory. Some of these topics represent independent discoveries of graph theory and others illustrate the conscious application of graph theory; some use only the diagrammatic representation provided by graphs and others use the full power of graph-theoretic results; all show the versatility of graph theory in any situation which can be viewed as a set of objects related to each other in pairs. In this chapter we

first present two applications of historical interest, one in chemistry and one in circuit theory; secondly, we discuss in detail the important interaction between graph theory and social psychology; and finally, we mention briefly applications to sociology, critical path analysis, economics, information theory, and bell-ringing.

1. Circuit theory

Kirchhoff,[1] a physicist interested in current and voltage flows in electrical network theory, invented graphs in 1847 to enable him to study the abstract structure of networks without regard to physical factors such as the lengths of wires or the effects of resistances and inductances, etc. He constructed network diagrams of two-terminal elements joined to each other at their terminals in which the sole important feature was the interconnections; that is, he associated a graph with each network, as shown in Figure 1.

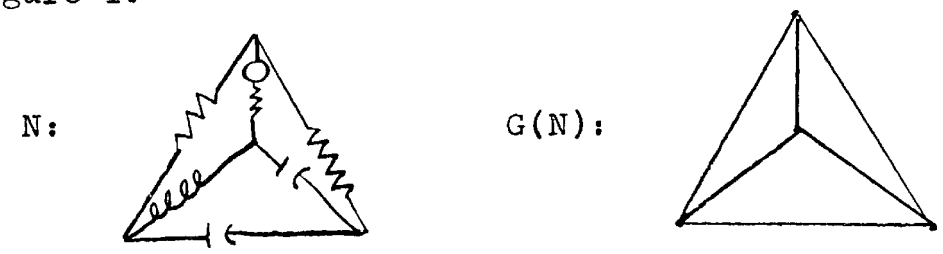


Figure 1. Graph of a network

To study the current and voltage flows in these networks, Kirchhoff assigned to each line a current and

---

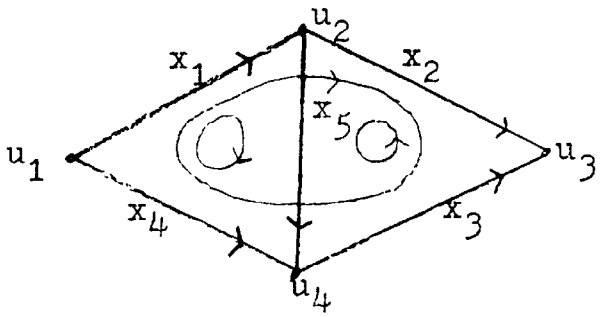
[1] G.Kirchhoff, Uber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung Galvanischer Ströme geführt wird, Annalen der Physik 72 (1847), 497-508. Translated by J.B.O'Toole in IEEE Trans. Circuit Theory 5(1958), 4-7.

voltage, together with a flow direction, with the currents flowing toward any point considered positive, and those flowing from it negative. His well-known laws state that

1) The sum of the currents at any point is zero.

2) The voltage drop around any circuit is zero.

These laws may be summarized in matrix form by letting F be the point-arc incidence matrix of the oriented graph in which f_{ij} is 1 if u_i is the terminal point of x_j, -1 if u_i is the initial point of x_j, and 0 otherwise. The cycle-arc incidence matrix C is obtained by assigning an orientation to each cycle and defining c_{ij} to be 1 if the arc x_j occurs in the cycle C_i with its direction matching that of the cycle, -1 if the direction is opposite to that of the cycle, and 0 if x_j is not in C_i. We show in Figure 2 the oriented graph of an electrical network together with its matrices F and C.



$$F = \begin{pmatrix} -1 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & 0 \end{pmatrix}$$

Figure 2. Oriented graph of an electrical network with its matrices F and C

Defining E and V to be the current and voltage vectors, respectively, of the arcs, Kirchhoff's laws may be expressed

matricially as

1) $F \cdot E = 0$.

2) $C \cdot V = 0$.

In this form, Kirchhoff's laws give rise to two sets of linear equations. For the example in Figure 2, we obtain

$$\begin{aligned} -e_1 - e_2 &= 0 \\ e_1 - e_2 - e_5 &= 0 \\ e_2 + e_3 &= 0 \\ -e_3 + e_4 + e_5 &= 0 \end{aligned}$$

and

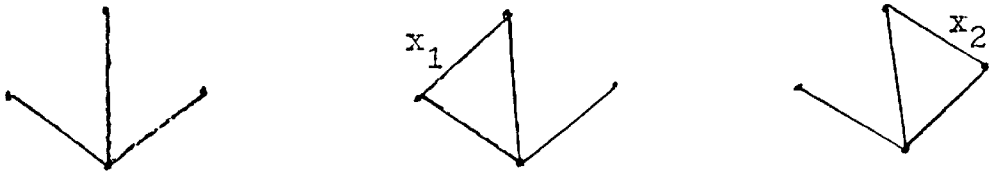
$$\begin{aligned} v_1 - v_4 + v_5 &= 0 \\ -v_2 + v_3 + v_5 &= 0 \\ v_1 + v_2 - v_3 - v_4 &= 0 \quad . \end{aligned}$$

Using graphical properties, Kirchhoff was able to show that for a connected network, exactly p-1 of the current equations and q-p+1 of the voltage equations are linearly independent; that is, he showed that the matrix F has rank p-1 and the matrix C has rank q-p+1. His method, in brief, was as follows. The determinant of any submatrix of F corresponding to a cycle in the graph of the network is zero, since each column has exactly one +1 and one -1; furthermore, the determinant of any submatrix corresponding to a subgraph of G which contains a cycle is zero. Therefore, the rank of F is at most the number of points in a spanning tree of D. He also showed that the determinant of a submatrix corresponding to a tree is not zero,[2] and

---

[2]The matrix-tree theorem is thus implicit in Kirchhoff's work.

thus the rank of F is p-1. To find the rank of C, we consider a maximal set of independent cycles of G, formed by choosing a spanning tree, and adjoining one additional line of the graph at a time. In Chapter I we showed that there are q-p+1 cycles in this set; label them 1,2,...,q-p+1. Labeling the line of G which created the cycle i by x_i, and the remaining lines in G arbitrarily, the q-p+1 square submatrix in the upper left hand corner of C has 1's and -1's on its diagonal and 0's elsewhere. Thus the rank of C is at least q-p+1. (See Figure 3.) It can also be established that this is actually an equality.[3]



spanning tree cycle 1, line x_1 cycle 2, line x_2

Figure 3. A maximal set of independent cycles for the graph in Figure 2

It is not difficult to show that the matrix F in Figure 2 has rank 3 and the matrix C has rank 2. Assignments of voltages and currents to the arcs of the network, resulting from a solution to the voltage and current equations, are shown in Figure 4. It can easily be verified that these assignments satisfy Kirchhoff's laws.[4]

---

[3]Further information on Kirchhoff and on graphical applications to electrical network theory are to be found in S.Seshu and M.B.Reed, Linear Graphs and Electrical Networks, Addison-Wesley, Reading, Mass., 1961. A detailed discussion of this method is to be found on pp. 62-71.

