NORTH-HOLLAND

MATHEMATICS STUDIES

121

ANNALS
OF
DISCRETE
MATHEMATICS

29

Matching Theory

L. LOVÁSZ
M.D. PLUMMER

NORTH-HOLLAND

MATCHING THEORY

NORTH-HOLLAND MATHEMATICS STUDIES 121
Annals of Discrete Mathematics (29)

NORTH-HOLLAND – AMSTERDAM • NEW YORK • OXFORD • TOKYO

MATCHING THEORY

by

L. LOVÁSZ
Department of
Computer Science,
Eötvös University
Budapest
Hungary

and

M. D. PLUMMER
Department of Mathematics
Vanderbilt University
Nashville U.S.A.

1986

NORTH-HOLLAND – AMSTERDAM • NEW YORK • OXFORD • TOKYO

ISBN: 0 444 87916 1

Publishers:

ELSEVIER SCIENCE PUBLISHERS B.V.
P.O. BOX 1991
1000 BZ AMSTERDAM
THE NETHERLANDS

and

AKADÉMIAI KIADÓ, BUDAPEST

Sole distributors for the U.S.A. and Canada:

ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 VANDERBILT AVENUE
NEW YORK, N.Y. 10017
U.S.A.

For the East European countries, Democratic People's Republic of Korea, People's Republic of China, People's Republic of Mongolia, Republic of Cuba and Socialist Republic of Vietnam:

AKADÉMIAI KIADÓ, BUDAPEST

PRINTED IN HUNGARY

Table of Contents

This Page Intentionally Left Blank

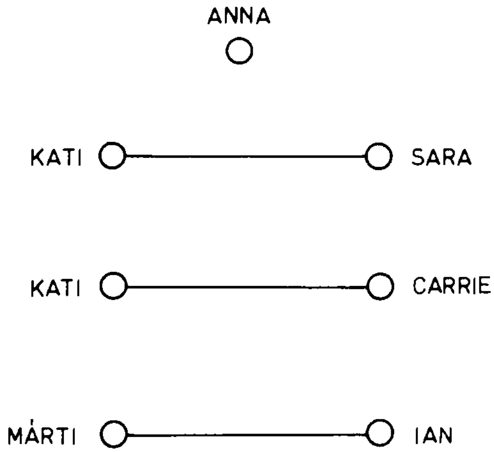


FIGURE P.1.

Preface

Suppose a pharmaceutical firm wishes to test n antibiotics on n volunteer subjects. However, preliminary screening has shown that certain subjects are allergic to certain of the drugs. Can an experiment be designed in which each subject takes exactly one of the antibiotics to which he/she is not allergic and each drug is taken by exactly one subject?

Let us model this situation using a bipartite graph G in which the two point classes consist of the n subjects and the n antibiotics respectively and let us agree to join subject to drug if and only if the subject is not allergic to the drug. Then the answer to the question posed is "yes" if and only if graph G has a **perfect matching** (or **1-factor**), that is, a pairing of subjects to drugs which uses each subject and each drug once and only once.

Next suppose one has two computers available and p jobs to be processed on these machines. We will assume that any job can be run on either machine. Indeed, we may assume that the computers are identical. Let us also suppose that the p jobs are partially ordered in the sense that for any two (different) jobs J_i and J_k, $J_i \leq J_k$ if J_i must be completed before J_k can be started (by either computer). If all jobs require an equal amount of time to complete, what is the shortest possible time sufficient to run all p jobs?

Let us model this situation using an undirected graph G as follows. Let the points of G be the jobs $J_1, \ldots, J_p$ and let us join J_i and J_k with a line if and only if they are incomparable in the partial order. (In other words, if they can be run simultaneously.) Now it is clear that to design an optimum schedule, we must use both machines simultaneously as often as possible; that is , we must find a matching of largest cardinality in G. (See Fujii, Kasami and Ninomiya (1969, 1971) and Coffman and Graham (1972).) This problem belongs to the class of so-called **maximum cardinality matching** problems. Note also that in this case, unlike the drug testing example above, the graph modeling the problem is no longer bipartite.

Now let us get our hands dirty! Suppose we are drilling for oil. Seismic (and/or other) information has been supplied to us which gives accurate locations of oil deposits. The existing technology at the time of this writing is such that one can drill until one deposit of oil is tapped and then drilling may be continued in the same hole to a second deposit situated at a location somewhat deeper, but not necessarily directly beneath, the first deposit. (See Figure P.2.). One can then bring up oil from both deposits through two concentrically placed pipes in the same drill hole. (For further technical comments see Devine (1973).)

One can associate a number with each pair of deposits representing the cost of tapping them both in one drilling operation. Practical impossibility of pairing can be reflected by assigning a very large (or even infinite) value to a pair. For the sake of simplicity, let us assume that an even number of deposits is under consideration so that all deposits may be simultaneously paired off.

The task is then to find a set of pairs (that is, a "perfect matching") of deposits in which all deposits are paired and so that the sum of the costs is minimum. In other words, we seek a **minimum weight perfect matching** in the associated complete graph.

To put our next example in a bit more facetious form, suppose that $2n$ students arrive at Nashpest University at the beginning of the school year and are to be assigned to n 2-person dormitory rooms. The housing office has given each student a roster of all $2n$ students and asked each student to check off those other students on the list with whom they are acquainted. This information is then fed into the University computer. But, alas, on registration day the computer "crashes"! (Such has been known to happen on occasion!) The Dean of Housing, in despair, randomly assigns two students to each room. What is the probability that only students acquainted with each other are paired by the Dean's

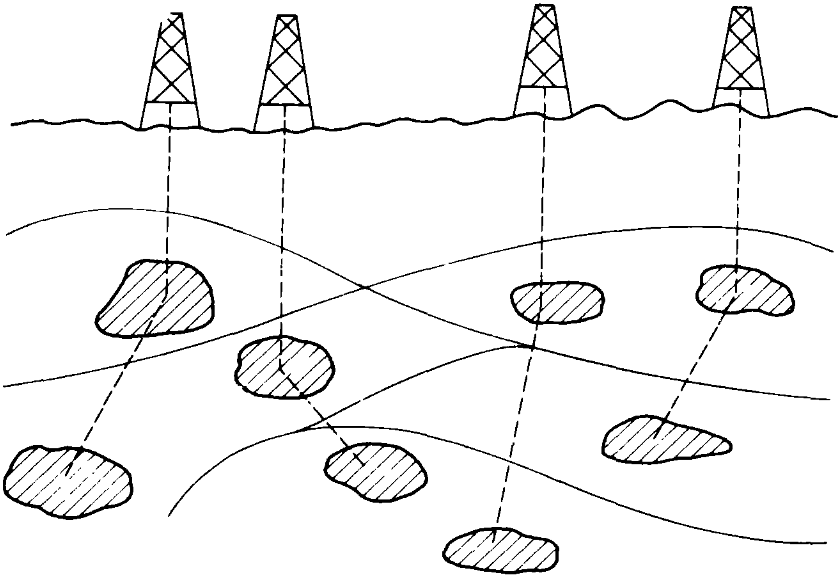


FIGURE P.2.

random assignment?