[4]R.L.Brooks, C.A.B.Smith, A.H.Stone, and W.T.Tutte have also applied Kirchhoff's laws to a graphical problem in The dissection of a rectangle into squares, Duke Math.J.,7(1940, 312-340.



Figure 4. Assignments of currents and voltages

2. Chemistry

The mathematician Cayley[5] is also credited with a discovery of graph theory, in connection with a problem in chemistry involving hydrocarbons. In a saturated hydrocarbon molecule, each carbon atom has valence 4 and each hydrogen atom valence 1; thus a saturated hydrocarbon can be represented by the formula C_nH_{2n+2}. Several examples are shown in Figure 5.

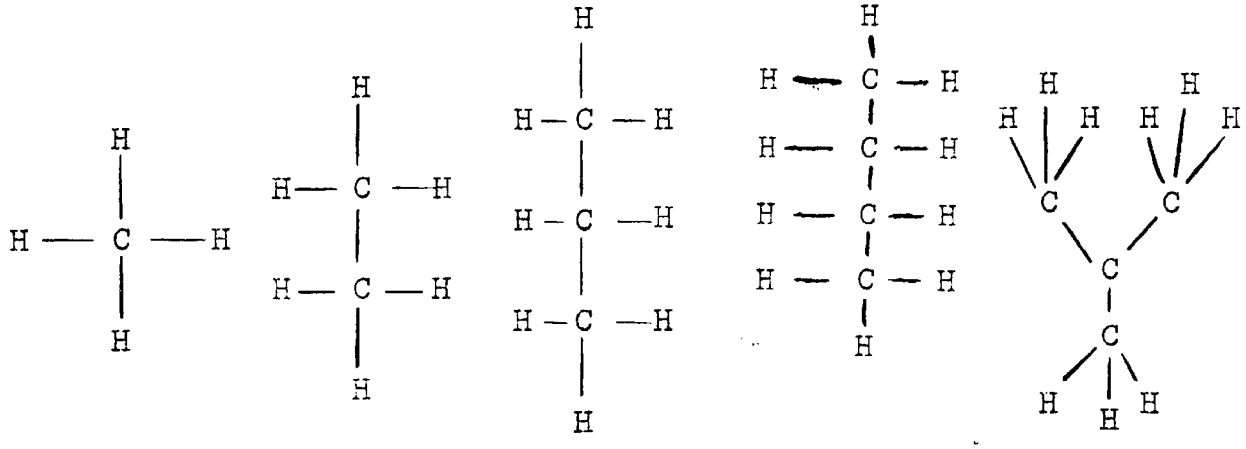


Figure 5. Hydrocarbon molecules

---

[5]A. Cayley, On the mathematical theory of isomers, The Collected Mathematical Papers of Arthur Cayley, Cambridge University Press, Cambridge, 1889-1897, XIII, 26-28.

Cayley's interest in hydrocarbons centered around the question of how many different hydrocarbons with a given number of carbon atoms were possible. He developed the idea of a tree and proved the general formula[6] for the number of labeled trees with n points. For small values of n he also calculated[7] the number of hydrocarbons, which can be represented by trees with every point having degree 4 or 1.

3. Social Psychology

In social psychology, graph theory has had a perhaps unique opportunity to develop in direct conjunction with an application. The social psychologist D. Cartwright and the graph theorist F. Harary have been studying topics of value to both fields such as, for example, "balance" in digraphs and in human relationships.[8]

The interest of social psychologists in graph theory began in the 1930's when Lewin,[9] in an attempt to explain human behavior, viewed the "life space" of an individual or group as a set of "cells" representing various factors in an individual's environment. A relationship or communication between two cells was indicated by "touching", and

---

[6]Cayley, A theorem on trees, XIII, 26-28.

[7]Cayley, On the analytical forms called trees, with applications to the theory of chemical combinations, IX, 427-460.

[8]D.Cartwright and F.Harary, Structural balance: a generalization of Heider's theory, Psychol. Rev.63(1956), 277-293.

[9]K.Lewin, Principles of Topological Psychology, trans. by F.Heider and G.M.Heider, McGraw Hill, New York, 1936.

Lewin attempted to study the "dynamic" interplay of the cells in an individual's life space by introducing the new field of "hodology", the study of paths. There was a great need for the new type of geometry, since Euclidean geometry was too rigid a discipline in which to view psychological concepts and topology was too general, since it provides no quantitative description of distances and direction necessary in the study of psychological forces.

Bavelas[10] gave Lewin's geometry more precise mathematical formulation, making certain changes in Lewin's basic concepts in order to provide a mathematical model which could be easily and widely used in psychological situations. He reduced the cell to a point and indicated "touching" by a line between points--thus he too, invented the graph. In doing so, he eliminated some of the problems inherent in Lewin's representation, such as that of depicting several mutually adjacent cells, a problem we have encountered before in connection with the Four Color Conjecture. The problem of how to define distance was solved by the graphical definition of distance as we know it, and direction was indicated by the use of arrows in place of lines. Bavelas studied distance and centrality in detail, applying them to empirical situations.

Lewin's and Bavelas's efforts were directed toward providing a foundation and a framework for psychological

---

[10]A. Bavelas, A mathematical model for group structures, Applied Anthropology 7(1948), 16-30.

experimentation. They were successful; we cite one study directly inspired by Bavelas, as an example. Leavitt[11] tested the effect of various communication patterns on the performance of groups by seating five people at a circular table and inducing various rules of communication on them, as shown in Figure 6. He then gave each member of the group one card from a set of six cards numbered from 1 to 6. Each person was asked to determine, through the established channels of communication, which card was left. The performance of each group showed a definite relationship to the communication pattern under which it operated. The person seated at the central point of the graph tended to become the leader of the group; in the cycle pattern, having no central point, a leader was less likely to develop. Furthermore, as a person's distance from the central point increased, his satisfaction with his role diminished greatly, the enjoyment of job being approximately equal among the members of the cycle pattern.

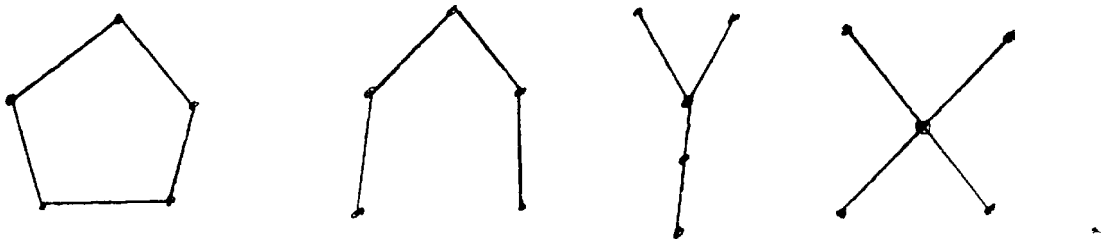


Figure 6. Communication channels

[11] H.J.Leavitt, Some effects of certain communication patterns on group performance, J.Abn.Soc.Psychol. 46(1951), 38-50.

Digraph theory has been developed greatly under Cartwright and Harary.[12] One particular topic which has interested them is the study of balance, or equilibrium, in human relations. Since human relations tend to break into three categories--positive, negative, or indifferent--a signed digraph in which some arcs are viewed as positive and some as negative is used. For example, a positive line might represent love, a negative line hate, and the absence of a line indifference. We use a dashed arc to represent negative lines, as shown in the examples in Figure 7.

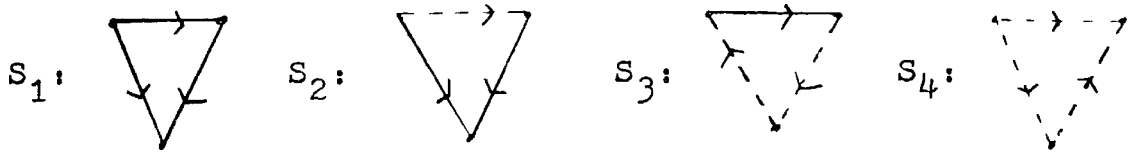


Figure 7. Signed digraphs

A signed digraph is balanced if every semicycle has an even number of negative lines. Thus S_1 and S_3 in Figure 7 are balanced. The sign of a semiwalk in a digraph is defined to be the product of the signs of its arcs. Using this definition, we give a graphical description of balance[13] in Theorem 1.

Theorem 1. The following statements are equivalent for a signed digraph S:

1) S is balanced, that is, every semicycle of S is positive.

---

[12]They have written a book on it: F.Harary, R.Norman, and D.Cartwright, Structural Models, Wiley, New York, 1965.

[13]Harary, Norman, and Cartwright, p.342.

2) For every pair of points, all semipaths joining them have the same sign.

3) The set V of points of S can be partitioned into two subsets (one of which may be empty) such that every positive arc of S joins two points in the same subset and every negative arc joins two points in different subsets.

Signed digraphs also have been used in empirical research on stability and balance within a group of people.[14]

4. Miscellaneous topics

Having discussed briefly three areas in which graph-theoretic applications have been important, we now mention several areas in which graphs have been used to a lesser, though not insignificant, extent. The first four topics illustrate how problems in such diverse areas as telephone networks, cryptography, bell-ringing, and printed circuits can be stated as questions about spanning trees, spanning cycles, eulerian trails, and planarity.

Related to a problem of the Bell Telephone System concerning the rate structure for leased line services, Prim[15] published a paper discussing algorithms for finding the shortest network connecting a given set of cities. The resulting network is, naturally, a spanning tree, and it

---

[14] See, for example, J.O.Morrissette, An experimental study of the theory of structural balance, Hum. Rel. 11 (1958), 329-354.

[15] R.C.Prim, Shortest connection networks and some generalizations, Bell System Tech. J. 36(1957), 1389-1401.

can be constructed from a weighted graph showing all possible connections between the cities by first choosing the shortest link of the network and then proceeding by selecting successively the next shortest links which do not form cycles with preceding links.

Spanning cycles have a curious application to the ancient English art of change-ringing. Motivated by mechanical limitations, the English developed a method of ringing church bells in mathematical patterns, rather than in melodies. A tower may have from five to twelve bells, numbered in order of increasing size. A "change" consists of ringing the bells one at a time in any order, say 14532768, and specific rules govern which changes may succeed a given change. These rules are symmetric in nature, since any change which may succeed a given change may also precede it, and thus a graph is obtained. A "composition" outlines a pattern by which each change is rung exactly once, returning in the end to the initial change, which is always the progression from the highest bell to the lowest bell, 12345678. Thus each composition describes a spanning cycle of the graph of changes. An unsolved problem of change-ringing as well as of graph theory is the question of how many compositions are possible on a given number of bells.

There is a particularly interesting family of rela- graphs whose eulerian trails have application to cryptog- raphy and information theory. These are the "teleprinter

diagrams" which have as their point sets all the binary sequences of length n. The teleprinter diagram T_0 is defined to be a single point with two loops, labeled 0 and 1. The points of T_1 are the arcs of T_0. At the point 0 is a loop labeled 00, at the point 1 is a loop labeled 11, and between the points are the two arcs (0,1), labeled 01, and (1,0), labeled 10. The teleprinter diagram T_n is the relagraph whose points are the arcs of T_{n-1} and which has the arc (x,y) if the last n-1 digits of x coincide with the first n-1 digits of y. The arc (x,y) is labeled with the n digits of x followed by the last digit of y. The first four teleprinter diagrams are illustrated in Figure 8. It is easy to see that each teleprinter diagram is a digraph plus loops, 00...0 and 11...1, so the calculation of the number of eulerian trails is essentially the same as for a digraph.

Given a message in the form of a sequence of letters from some finite alphabet, say {0,1}, written $S = a_1a_2\ldots$, one can describe the walk in T_1 determined by S as point a_1 followed by a_2, etc. In T_2 the point sequence of the walk is $a_1a_2, a_2a_3, a_3a_4, \ldots$, and in T_3 it is $a_1a_2a_3, a_2a_3a_4, \ldots$ For example, the sequence 1000101110 in T_2 is the walk whose point sequence is 10,00,00,01,10,01,11,11,10. The determination of eulerian trails in teleprinter diagrams is a way of compiling useful information toward the solution of the code in which the massage is given. A formula for the number of eulerian trails in a teleprinter diagram has been

given by de Bruijn.[16]

Theorem 2. The number of eulerian trails in the teleprinter diagram T_n is $2^{2^{n-1}-1}$.