Let us build a graph G with the $2n$ students as points and agree to join two students if they are acquainted with each other. The answer to the room assignment question above is then clearly $\Phi(G)/(2n-1)!!$ where $\Phi(G)$ denotes the number of perfect matchings in G.

Sometimes problems which are not directly convertible into matching problems can nevertheless be solved with the help of matching theory. **The Chinese Postman Problem** (see Meigu Guan (1962)) was first presented as follows. A postman must deliver mail along all streets of a town. How can he leave the post office, perform the "swift completion of his appointed rounds" and return to the post office having traversed a minimum distance?

Here there is no longer an obvious translation of the problem directly into a matching problem. But we shall see later that this problem can be efficiently solved by first solving a set of shortest path problems and then solving a certain minimum weight perfect matching problem.

Let us now modify the last problem a bit. Suppose our postman's task is to pick up mail from boxes located at street intersections only.

Again, he wants to find a shortest route. The main difference between this problem and the previous one is that now he has to visit all *intersections*, but does not have to traverse all the *streets*. In a more common formulation of this problem, a travelling salesperson has to visit all capitals of states in the U.S. (or all the counties of Hungary) in such a way as to minimize the length (or time, or cost) of the route. This problem is known as the **Travelling Salesman Problem** and it has received quite a lot of attention in the "real" world.

As a special case, the salesperson may only be interested in finding a tour which avoids certain connections. In such a case we retain the lines representing allowable connections and delete all other lines. The problem then reduces to finding a Hamilton cycle (that is, a cycle through all the points) in the given graph. This problem is called **The Hamilton Cycle Problem.**

Superficially, the Travelling Salesman Problem seems quite similar to the Chinese Postman Problem and to the Weighted Perfect Matching Problem. However, the Travelling Salesman Problem (and even its special case, the Hamilton Cycle Problem) turn out to be significantly more difficult. More particularly, the Chinese Postman Problem is solvable in polynomial time, but the Travelling Salesman Problem is "NP-hard" (i.e., it belongs among the hardest combinatorial problems). So here we have an example of a generalization of matching to a truly more difficult problem.

But matching theory is often used as a stepping stone to the Travelling Salesman Problem. A "tour" for a salesperson is a connected spanning subgraph in which all points have degree 2. If we drop the connectedness assumption we arrive at the concept of a **2-factor**. In this way the 2-factor is a natural extension of the notion of a perfect matching and in fact the 2-factor problem can be reduced to finding a perfect matching. In this light the Matching Problem can be viewed as a "relaxation" of the Travelling Salesman Problem.

As our last example, let us describe an important and well-known combinatorial problem, but one not usually thought of as being related to matching. A telephone company wishes to provide service to each of n cities so that any city may call any other. How can the cities be connected so that construction costs are minimized?

Here the solution is clearly a minimum weight spanning tree. This problem can be rather easily solved by the so-called Greedy Algorithm (Borůvka (1926a, 1926b), Kruskal (1956)) in which at each step one selects the cheapest "allowable" line.

But this "greedy" procedure does not work for the minimum weight perfect matching problem and in fact matching problems are nearly always more difficult to solve. Why? In modern parlance the reason lies in the fact that the spanning trees of a graph form a "matroid", whereas the perfect matchings do not. Nevertheless we shall see that matroids do indeed arise at many points in the study of matchings.

So we see that matching and some of its close relatives model rather a lot of applied problems. But does this warrant writing a book on the subject? In a sense, matching problems seem somehow to be of the "proper" level of difficulty. On the one hand, most of them are solvable problems, but to solve them certainly requires non-trivial methods. Indeed matching theory has played a catalytic role over the past hundred years or so in developing a number of new and more general combinatorial methods. To see more concrete manifestations of this, let us take a look at some history.

It is fascinating (if not downright dangerous!) to probe the roots of any family tree. Although it can be argued that such famous names as Euler, Kirchhoff and Tait can be found in the historical shadows of matching theory, we shall take as the two principal "founders" of the discipline the Dane, Julius Petersen and the Hungarian, Dénes König. Although their interests certainly overlapped, it is perhaps helpful to identify Petersen with the earliest study of *regular* graphs (that is, graphs having the same degree at each point) and König with *bipartite* graphs.

In an 1891 paper, Petersen considered an algebraic factorization problem due to Hilbert (1889) and reformulated it as a factorization problem for graphs. He set as his general problem the task of deciding which regular graphs have a non-trivial factorization into smaller regular spanning subgraphs the union of which is the parent graph.

He then proceeded to prove that any graph regular of even degree can be expressed as the union of line-disjoint 2-factors. This result is intimately related to the famous result of Euler who showed in his celebrated paper on the Königsberg Bridge Problem (1736) that one can traverse all lines of a graph once and only once and return to the starting point without stopping if and only if the graph under consideration is connected and has all degrees even. Petersen does not mention Euler in his paper and indeed we do not know if he was aware of the Euler result, already at that time more than 150 years old.