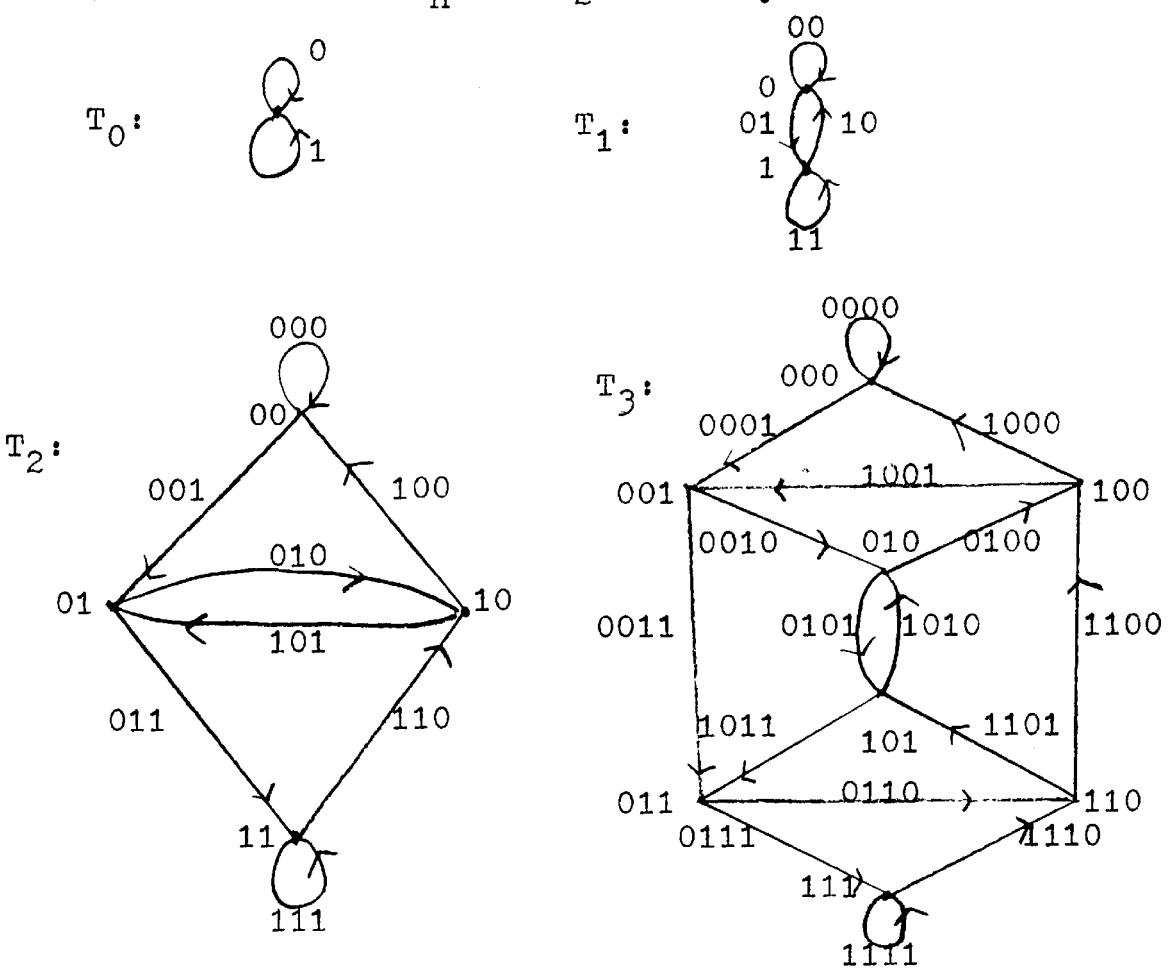


Figure 8. Teleprinter diagrams

The theory of planar graphs has a concrete application to printed circuits. Since the wires printed on a plastic (nonconducting) plate are not insulated, no two of them must cross except at a definite junction. This is usually accomplished by printing the wires on both sides of the plate and passing connecting wires through holes in the

---

[16] N.G.deBruijn, A combinatorial problem, Indag. Math. 8(1946), 461-467.

plate at the desired points. A circuit can be printed on a single side if and only if it corresponds to a planar graph.[17] The problem raised by printed circuits has led to a new graphical concept, the thickness of a graph G, which is the minimum number of planar graphs whose union is G; it has been studied by Beineke and Harary.[18]

In some cases, the diagrammatic representation provided by graphs serves as the most important contribution of graph theory to a given field. In sociology, the analysis of kinship systems uses a graphical representation, in critical path analysis a digraph depicts the precedence relation among jobs in a project, and in economics graphs show the flows of goods and services. In each of these areas, graphs are a gross oversimplification of the empirical situation, but the graphical representation serves a useful purpose in diagramming the fundamental structure of the system.

In discussing kinship systems, White[19] proposes the following axioms.

1) The entire population is partitioned into n clans.

2) There is a permanent rule fixing the single clan among whose women the men of a given clan must find their wives.

---

[17]Further information along these lines may be found in R.Busacker and T.Saaty, Finite Graphs and Networks, McGraw Hill, New York, 1965, 195ff.

[18]L. Beineke and F.Harary, On the thickness of the complete graph, Bull. Amer. Math. Soc. 70(1964), 618-620.

[19]H.C.White, An Anatomy of Kinship, Prentice-Hall, Englewood Cliffs, N.J., 1963, p.34.

3) Men from 2 different clans cannot marry women of the same clan.

4) All children of a couple are assigned to a single clan, uniquely determined by the clans of their mother and father.

5) Children whose fathers are in different clans must themselves be in different clans.

6) A man can never marry women of his own clan.

7) Every person in the society has some relative by marriage and descent in each other clan, i.e., the society is not split into unrelated groups.

8) Whether two people who are related by marriage and descent links are in the same clan depends only on the kind of relationship, not on the clan to which either one belongs.

Marriage and descent rules are conveniently represented by (0,1) permutation matrices W and D, respectively, restricted only by axioms 6,7,and 8. Axiom 6 implies that W has zeros on its diagonal, and Axiom 8 implies that if $d_{ii} = 1$ for any i, then $d_{ii} = 1$ for all i, and furthermore, $D_m = I$ for some $m \leq n$. The product WD is meaningful, specifying for a man, the clan to which the children of his wife's brother belong. The product DW, on the other hand, gives the clan into which a man's son must marry. In general, the products $W^a D^b$ give the clans of various relatives, D^{-1} gives the clan of a child's father, and W^{-1} the clan of a woman's husband.

In discussing marriage taboos, a primary consideration

is usually whether first cousins can marry. A boy and a girl are called bilateral cousins if the father of each is is the brother of the other one's mother. They are matrilateral cousins if the mother of the boy is the sister of the girl's father, and patrilateral cousins if the boy's father is the brother of the girl's mother. In terms of the matrices W and D, bilateral cousins are allowed to marry if $W^2 = I$ and $DW = WD$. Matrilateral cousins are allowed to marry if $W^2 \neq I$, but $WD = DW$, and patrilateral cousins if $W^2 \neq I$, but $WD = DW^{-1}$.

The kinship system can easily be represented by a digraph in which the points represent the clans. The arcs are colored red or blue according to which rule, marriage or descent, is being applied. Axioms 2 and 3 imply that there is exactly one red arc entering every point and at most one leaving it; axioms 4 and 5 imply that there is exactly one blue arc leaving every point and at most one blue arc entering it; and axiom 6 stipulates that there are no loops. Some examples are shown in Figure 9.

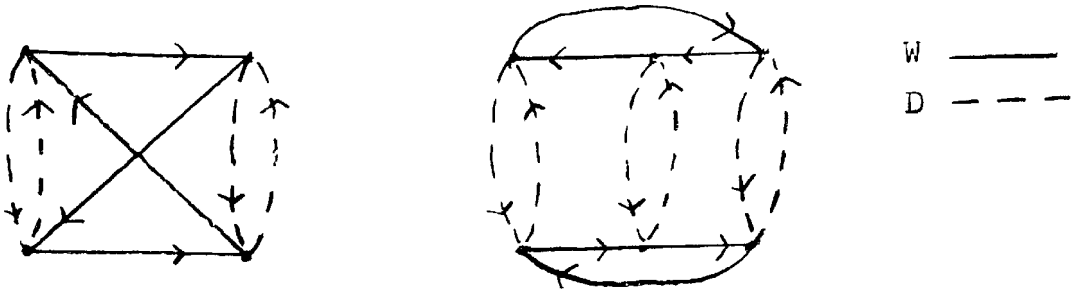


Figure 9. Kinship systems

The graphical implications of axioms 7 and 8 are that the digraph must be strongly connected, and any sequence of red and blue arcs which forms a cycle at one point must form a cycle at every point.