Petersen quite accurately observed that factorization of graphs regular of *odd* degree is a more difficult problem. He then proved that any connected 3-regular graph having no more than two cutlines has a perfect

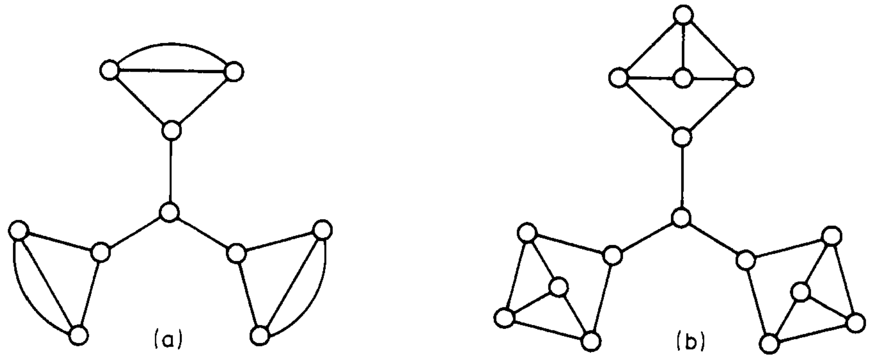


FIGURE P.3. Sylvester's graphs

matching (and thus is decomposable into a 1-factor and a 2-factor). Petersen pointed out that two cutlines was best possible in the sense that there are 3-regular graphs with three cutlines and no perfect matching. He gave the example in Figure P.3(a) and attributed it to Sylvester. (Figure P.3(b) contains the corresponding graph without multiple lines.)

Petersen's proof of his decomposition theorem for 3-regular graphs was successively simplified by Brahana (1917-18), Errera (1922) and by Frink (1925-26). The last of these contained a slight error which was corrected in König's book (1936). Petersen's work was extended to other regular graphs by Bäbler (1938, 1952, 1954), Gallai (1950) and by Belck (1950). This line of research culminated in Tutte's work which we shall discuss later.

Petersen also became interested in the famous Four Color Conjecture which had been formulated about fifty years before. We make no attempt in our book to chronicle the work on this problem, but refer the reader to Biggs, Lloyd and Wilson (1976) for a lively account of the early years of this conjecture. The conjecture says that the regions of any planar map (that is, the faces of any plane graph) can be colored in four colors in such a way that no two regions sharing a common boundary line receive the same color.

To connect this problem to Petersen we must introduce the Scot, P.G. Tait to our story. It already had been observed by Cayley and Kempe that in order to prove the Four Color Conjecture true in general, it would be enough to prove it true for 3-regular (that is, "cubic") planar graphs. In a series of three papers (1878-1880a, 1878-1880b, 1880) Tait then took up the problem (in a rather foggy manner). He made the important observation (1878-1880b) that 4-coloring the regions of a cubic planar map was equivalent to 3-coloring the *lines* of the map and then he turned his attention to factoring cubic planar graphs.

Tait provided an example to show that if a cubic graph has a cutline it may not be factorizable into three perfect matchings. But he then claimed that if the cubic graph were "polyhedral" such a factorization could always be carried out. (Steinitz (1922) later showed that "polyhedral" means precisely 3-connected and planar.) He then outlined various approaches to prove his claim which, of course, did nothing of the kind.

Tait also made a second conjecture, namely that every cubic polyhedral graph contains a Hamilton cycle. He then pointed out, accurately enough, that if this conjecture were true, 3-line colorability of all cubic polyhedral graphs would follow immediately for one could color the lines of the (necessarily even) Hamilton cycle with two colors and all other lines with the third color. This second conjecture of Tait enjoyed a long life too until Tutte (1946) found the first counterexample.

Of course had Tait's second conjecture been true, the Four Color Conjecture would have been proved. But as all mathematicians know, this celebrated problem remained open for nearly 100 years more until it was finally settled in the affirmative by Appel and Haken (1977) and Appel, Haken and Koch (1977).

But let us now return to Petersen who in 1898 published a second paper in which he responded to Tait's remarks on cubic graphs. The most significant contribution made to graph theory in this paper was undoubtedly the introduction of a non-planar graph which was cubic and had no cutlines, but which could not be decomposed into three disjoint perfect matchings. This graph, nowadays deservedly called the "Petersen graph", is perhaps the single most famous graph in existence (or the most notorious, depending upon what it has done to the reader's conjectures!) The graph is shown in Figure P.4. The reader will encounter this graph several times in the present book as well.

In the meantime, however, a second stream of historically important results relevant to matching had begun to flow. The emphasis this time was on *bipartite* graphs. The German algebraist, Frobenius, who had become interested in reducibility properties of determinants, proved the following result (1912). (In order to keep things sorted out subsequently, let us call this Theorem F-1.) Consider an $n \times n$ matrix M such that every entry is either zero or a variable, all variable entries being different from each other. Then the determinant of the matrix is a reducible (that is, factorizable) polynomial of these variables if and only if there is an

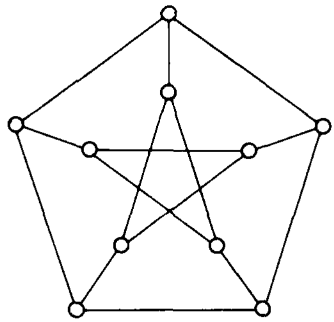


FIGURE P.4. The Petersen graph

integer $p, 0 \leq p \leq n$, and a permutation of the rows and columns of M which results in a block of zero entries of size $p \times (n-p)$. Frobenius' proof of F-1 was more complicated than need be and König (1915), having cast the problem in terms of bipartite graphs, gave a shorter proof.

In retrospect, König's translation of the above problem into graph-theoretical terms was indeed quite straightforward, but was to prove to be of considerable significance. Let the rows of a square matrix M be $r_1, \ldots, r_n$ and the columns be $c_1, \ldots, c_n$. Form a bipartite graph by joining r_i to c_j if and only if entry m_{ij} of matrix M is not zero. It is now immediate that there is a one-to-one correspondence between non-zero expansion terms in the determinant of M and perfect matchings in the bipartite graph constructed from M as above.

We should remark parenthetically here that relationships between graphs and linear algebra go much deeper than the matrix-bigraph model just mentioned. We will treat a number of these relationships in this book.

In the following year in two nearly identical papers — one in German (1916a), the other in Hungarian (1916b) — König proved that every doubly stochastic matrix with non-negative entries must have a non-zero term in its determinant. (An $n \times n$ matrix is **doubly stochastic** if the $2n$ row and column sums all have the same value.) The proof given is for integer matrices, but König pointed out that the result is clearly extendable to the rationals and then to the reals. In the proof he gave for integers he considered once again the corresponding bipartite graph and observed that the matrix is doubly stochastic if and only if this bigraph is regular. Also in his twin 1916 papers, König proved that every bipartite graph which is regular of degree k is the union of k disjoint perfect matchings. It is a trivial observation — once these results are proved — that any doubly stochastic matrix with non-negative elements must be a convex sum of permutation matrices. Couched in these terms, the result was to be independently rediscovered some forty years later by Birkhoff (1946) and von Neumann (1953) and has come to be known as the Birkhoff–von Neumann Theorem.