White studied the graphical structure of various marriage systems, introducing such concepts as the "dual" system, in which the marriage and descent lines are interchanged. He then examined several tribes--Kariera, Arunta, Murngin, Purum--placing them in various kinship categories.

In critical path analysis,[20] a primary concern is the scheduling of tasks in a project in some efficient manner so that the project may be completed as smoothly and quickly as possible. In analyzing a project, the planner compiles a list of all the jobs involved and for each job, the time necessary to complete it and the jobs which must precede it. It is helpful to arrange this information on a digraph, or more usually, a Hasse diagram (since a cycle would render completion of the project impossible), with the time durations indicated on the lines preceding each job. As an example, we give the following hypothetical analysis of the process of preparing an article for publication, together with its associated Hasse diagram, in Figure 10. If any two jobs, neither of which precedes the other, can be done by different people, then the minimum

---

[20] J.E.Kelley, Jr., Critical-path planning and scheduling: mathematical basis, Operations Res. 9(1961), 296-320.

length of time necessary to complete the project is the (weighted) length of the longest path from "start" to "finish".

job	description	immediate predecessors	time(days)
a	start	none	0
b	write body of article	a	3
c	write introduction	b	2
d	write conclusion	b	1
e	draw diagrams	b	1
f	compile bibliography	c,d	1
g	revise article	c,d,e	2
h	finish	f,g	0

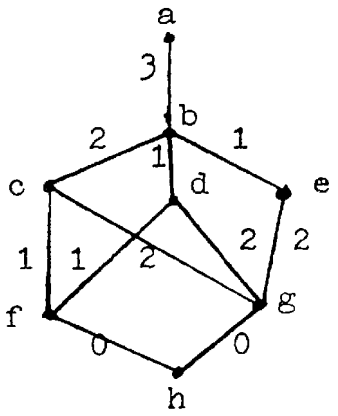


Figure 10. Analysis and Hasse diagram of a project

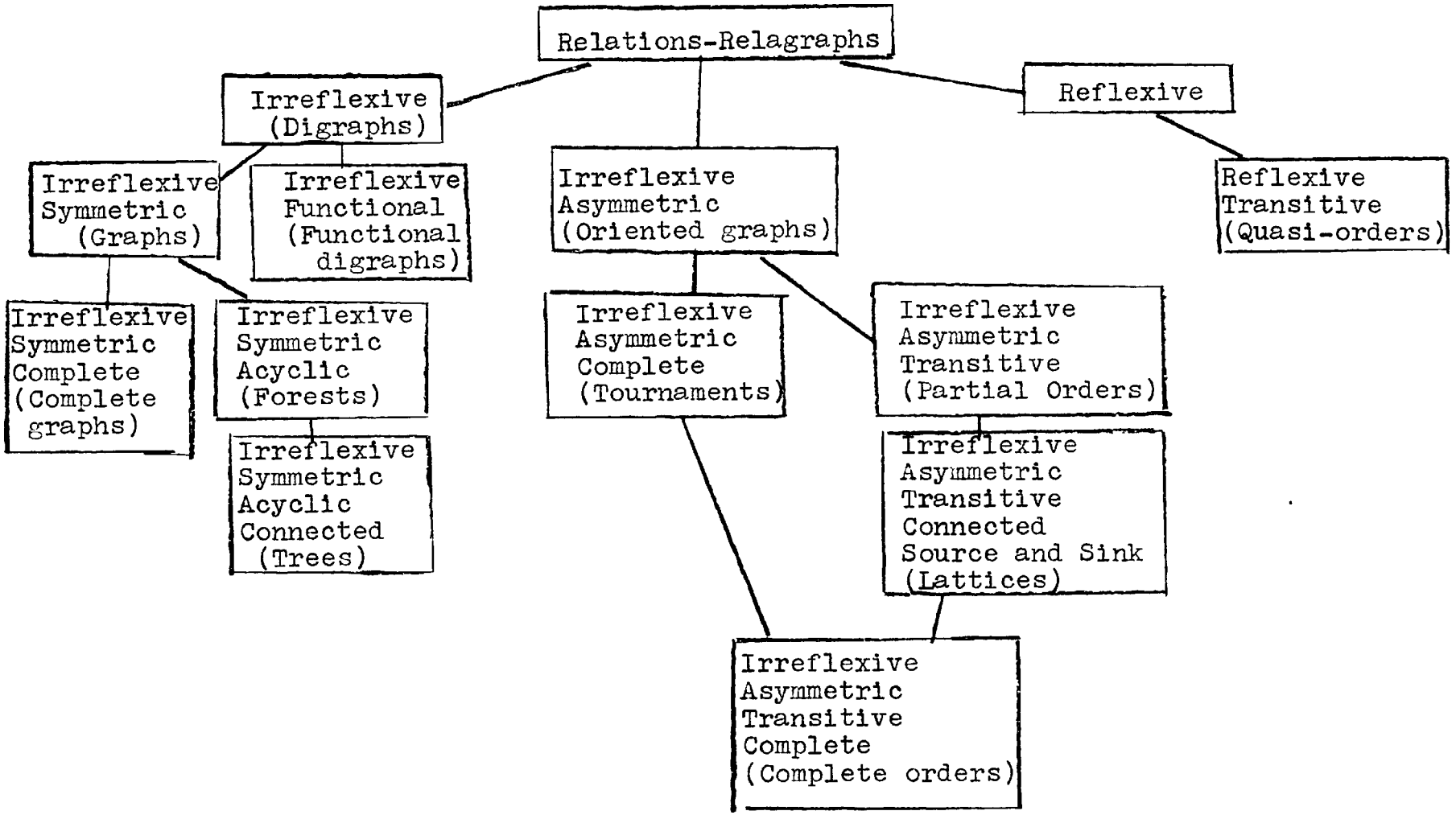
A significant large-scale model used in economics, called the Leontief input-output model,[21] provides an interesting application of digraphs. The points of the digraph represent industries and households, or, more generally, producers and consumers, and the arcs represent the flow of goods, labor, or funds. (Transactions within

---

[21] W.W.Leontief, The Structure of the American Economy, Oxford University Press, New York, 1951.

an industry are not included.) These models are used to study general flows in the economy such as employment, income, etc. When numbers are placed on the arcs, the model provides a logical method of viewing statistics. The model also shows, for example, how a strike in one sector of the economy affects other sectors. The size of the model, however, usually makes a matrix representation more convenient than the digraph.

This list of the uses of graph theory could be extended to include topics in biology, organization theory, genealogy, theoretical physics, and many other areas, but the methods of application would not differ. In each topic a set of objects--people, animals, jobs--provides the point set of the graph and a binary relation--authority, prey, status level--between these objects is singled out to induce the lines or arcs. A new type of graph, such as a signed digraph or a graph whose lines have been assigned weights, may be invented to handle a given situation. After the appropriate diagram is selected, the applications vary greatly in the extent to which they utilize graphical results; it depends on the type of problem which is important in each area. Nevertheless, graph theory provides a sound logical framework in which certain properties can be viewed and the full strength of this representation in most areas is yet to be determined.



APPENDIX. Chart showing relations studied

BIBLIOGRAPHY

Books

C.Berge, The Theory of Graphs and its Applications, translated by Alison Doig, Methuen, London, 1962.

C. Berge and A.Ghouila-Houri, Programming, Games, and Transportation Networks, Methuen, London, 1965.

G.Birkhoff, Lattice Theory, Amer. Math. Soc. Colloquium Publ., 25, 1948.

R.H.Bruck, A Survey of Binary Systems, Springer-Verlag, Berlin, 1958.

R.G.Busacker and T.L.Saaty, Finite Graphs and Networks, McGraw Hill, New York, 1965.

E.Cassignol and Y.Chow, Linear Signal-flow Graphs and Applications, Wiley, New York, 1962.

R.Courant and H.Robbins, What is Mathematics?, Oxford University Press, New York, 1941.

H.S.M.Coxeter and W.O.J.Moser, Generators and Relations for Discrete Groups, Springer-Verlag, Berlin, 1957.

R.A.Fisher and F.Yates, Statistical Tables for Biological, Agricultural, and Medical Research, Hafner, New York, 1963.

C.Flament, Applications of Graph Theory to Group Structure, Prentice-Hall, Englewood Cliffs, N.J., 1963.

L.R.Ford, Jr., and D.R.Fulkerson, Flows in Networks, Princeton University Press, Princeton, 1962.

I.Grossman and W.Magnus, Groups and Their Graphs, Random House, New York, 1964.

M.Hall, Combinatorial Theory, Blaisdell, Waltham, Mass., 1967.

F.Harary, Theory of Graphs, forthcoming.

F.Harary, R.Norman, and D.Cartwright, Structural Models, Wiley, New York, 1965.

J.G.Kemeny and J.L.Snell, Finite Markov Chains, Van Nostrand, New York, 1960.

_____, Mathematical Models in the Social Sciences, Ginn, New York, 1962.

D.König, Theorie der Endlichen und Unendlichen Graphen, Academische Verlagsgesellschaft M.B.H., Leipzig, 1936.

W.W.Leontieff, The Structure of the American Economy, Oxford University Press, New York, 1951.

K.Lewin, Principles of Topological Psychology, translated by F.Heider and G.M.Heider, McGraw Hill, New York, 1936.

E.S.Ljapin, Semigroups, Translations of Mathematical Monographs, 3, Amer. Math. Soc., Providence, R.I., 1963.

H.B.Mann, Analysis and Design of Experiments, Dover, New York, 1949.

O.Ore, The Four-Color Problem, Academic Press, New York, 1967.

_____, Graphs and Their Uses, Random House, New York, 1963.

_____, Theory of Graphs, Amer. Math. Soc., Colloquium Publ., 38, 1962.

J.Riordan, An Introduction to Combinatorial Analysis, Wiley, New York, 1958.

D.E.Rutherford, Introduction to Lattice Theory, Hafner, New York, 1965.

H.J.Ryser, Combinatorial Mathematics, Carus Mathematical Monographs, 14, Mathematical Association of America, Buffalo, 1963.

S.Seshu and M.B.Reed, Linear Graphs and Electrical Networks, Addison-Wesley, Reading, Mass., 1961.

D.Shanks, Solved and Unsolved Problems in Number Theory, Vol. I, Spartan Books, Washington, D.C., 1962

C.E.Shannon and W.Weaver, The Mathematical Theory of Communications, University of Illinois Press, Urbana, 1949.

T.Storer, Cyclotomy and Difference Sets, Markham, Chicago, 1967.

H.C.White, An Anatomy of Kinship, Prentice-Hall, Englewood Cliffs, N.J., 1963.

Articles

Chapter 1

T.van Aardenne-Ehrenfest, and N.G.deBruijn, Circuits and trees in oriented graphs, Simon Stevin 28(1951), 203-217.

L.R.Alvarez, Undirected graphs realizable as graphs of modular lattices, Canad. J. Math. 17(1965), 923-932.

S.S.Anderson and F.Harary, Trees and unicyclic graphs, The Math. Teacher 60(1966), 345-348.

L. Beineke and F.Harary, On the thickness of the complete graph, Bull. Amer. Math. Soc. 70(1964), 618-620.

R.L.Brooks, C.A.B.Smith, A.H.Stone, and W.T.Tutte, The dissection of rectangles into squares, Duke Math. J. 7(1940), 312-340.

I.Fary, On straight line representation of planar graphs, Acta Sci. Math. (Szeged) 11(1948), 229-233.

F.Harary, The number of functional digraphs, Math. Ann. 138 (1959), 203-210.

F.Harary and L.Moser, The theory of round robin tournaments, Amer. Math. Monthly 73(1966), 231-246.

F.Harary, E.Palmer, and C.Smith, Which graphs have only self-converse orientations?, Canad. Math. Bull, 10(1967), 425-429.

F.Harary and W.T.Tutte, A dual form of Kuratowski's theorem, Bull. Amer. Math. Soc. 71(1965), 168.

M.C.Jordan, Sur les assemblages de lignes, J.Reine Angew. Math. 70(1869), 185-190.

G.Kirchhoff, Uber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung Galvanischer Ströme geführt wird, Annalen der Physik, 72(1847), 497-508. Translated by J.B.O'Toole in IEEE Trans. Circuit Theory 5(1959), 4-7.

G.Kuratowski, Sur le problème des courbes gauches en topologie, Fund. Math. 15(1930), 271-283.

J.W.Moon, On subtournaments of a tournament, Canad. Math. Bull. 9(1966), 297-301.

L.Redei, Ein Kombinatorischer Satz, Acta Sci. Math. (Szeged) 7(1934), 39043.

H.E.Robbins, A theorem on graphs; with an application to a problem of traffic control, Amer. Math. Monthly 46 (1939), 281-283.

C.Smith and W.T.Tutte, On unicursal paths in a network of degree 4, Amer. Math. Monthly 48(1941)233-237.

W.T.Tutte, The factors of graphs, Canad. J. Math. 4(1952) 314-328.

E.S.Wolk, The comparability graph of a tree, Proc. Amer. Math. Soc. 13(1965), 789-795.

_____, A note on "The comparability graph of a tree", Proc. Amer. Math. Soc. 16(1965), 17-20.

Chapter 2

G.D.Birkhoff, A determinant formula for the number of ways of coloring a map, Ann. of Math. 14(1912), 42-46.

H.R.Brahana, The four color problem, Amer. Math. Monthly 30(1923), 234-243.

R.L.Brooks, On coloring the nodes of a network, Proc. Amer. Cambridge Philos. Soc. 37(1941), 194-197.

P.Franklin, The four color problem,Galois Lectures, Scripta Mathematica Library No 5, (1941), 49-85.

F.Harary and W.T.Tutte, A dual form of Kuratowski's theorem, Bull. Amer. Math. Soc. 71(1965), 168.

A.B.Kempe, On the geographical problem of the four colors, Amer. J. Math. 2(1879), 193-200.