It is in the 1916 papers that one finds the first proof of yet another famous theorem for bipartite graphs. We shall call it the König Line Coloring Theorem. (See Theorem 1.4.18 of the present book.) This result says that in any (not necessarily regular) bipartite graph the chromatic index is equal to the maximum degree of all the points.

At this point the plot begins to thicken a bit! In 1917, Frobenius published his own simpler proof of Theorem F-1. He obtained his proof using the following lemma which we shall denote by F-2. Consider an $n \times n$ matrix M in which every entry is either zero or a variable, different entries being different variables. Suppose that the determinant of M vanishes identically as a polynomial of its non-zero entries. Then there is an integer $p, 0 < p < n$, and permutations of the rows and columns of M such that the resulting matrix contains a block of zeros of size $p \times (n-p+1)$.

It is somewhat surprising that the question as to whether a determinant is *irreducible* was studied much earlier by Frobenius than the more natural question as to whether it is *identically zero*. We all learned in school that a polynomial is identically zero if and only if upon expanding it, all terms cancel. From today's more algorithmic point of view, this is not a satisfactory test since it may involve exponentially long computations. But from the classical point of view, irreducibility was a much more intriguing question.

Frobenius did not cite König's 1915 proof of Theorem F-1, even though König contended that he had sent his proof to Frobenius in German translation. (See footnote 2, page 240 of König's book (1936).) However, Frobenius did cite König's paper (1916a), but he then dismissed König's graph-theoretic formulation of these determinant questions as being of little value, thereby heaping a few more coals upon the fire! For König's ripost the interested reader is referred to the above-mentioned footnote or to König (1933). For further commentary on this issue, see Schneider (1977).

Lemma F-2, in addition to its role in obtaining the short proof of F-1 referred to above, is of considerable interest in its own right, for if one translates it into the language of bigraphs, it gives a necessary and sufficient condition for a bipartite graph to have a perfect matching. In this guise it has come to be called the Marriage Theorem.

Suppose we have n men and n women and we wish to arrange n marriages (without bigamy, polyandry or homosexual relationships). Let us suppose further that we wish to marry only men and women who are acquainted with each other. The Marriage Theorem states that this is possible if and only if for each $k, 1 \leq k \leq n$, each set of k men collectively knows at least k women.

This theorem was the forerunner of one of the best known results in bipartite matching, the Theorem on Distinct Representatives due to Philip Hall (1935). Hall's Theorem was first stated in terms of sets and may be simply expressed as follows. Let $S_1, \ldots, S_n$ be a finite collection of sets. Then there is a set of distinct elements $x_1, \ldots, x_n$ such that $x_i \in S_i$ if and only if for each $k, 1 \leq k \leq n$, the union of any k of the S_i's contains at least k elements. It is clear that the Marriage Theorem of Frobenius follows from this result immediately. (In fact, the two are equivalent, but more about that in Chapter 1.)

At this point we must backtrack about four years to 1931 when the first proof appeared of the so-called König Minimax Theorem for bipartite graphs (Theorem 1.1.1 of our book). (See König (1931) and (1933) for Hungarian and German language versions respectively.) This result says that in a bipartite graph G the size of a largest matching is equal to the size of a smallest set of points which together touch every line of G. In the same year, Egerváry (1931) generalized this result to graphs with non-negative weights on each line. Such "minimax" results as these (the classical theorem of Menger (1927) on graph connectivity and the Max-Flow Min-Cut Theorem of Ford and Fulkerson (1956) and Elias, Feinstein and Shannon (1956) on network flows are but two of many others) are of great significance in our opinion and we shall emphasize this type of result throughout the book. Indeed the importance of such results grows daily in various branches of combinatorics, due largely to increasing use of linear programming to formulate and solve many combinatorial problems. (See Schrijver (1983a) for an excellent up-to-date survey of minimax results in combinatorics.)

In the first textbook on graph theory ever written, König (1936) showed that Menger's theorem on connectivity, the Marriage Theorem of Frobenius and P. Hall's Theorem on Distinct Representatives all follow from his minimax theorem. In fact, we shall see in Chapter 1 that these four results, together with the Max-Flow Min-Cut Theorem and an important theorem on partially ordered sets due to Dilworth (1950), are all *equivalent*.

For further interesting historical comments on the development of matching theory prior to the appearance of König's book in 1936, the reader is referred to the book of Biggs, Lloyd and Wilson (1976). Also Gallai has written a very interesting account of the life and work of König (1964b, 1978).

Very few papers on matching theory appeared during the years of the Second World War. In 1942, however, Rado published a paper in which he generalized P. Hall's theorem to independent systems of distinct representatives in Euclidean vector spaces and thereby established the first link between matchings and *matroids*. In 1945, Marshall Hall applied the notion of systems of distinct representatives to extend Latin rectangles to Latin squares.

In the immediate postwar period several exciting results burst upon the scene and from these matching theory received a vigorous boost. In 1947, Tutte proved a theorem characterizing those general (that is, non-bipartite) graphs with perfect matchings. (See Theorem 3.1.1 of this book.) This elegant result which, with the advantage of hindsight, is the "natural" generalization of the bipartite Marriage Theorem has become the cornerstone of matching theory in the non-bipartite case. Suppose a connected graph G contains a set S of points such that $G - S$ has more than $|S|$ components having an odd number of points each. It is then clear that G cannot contain a perfect matching. The crucial contribution of Tutte's Theorem was to prove the converse true; that is, if G has no separating set S with this property, then G must have a perfect matching. This theorem was destined to become an archetypal example of a "good characterization" in the language of algorithmic complexity theory. The latter discipline was born only in the 1970's as mathematicians scrambled to keep up with the explosive entry of the computer upon the scientific scene along with the accompanying surge of interest in algorithms. But before considering computers and algorithms per se, let us mention a few more important non-algorithmic results forthcoming during the two decades immediately following World War II.

In the early 1950's, Tutte proved his so-called "f-factor theorem" which has its roots back in the 1890's in the earlier work of Petersen. The concept of an f-factor and the closely related "degree-constrained subgraph problem" were also studied by Ore and Gallai in the 1950's and early 1960's. In 1955, Ore published his "defect" version of P. Hall's Theorem for bipartite graphs and in 1958, Berge obtained the analogous "defect" version of Tutte's perfect matching theorem for the non-bipartite case.

The first two textbooks in graph theory after König were due to Berge and Ore respectively. Berge's book appeared in its French edition in 1958 and in an English translation in 1962, the same year that Ore's book in English was published. Together they introduced graph theory to a much wider audience (including the second author of the present book) and their appearance set the stage for the incredible growth of graph theory in both breadth and depth seen in the last two decades.

In the late 1950's and early '60's, Dulmage and Mendelsohn published a series of papers in which they worked out a canonical decomposition theory for bipartite graphs in terms of maximum matchings and minimum point covers. Their work was motivated by questions concerning matrices.

The year 1964 marked the appearance of a paper by Gallai which contained one of the central results of this book. (See Chapter 3.) In this paper, Gallai established the existence of a *canonical* decomposition theory of any graph in terms of its maximum matchings. An efficient method to obtain this decomposition was provided by the polynomial matching algorithm for general graphs due to Edmonds which appeared in 1965. Hence we have chosen to call this important result the Gallai-Edmonds Structure Theorem.

One of several important degenerate cases for the Gallai-Edmonds theorem arises when the graph in question has a *perfect* matching. However, Kotzig had already begun to lay the foundations for a canonical treatment of these graphs in a series of papers which appeared in 1959 and 1960. It is unfortunate that these important papers remained more or less unnoticed since they were written in Slovak. In these publications, Kotzig introduced a certain binary relation on the point set of any graph having a perfect matching. For an important special class of such graphs, the so-called "elementary" graphs, this relation is an equivalence relation and thus induces a canonical partition of the point set.

But let us now return once more to the period immediately following World War II and take a second historical tack. Computers immediately focused attention on the development of algorithms naturally enough, but if we turn our attention back to matching theory, we see that a fundamental algorithmic question has been with us since the earliest days of the subject. Its importance is perhaps belied by its simplicity of statement: how do you find a perfect (or maximum) matching?

It is no surprise that the first matching algorithm did not spring forth "fully coded" from any one forehead! The rudiments for finding a maximum matching in a bipartite graph had already appeared in the works of König and Egerváry in the 1930's. Kuhn (1955) and M. Hall (1956) presented the first formal procedures for finding a perfect matching in a bigraph. It seems to have been Kuhn who at this time first used the phrase "Hungarian Method" to distinguish algorithms of this type.

At almost the same time, Ford and Fulkerson published the first papers on the theory of network flows (1956, 1957, 1962). Flow theory immediately became a substantial new tool in combinatorial applications of all kinds. Flows can be easily visualized and for our purposes flow theory is of great importance because it can be used to prove most results in bipartite matching.

Matching in non-bipartite graphs turned out to be substantially more difficult and almost another decade passed before Edmonds (1965a) found the first efficient algorithm to find a maximum matching in such a graph. This algorithm also motivated Edmonds to propose *polynomial time* as a measure of "goodness" of algorithms, a point of view which has proved extremely fruitful in theoretical computer science.

At this point we must pursue yet another historical branch. Another product of the early post-war years which has had enormous impact upon not only mathematics itself, but upon nearly every quantitative area of science, is *linear programming.*

In the late 1930's, Kantorovich seems to have been the first to cast linear programming as a mathematical theory in its own right, but his work remained unnoticed in the West. Moreover, no complete algorithm for solving a linear program had yet appeared.

Motivated by World War II planning activities, Dantzig and von Neumann independently discovered and developed the new subject. The important concept of *duality* was introduced by von Neumann (1947). Dantzig (1951) gave the infant discipline a giant practical boost when he introduced the algorithm known as the Simplex Method. This method has solved nearly all real-life linear programming problems far more efficiently than any other method known before or since. For a more extensive historical review of linear programming, see Dantzig's book (1963) and his more recent historical article (1983).

A link between linear programming and matching theory was soon discovered. In 1955, Kuhn published the first of several papers (1955, 1956) in which he cast *bipartite* matching — weighted and unweighted — in the primal-dual setting of linear programming for the first time. To obtain a combinatorial minimax theorem from linear programming duality, one needs a sufficient condition for the integrality of the optimum solutions. If the graph is bipartite integrality follows easily. The first more general sufficient condition for integrality to be discovered — *total unimodularity* — was found by Hoffman and Kruskal (1956). In 1958, Gallai used unimodularity and the duality theorem of linear programming to derive a number of minimax results including the Menger, Dilworth and Egerváry theorems as well as Max-Flow Min-Cut. Hoffman (1960) also accurately predicted that linear programming would become an important general tool in handling combinatorial optimization problems.

The extension of the linear programming approach to the case of *non*-bipartite graph matching turned out to be quite difficult, however, and it required the introduction of a new technique, namely the technique of describing the convex hull of incidence vectors of matchings by linear inequalities. Such a set of linear inequalities was found by Edmonds (1965b). His result allows us to obtain various minimax theorems in matching theory as special cases of the Duality Theorem of linear programming. This approach initiated the study of other combinatorially defined polyhedra and has led to a whole new branch of combinatorial mathematics — polyhedral combinatorics. (For a very recent survey of this discipline see Pulleyblank (1983).)

It is also a natural idea to try to combine this result with some linear programming algorithm to obtain a maximum matching algorithm. This is not straightforward because of the large number of inequalities involved. However, a recently discovered new method for solving linear programs, the so-called *Ellipsoid Method* (Šor (1970, 1977), Judin and Nemirovskiĭ(1976), Hačijan (1979)) can be used to turn a polyhedral description of the convex hull of matchings into a polynomial-time algorithm for maximum matching. (For more details see Grötschel, Lovász and Schrijver (1981).)