K.O.May, The origin of the four-color conjecture, Isis 56(1965), 346-348.

J.Petersen, Sur le theorème de Tait, L'Intermédiare des mathématiciens 5(1898), 225-227.

_____, Die Theorie der regularen Graphen, Acta Math. 15(1891), 193-220.

P.Stäckel, Uber Nachargebiete im Raum, Zeitschrift fur Math. und Physik 42(1897), 275-276.

P.G.Tait, On the colouring of maps, Proc. Roy. Soc. Edinburgh Sect A 10(1880), 501-502, 729.

K.Wagner, Uber eine Eigenschaft der ebenen Komplexe, Math. Ann. 114(1937), 570-590.

H.Whitney, Non-separable and planar graphs, Trans. Amer. Math. Soc. 34(1932), 339-362.

Chapter 3.

R.P.Dilworth, A decomposition theorem for partially ordered sets, Ann. of Math. 51(1950), 161-166.

G.A.Dirac, Extensions of Menger's theorem, J. London Math. Soc. 38(1963), 148-161.

L.Dulmage and N.S.Mendelsohn, Some generalizations of the problem of distinct representatives, Canad. J. Math. 10(1958), 230-241.

P.Elias, A.Feinstein, and C.Shannon, A note on maximum flow through a network, IEEE Trans. Information Theory, 2(1956), 117-119.

L.R.Ford, Jr., and D.R.Fulkerson, Network flows and systems of representatives, Canad. J. Math. 10(1959), 78-84.

D.R.Fulkerson, On Dilworth's theorem, Proc. Amer. Math. Soc. 7(1956), 701-702.

P.Hall, On representatives of subsets, J. London Math. Soc. 10(1935), 26-30.

P.R.Halmos and H.E.Vaughan, The marriage problem, Amer J. Math. 72(1950), 214-215.

A.J.Hoffman and H.W.Kuhn, On systems of distinct representatives, in Linear Inequalities and Related Systems, Princeton University Press, Princeton, 1956, 199-206.

D.König, Graphok es matrixok, Matematikai es Fizikai Lapok 38(1931), 116-119.

H.B.Mann and H.J.Ryser, Systems of distinct representatives, Amer. Math. Monthly 60(1953), 397-401.

K.Menger, Zur Allgemeinen Kurventheorie, Fund Math. 10 (1927), 96-115.

L.Mirsky and H.Perfect, Systems of representatives, J.Math. Anal. Appl. 15(1966), 520-568.

H.Whitney, Congruent graphs and the connectivity of graphs, Amer. J. Math. 54(1932), 150-168.

Chapter 4

A.S.Amitsur and J.Levitski, Minimal identities for algebras, Proc. Amer. Math. Soc. 1(1950), 449-463.

L.Collatz and U.Sinogowitz, Spektren endlichen Graphen, Abh. Math. Sem. Univ. Hamburg 21(1957), 63-77.

F.Harary, The determinant of the adjacency matrix of a graph, SIAM Rev 4(1962), 202-210.

_____, A graph-theoretic method for the complete reduction of a matrix with a view toward finding its eigenvalues, J.Math. Physics 38(1959), 104-111.

A.J.Hoffman, On the polynomial of a graph, Amer. Math. Monthly 70(1963), 30-36.

A.J.Hoffman and M.H.McAndrew, The polynomial of a directed graph, Proc. Amer. Math. Soc. 10(1965), 303-309.

B.Kostant, A theorem of Frobenius, a theorem of Amitsur-Levitski, and cohomology theory, J.Math. Mech. 7(1958), 237-264.

R.G.Swan, An application of graph theory to algebra, Proc. Amer. Math. Soc. 14(1963), 367-373.

Chapter 5

R.Artzy, Cayley diagrams of binary systems, Duke Math. J. 28(1961), 491-495.

A.Cayley, The theory of groups: a graphical representation, Amer. J. Math. 1(1879), 174-176.

C.-Y. Chao, On a theorem of Sabidussi, Proc. Amer. Math. Soc. 15(1964), 291-292.

R.Frucht, Graphs of degree three with a given abstract group, Canad. J. Math. 1(1949), 365-378.

_____, Herstellung von Graphen mit vorgegebener abstrakter Gruppe, Compositio Math. 6(1938), 239-250.

F.Harary and E.Palmer, The smallest graph whose group is cyclic, Czech.Math.J. 16(91)(1966), 70-71.

R.L.Hemminger, Directed graphs with transitive, abelian groups, Amer. Math. Monthly 74(1967), 1233-1234.

_____, On the group of a directed graph, Canad. J. Math. 18(1966), 211-220.

H.Izbicki, Reguläre Graphen 3.4. und 5 Grades mit vorgegebenen abstrakten Automorphismengruppen, Farbenzahlen, und Zusammenhängen, Monatsh. Math. 61(1957), 42-50.

I.N.Kagno, Linear graphs of degree ≤ 6 and their groups, Amer. J. Math. 68(1946), 505-520, corrections in 69(1947), 872, and 77(1955), 392.

W.Maschke, The representation of finite groups, Amer. J. Math. 18(1896), 156-194.

J.W.Moon, Tournaments with a given automorphism group, Canad. J. Math. 16(1964), 485-489.

G.Sabidussi, The composition of graphs, Duke Math. J. 26(1959), 693-696.

_____, Graphs with given group and given graph-theoretical properties, Canad. J. Math. 9(1957), 515-525.

_____, On the minimum order of graphs with given automorphism group, Monatsh. Math. 63(1959), 124-127.

Chapter 6

L.D.Baumert, Hadamard matrices of orders 116 and 232, Bull. Amer. Math. Soc. 72(1966), 237.

L.D.Baumert, S.W.Golomb, and M.Hall, Discovery of an Hadamard matrix of order 92, Bull. Amer. Math. Soc. 68(1962), 237-238.

L.D.Baumert and M.Hall, A new construction for Hadamard matrices, Bull. Amer. Math. Soc. 71(1965), 169-170.

J.W.Bergquist, Difference sets and congruences modulo a product of primes, Ph.D.Dissertation, University of Southern California, 1963.

R.C.Bose and S.S.Shrikhande, On the construction of sets of mutually orthogonal Latin squares and the falsity of a conjecture of Euler, Trans. Amer. Math. Soc. 95(1960), 191-209.

_____, On the falsity of Euler's conjecture about the nonexistence of two orthogonal Latin squares of order $4t+2$, Proc. Mat. Acad. Sci. U.S.A. 45(1959), 734-737.

R.C.Bose, S.S.Shrikhande, and E.T.Parker, Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture, Canad. J. Math. 12(1960), 189-203.

R.H.Bruck and H.J.Ryser, The nonexistence of certain finite projective planes, Canad. J. Math. 1(1949), 88-93.

S.Chowla, A property of biquadratic residues, Proc. Nat. Acad. Sci. India 14(1944), 45-46.

S.Chowla and H.J.Ryser, Combinatorial problems, Canad.J. Math. 2(1950), 93-99.

H.Ehlich, Neue Hadamard-Matrizen, Arch. Math. 16(1965), 34-35.