Besides the existence and optimization problems mentioned above, there are other important aspects of matching theory. The study of the Ising model for ferromagnetic materials led Kasteleyn (1961, 1963) to the problem of *enumerating perfect matchings* in graphs and he was able to solve this problem for planar graphs. Later, L. Valiant (1979a) proved that the enumeration problem for perfect matchings is NP-hard in general, but useful upper and lower bounds have been obtained. Recent work of Heilmann and Lieb (1970, 1972) and Godsil (1981b) on the generating function for the number of matchings of various sizes relates this problem to determinants and thereby recalls the work of Frobenius dating back to the beginning of this century.

Our historical sketch now brings us to the present book. The first chapter deals with bipartite matching, as was the case historically. In particular, we present first the König Minimax Theorem and then the Marriage Theorem of Frobenius and the result on distinct representatives due to P. Hall. The bipartite matching algorithm know as the Hungarian Method is presented next. For treating bigraphs not having perfect matchings, the concepts of deficiency and surplus are then introduced. The fact that they are supermodular and submodular functions respectively, is discussed and this leads us to consider such functions in the more general framework of matroids. We close the chapter with some of the many consequences of the König–Hall–Frobenius results; König's Line Coloring Theorem and Dilworth's Theorem being only two examples.

In Chapter 2 we develop enough from the theory of network flows to show that most bipartite matching results can be expressed and solved within this framework. In addition the idea of a flow-equivalent tree which is needed later is introduced in this chapter.

Chapter 3 contains fundamental results for the non-bipartite case such as Tutte's Theorem on perfect matchings and Berge's "defect" version thereof. Next we develop the Gallai-Edmonds Structure Theory. The Gallai-Edmonds theory helps us reduce the study of the structure of maximum matchings to three disjoint classes of more special graphs: factor-critical graphs, positive surplus bipartite graphs and graphs having perfect matchings. The last of these three families is further reduced to the study of those which are "elementary". A graph with a perfect matching is **elementary** if the union of all its perfect matchings forms a connected subgraph.

In Chapter 4 we study elementary bipartite graphs and in Chapter 5 we undertake the investigation of elementary graphs in general. We produce a further decomposition of the latter class into smaller elementary *bipartite* graphs and into a new type of elementary graph called "bicritical". This brings us to a frontier, so to speak, as far as decompositions which are "canonical" are concerned. Factor-critical and bicritical graphs are taken up next and quite a lot of useful information about both families is gained from the so-called "ear decomposition" results presented. But such ear decompositions, although quite useful, are not *canonical*. Canonical theories for the decomposition of factor-critical and bicritical graphs do not yet exist.

In Chapter 6 we generalize the idea of matchings to subgraphs having all degrees at most two — the so-called **2-matchings**. We see that the corresponding generalization of König's Minimax Theorem to 2-matchings holds for *all* graphs and not just for those which are bipartite! The 2-matching analogues of elementary and bicritical graphs are introduced.

In many ways these 2-matchings are easier to handle than ordinary (1-) matchings; so it is quite reasonable to ask, for example, how maximum 2-matchings might be used to obtain maximum 1-matchings. The answer is not yet known. In another direction of generalization, we show how 2-matchings can be used to give good characterizations of non-bipartite graphs which satisfy the König minimax equation.

Next we discuss the Chinese Postman Problem from the point of view of 2-matchings and finally, we close out the chapter with a collection of other problems which are reducible to matching problems of one kind or another.

König's Minimax Theorem is a special integer-valued instance of the more general Duality Theorem of linear programming. In Chapter 7 we formulate more general matching problems as linear programs. We present Edmonds' approach to determining the facets of the associated convex polytope spanned by the binary incidence vectors of *all* matchings in a graph. We shall study this so-called *matching polytope* $M(G)$ as well as an assortment of related polyhedra: the *fractional matching polytope*, $FM(G)$, the *vertex packing polyhedron* $VP(G)$ and equivalently, the *point cover polyhedron* $PC(G)$, the *fractional point cover polyhedron* $FPC(G)$ and, finally, the *perfect matching polytope* $PM(G)$. Our knowledge of these varies from considerable for the matching polytope $M(G)$ to very little in the case of $VP(G)$. We shall return to $VP(G)$ in Chapter 12.

König's initial investigations of perfect matchings were motivated by related questions for determinants and in Chapter 8 we return to this relationship and extensions thereof. Because of algebraic sign problems with the expansion terms, the determinant cannot be used to enumerate perfect matchings in a bigraph in general. This difficulty can be overcome, in a sense, by switching to the *permanent* function, but then new problems replace the old. The permanent is notoriously hard to handle! Nevertheless, we study bounds of various types for the permanent function and use them to obtain bounds for the number of perfect matchings in a regular bigraph.

We extend our considerations to the non-bipartite case by introducing a third matrix function–the *Pfaffian*. Some probabilistic considerations are then discussed before moving on to a discussion of the *matching polynomial*. This polynomial is related to the better-known "characteristic polynomial" of a graph, and in fact for some graphs — trees for example — the two coincide. The matching polynomial is then applied to two topics from theoretical chemistry and physics, so-called *topological resonance energy*, as well as to the *Ising model* for magnetic materials.

Whereas the bounds on the permanent developed earlier apply only to graphs which are regular, we close this chapter by applying results from earlier chapters to obtain new lower bounds for graphs which are not necessarily regular.

In Chapter 9 we present and analyze the *Edmonds Matching Algorithm.* Three other algorithmic approaches are also discussed. Two of these turn out to be polynomial, namely a routine based upon the Gallai-Edmonds Decomposition and a second using the very recently developed Ellipsoid Method for linear programming. The third routine, due to Padberg and Rao (1982) is not polynomial in the worst case, but it seems to be competitive with the algorithm of Edmonds in practice.

In Chapter 10 we generalize the degree one restriction of perfect matchings by investigating spanning subgraphs having a prescribed degree $f(v)$ at each point v. We call these *f-factors.* We begin by showing that in fact one can reduce the "f-factor problem" to the "1-factor problem", that is, to the perfect matching problem. Then we proceed to obtain results for f-matchings analogous to the Gallai-Edmonds decomposition results for 1-matchings.

A classical question in graph theory asks: when is a given sequence of non-negative integers realizable as the sequence of degrees of a graph? This may be viewed as a special case of the f-factor problem where the graph in question is complete. We address questions of this type as we bring Chapter 10 to a close.

In Chapter 11 we discuss some extensions of (non-bipartite) matching theory to the more general setting of matroids. (The reader will have seen by the time he reaches Chapter 11 that we have been only partially successful at keeping matroid theory out of the first ten chapters!) We formulate a number of matroid problems which turn out to be equivalent and which are collectively called the *Matroid Matching Problem.* Examples drawn from the disciplines of architecture and electrical engineering show that this extension of matching theory has important applications.

In our final chapter we take up the study of *vertex packing*, that is, the study of independent sets of *points.* Since matchings in any graph G correspond to independent sets of points in the associated line graph $L(G)$, one can take the point of view that all matching problems are just problems about vertex packing in the special subclass of all graphs — the line graphs. Unfortunately this does not help much! Our hopes are further dashed upon learning that the vertex packing problem is NP-complete, whereas matching, as we know, is polynomial.

But progress has been made on vertex packing and a number of interesting results have been obtained, mainly in the study of so-called *τ-critical graphs.* We present most of what is currently known about these interesting, but difficult, graphs including a finite basis theorem.

We next revisit the vertex packing polytope. Facet determination, as was done for the matching polytope in Chapter 7, seems hopeless here, but for bipartite graphs and for the more general class of "perfect" graphs this polytope has a nice description which we present.

It turns out that perfect graphs are also closely related to another natural generalization of matching, namely *hypergraph matching.* A hypergraph is a generalization of the idea of a graph in which a line may have more than two endpoints. The matching problem can be generalized to hypergraphs in a natural way: find the maximum number of disjoint "lines" in a hypergraph. While this problem is NP-complete, various generalizations of König's Minimax Theorem to hypergraphs do exist. We discuss briefly this problem of hypergraph matching.

One interesting class of graphs which contains the line graphs and for which vertex packing is polynomially solvable is the class of *claw-free* graphs. We conclude the chapter — and the book — by presenting an algorithm for this problem.

Now we must say a few words about the "Boxes" inserted throughout the text. No book on any mathematical discipline can be truly "linear" in its development. Branching in various directions is always possible, frequently desirable and sometimes inevitable. So the pruning shears must be ruthlessly employed! (Indeed, the reader may have preferred a greater degree of ruthlessness!)

Matching theory has not developed in a vacuum. Indeed it has often been in attendance when many of the exciting new concepts in combinatorial optimization have been born. It provided an archetypal minimax theorem which in turn places it close to the birth of duality theory in linear programming. The matching polytope was the first non-trivial polyhedron to be studied by Edmonds as he broke new ground in the areas of facet determination and "good" characterizations. His non-bipartite matching algorithm is a landmark which showed that matching, although certainly not an easy problem, was solvable in polynomial time.

We cannot pursue all these related areas in detail in this book of course. But we have decided to insert boxes of material at various points to provide more background information for the reader. Usually in these boxes we present, in a condensed manner, background material which may be useful to some readers, but which may be well-known to others. For the reader with more background in these topics, we have tried to insert these boxes in such a way that they may be skipped without unduly disrupting the flow of our presentation. At the other extreme, the reader with little — or no — background in these areas will frequently want to read further and we hope that these boxes, together with the references cited, will help in these efforts.

We have also included a brief section on basic terminology, an index of terms and an index of symbols to help the reader translate this book into English! It is likely that no two graph theorists agree on terminology (the two authors certainly do not!), but we hope that you, gentle reader, can learn to live with ours. As for us, we are rather proud to say that our friendship has survived this acid test of terminology selection!

A "pruned" bibliography follows the text. The term "pruned" is used here because in the final analysis we have decided for the sake of space to include only those references on matching which we cite in the text. Many other relevant papers exist, however, and a much more extensive bibliography on matching will be published separately by the authors.

A few remarks are now in order about the format of the book. Most definitions appear in the body of the text and will be set in **bold face** type when they occur for the first time. Sometimes, however, these bold face expressions will appear more than once; for example, when a concept must be recalled from some distant point earlier in the book. All these definitions and others can be found in our index of terms.

Our numbering scheme will be as follows. Theorems, lemmas, corollaries and exercises are all in the same basket as far as numbering is concerned. A string like "$x.y.z.$ Theorem" will refer to theorem z in Section y of Chapter x. Equations and inequalities and a few other displayed strings will be numbered separately, but similarly. Such strings will have three digit number sequences too, but will always be parenthesized. For example, "$(x.y.z)$" will refer to displayed equation (or inequality, etc.) number z in Section y of Chapter x. For further clarity, we endeavor to always say "equation $(x.y.z)$" when referring to this equation in the body of the text and similarly for inequalities, etc.

Finally, figures and tables will be numbered separately from lemmas, theorems, corollaries, exercises, equations, etc. and from each other. The string "Figure $x.y.z$" refers to the zth figure in Section y of Chapter x. Our symbol for the end of a proof (either presented or omitted) of a theorem, lemma or corollary will be the now common mathematical insect called the "black slug" and denoted by "■".

The idea of writing this book probably dates back to some time during 1975-76 when the second author visited Budapest to continue joint research begun with the first author in Nashville in 1972-73. Actually, neither of us quite remembers — or is willing to admit — when such a fatuous idea first occurred! Since that time we have pursued our task — and sometimes each other — in various parts of the world. Hence this peripatetic pair of authors has many people in many places to thank.

This book was done with the TEX editing system on the DEC 1099 computer system at Vanderbilt University. The TEX aspect of this project would never have gotten off the ground without the deep knowledge of "TEXpert" Brendan McKay who gave many hours of his time unselfishly in helping the second author to become acquainted with the system. In addition, McKay designed most of the macros pertinent to the book. Joan McKay typed most of the manuscript into the computer using the TEX system and did so promptly and in a remarkably error-free fashion. We also want to thank Flo Worden and Ruby Moore for some supplemental typing. Maria Perkins and David Palmer of the Vanderbilt Computer Center were most helpful with TEX and other computer related problems. Kathy Goforth has our gratitude for writing the original program to handle the bibliography.

We are deeply appreciative for valuable manuscript reading and commentary by Steven Campbell, András Frank, Chris Godsil, Ivan Gutman, Péter Pál Pálfy, Bill Pulleyblank, András Recski, Larry Schaad, András Sebő and Lex Schrijver. Special thanks are due to Derek Holton and his Honours Class in Mathematics at the University of Melbourne who endured a series of lectures in 1979 by the second author which were drawn from a preliminary draft of the first part of the book. This same author wishes to thank the MÁV (Hungarian State Railways) for keeping the Szeged Expressz running on time!