K.Goldberg, Hadamard matrices of order cube plus 1, Proc. Amer. Math. Soc. 72(1966), 237.

M.Hall, Cyclic projective planes, Duke Math. J. 14(1947), 1079-1090.

_____, A survey of combinatorial analysis, in Kaplansky, Some Aspects of Analysis and Probability, Wiley, New York, 1958, 37-104.

_____, Uniqueness of the projective plane with 57 points, Proc. Amer. Math. Soc. 4(1953), 912-916, corrections in 5(1954), 994-997.

M.Hall, J.D.Swift, and R.J.Walker, Uniqueness of the projective plane of order eight, Math. Tables Aids Comput. 10(1956), 186-194.

A.J.Hoffman, On the line graph of a projective plane, Proc. Amer. Math. Soc. 16(1965), 197-202.

R.Laskar, and T.A.Dowling, A geometric characterization of the line graph of a projective plane, Institute of Statistics Mimeo Series No. 516, University of North Carolina at Chapel Hill, 1967.

E.Lehmer, On residue difference sets, Canad. J. Math, 5 (1953), 425-432.

H.W.Norton, The 7 x 7 squares, Annals of Eugenics, 9 (1939), 267-307.

R.E.A.C.Paley, On orthogonal matrices, Jour. Math. and Phys. 12(1933), 311-320.

E.T.Parker, Orthogonal Latin squares, Proc. Nat. Acad. Sci. U.S.A. 45(1959), 859-862.

W.A.Pierce, The impossibility of Fano's configuration in a projective plane with eight points per line, Proc. Amer. Math. Soc. 4(1953), 908-912.

A.Sade, An omission in Norton's list of the 7 x 7 squares, Ann. Math. Statist, 22(1951), 306-307.

J.Singer, A theorem in finite projective geometry and some applications to number theory, Trans. Amer. Math. Soc. 43(1938), 377-385.

E.Spence, A new class of Hadamard matrices, Glasgow Math. J. 8(1967), 59-62.

R.G.Stanton, and D.A.Sprott, A family of difference sets, Canad. J. Math. 10(1958), 73-77.

G.Tarry, Le problème de 36 officieurs, Compte Rendu de l'Association Francaise pour l'Avancement de Science Naturel, 1(1960), 122-123; 2(1901), 170-203.

J.A.Todd, A combinatorial problem, Jour. Math. and Phys. 12(1933), 321-333.

O.Veblen and J.H.M.Wedderburn, Non-desarguesian and non-pascalian geometries, Trans. Amer. Math. Soc. 8(1907), 379-388.

M.B.Wells, The number of Latin squares of order eight, J.Combinatorial Theory, 3(1967), 98-99.

A.L.Whiteman, A family of difference sets, Illinois J. Math. 6(1962), 107-121.

J.Williamson, Hadamard's determinant theorem and the sum of four squares, Duke Math. J. 11(1944), 65-81.

Chapter 7

S.S.Anderson and G.Chartrand, The lattice-graph of the topology of a transitive directed graph, to appear in Math. Scand.

M.Behzad and H.Radjavi, Another analog of Ramsey numbers, unpublished manuscript, n.d.

_____, The line analog of Ramsey numbers, Israel J. Math. 5(1967), 93-96.

_____, On a problem of Erdos, unpublished manuscript, n.d.

R.L.Brooks, C.A.B.Smith, A.H.Stone, and W.T.Tutte, The dissection of rectangles into squares, Duke Math. J. 7(1940), 312-340.

D.W.Crowe, The n-dimensional cube and the tower of Hanoi, Amer. Math. Monthly 63(1956), 29-30.

P.Erdos, Graph theory and probability, I, Canad. J. Math. 11(1959), 34-38.

_____, Graph theory and probability, II, Canad. J. Math. 13(1961), 346-352.

_____, Remarks on a theorem of Ramsey, Bull. Res. Council Israel 7(1957), 21-24.

P.Erdos and G.Szekeres, A combinatorial problem in geometry, Compositio Math. 2(1935), 463-470.

L.Euler, Solutio problematis ad geometriam situs pertinentis, Commentarii Academiae Petropolitanae 8(1736), 128-140, translated in Sci. Amer. 189 (July, 1953), 66-70.

J.W.Evans, F.Harary, and M.S.Lynn, On the computer enumeration of finite topologies, Comm. A.C.M. 10(1967), 295-297.

J.E.Graver and J.Yackel, An upper bound for Ramsey numbers, Bull. Amer. Math. Soc. 72(1966), 1076-1079.

R.E.Greenwood and A.M.Gleason, Combinatorial relations and chromatic graphs, Canad. J. Math. 7(1955), 1-7.

J.G.Kalbfleisch, Chromatic graphs and Ramsey's theorem, Ph.D. Dissertation, University of Waterloo, Canada, 1966.

_____, Construction of special edge-chromatic graphs, Canad. Math. Bull. 8(1965), 575-584.

_____, On an unknown Ramsey number, Michigan Math. J. 13(1966), 385-392.

F.P.Ramsey, On a problem of formal logic, Proc. London Math. Soc. 30(1929), 264-286.

Chapter 8

A.Bavelas, A mathematical model for group structures, Applied Anthropology, 7(1948), 16-30.

L. Beineke and F.Harary, On the thickness of the complete graph, Bull. Amer. Math. Soc. 70(1964), 618-620.

N.G.deBruijn, A combinatorial problem, Indag. Math. 8(1946), 461-467.

A.Cayley, On the analytical forms called trees, with applications to the theory of chemical combinations, The Collected Mathematical Papers of Arthur Cayley, Cambridge University Press, Cambridge, 1889-1897, IX, 427-460.

_____, On the mathematical theory of isomers, The Collected Mathematical Papers of Arthur Cayley, Cambridge University Press, Cambridge, 1889-1897, IX, 202-204.

_____, On the theory of the analytical forms called trees, The Collected Mathematical Papers of Arthur Cayley, Cambridge University Press, Cambridge, 1889-1897, XIII. 242-246.

_____, A theorem on trees, The Collected Mathematical Papers of Arthur Cayley, Cambridge University Press, Cambridge, 1889-1897, XIII, 26-28.

J.E.Kelley,Jr., Critical-path planning and scheduling: mathematical basis, Operations Res. 9(1961), 292-320.

G.Kirchhoff, Uber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung Galvanischer Ströme geführt wird, Annalen der Physik, 72(1847), 497-508. Translated by J.B.O'Toole in IEEE Trans. Circuit Theory 5(1958), 4-7.

H.J.Leavitt, Some effects of certain communication patterns on group performance, J.Abn. Soc. Psychol. 46(1951), 38-50.

W.W.Leontieff, The structure of development, Sci. Amer. (Sept., 1963), 148-166.

J.C.Morrissette, An experimental study of the theory of structural balance, Hum. Rel. 11(1958), 329-354.

A.Nerode and H.Shank, An algebraic proof of Kirchhoff's network theorem, Amer. Math. Monthly 68(1961), 244-247.

R.C.Prim, Shortest connection networks and some generalizations, Bell System Tech. J. 36(1957), 1389-1401.