Our warmest gratitude is extended to Pál Erdős and Tibor Gallai who were both "present at the creation" of matching theory and whose personal recollections were unselfishly shared with both authors on a number of occasions.

For financial assistance and hospitality of various kinds we want to thank IREX (International Research Exchange), NKI (the Hungarian Cultural Institute), NAS (the National Academy of Sciences), MTA (the Hungarian Academy of Sciences), SFB (the Sonderforschungsbereich of

West Germany), MKI (the Mathematics Institute of the Hungarian Academy of Sciences), Vanderbilt University, Attila Jozsef University, Loránd Eötvös University, The University of Bonn, the University of Waterloo and the University of Melbourne.

We are grateful too for special funds made available by the Vanderbilt University Research Council and the Vanderbilt Natural Science Fund.

Our deepest thanks must be reserved for our two families who have cheerfully endured the peregrinations forced upon them so that the two authors could work together as well as the absences of the authors for similar reasons on many occasions. In the final analysis, it is the close friendship that all nine of us enjoy that counts for more than any book possibly could.

This Page Intentionally Left Blank

Basic Terminology

We present here a concise collection of those basic definitions in graph theory we need to get started. Additional terminology will be presented later in the book as needed.

An **undirected graph** (or simply a **graph**) G consists of a finite non-empty set of elements $V(G)$ called **points** and a multi-set of unordered pairs of points $E(G)$ called **lines**. Please note that we are allowing "multiple" or "parallel" lines here, unless otherwise specified. When multiple lines are *not* allowed, we shall call the corresponding graph **simple**. Also, we will not allow **loops** , i.e., lines of the form uu, unless otherwise specified. Unless stated otherwise, p will denote $|V(G)|$ and q will denote $|E(G)|$.

If uv is a line in graph G, line uv is said to **join** points u and v, to be **incident** with points u and v, and points u and v are said to be **adjacent**. Two lines which share a point are also said to be **adjacent**. The set of lines with exactly one endpoint incident with a point in X will be written $\nabla(X)$ and the set of lines with one endpoint in X and the other in Y will be written $\nabla(X,Y)$. The **complement** of a graph G, denoted $\overline{G}$, is that graph having the same point set as G, but in which two such points are adjacent if and only if they are not adjacent in G.

A one-to-one function f mapping $V(G)$ onto $V(H)$ is called an **isomorphism** if f and f^{-1} preserve the number of lines joining each pair of points. If such an isomorphism exists, graphs G and H are said to be **isomorphic**. An isomorphism of graph G onto itself is called an **automorphism** of G. The set of all automorphisms of a graph G under the operation of composition constitute a group called the **automorphism group** of G and denoted by $\mathrm{Aut}(G)$.

The number of lines in graph G incident with a point u is called the **degree** of u (in G) and denoted by $\deg_G(u)$. If graph G is understood, we shall sometimes abbreviate this to $\deg(u)$. A graph in which all degrees are equal to k is said to be **k-regular** and if G is k-regular for some k, we simply say that G is **regular**. A graph which is 3-regular is often called **cubic**.

An alternating sequence of points and lines, beginning and ending with points, is called a **walk**. If all lines in a walk are distinct, the walk is called a **trail**, and if, in addition, the points are also distinct, the trail is a **path**. A family of paths which have no points in common, except possibly their endpoints, will be called **openly disjoint**. The **length of a**

walk is the number of occurrences of lines in it. The **distance** between points u and v, written $d(u,v)$, is the length of any shortest path joining them. If P is a path and u and v are any two points on P, then $P[u,v]$ denotes the subpath of P having endpoints u and v.

A walk or trail in which the first and last points are the same will be said to be **closed**. If a graph contains a closed trail which includes all the lines of G the trail is called an **Euler trail** of G and a graph containing an Eulerian trail is said to be **Eulerian**.

We shall stick by tradition in avoiding the term "closed path", and instead we shall define a **cycle** to be any path of length at least two, together with a line joining the first and last points. The **length** of a cycle will also be the number of lines it contains. A cycle of length n will be called an **n-cycle**. A line joining two points of a cycle, but not itself a line of the cycle, is a **chord** of the cycle. A cycle which includes every point of a graph G is called a **Hamilton** cycle of G. The length of any shortest cycle in a graph G is called the **girth** of G and denoted by $\text{girth}(G)$.

If G is a graph and H is also a graph the points and lines of which are points and lines of G, then H will be called a **subgraph** of G. If H is a subgraph of G and if every line joining two points of H which lies in G also lies in H, we call H an **induced** subgraph of G. If X is a set of points in graph G, then $G[X]$, the subgraph of G **induced by** X, is the induced subgraph of G having point set X. A subgraph H of G is said to be **spanning**, if $V(H) = V(G)$. A spanning subgraph regular of degree n is called an **n-factor**.

A graph in which every pair of points are adjacent is said to be **complete**, and the complete graph on n points is denoted by K_n. A maximal complete subgraph of graph G is called a **clique** of G. A subgraph H is said to be **excluded** with respect to a property *PROP*, if no graph with property *PROP* has H as a subgraph.

A set of points or lines S in a graph is said to be **minimal** with respect to property *PROP*, if the set has property *PROP*, but no proper subset of S has property *PROP*. Set S is said to be **minimum** with respect to property *PROP* if, among all subsets of G having property *PROP*, S is one having smallest cardinality The terms **maximal** and **maximum** are defined analogously.

A graph is **connected** if every two points are joined by a path. A maximal connected subgraph of G is called a **component** of G. Components are **even** or **odd** according to whether their point sets have even or odd cardinality.

If the point set of a graph G can be partitioned into two disjoint non-empty sets, $V(G) = A \cup B$, such that all lines of G join a point of A to a point of B, we call G **bipartite** and refer to $A \cup B$ as the **bipartition** of G. In this case we shall also sometimes call the sets A and B the **color classes** of G. A bipartite graph is often also referred to as a **2-colorable** graph or **bigraph**. A special bipartite graph which we shall have occasion to use is $K_{m,n}$, the **complete bipartite graph** having color classes of size m and n and in which every point in each color class is adjacent with every point in the other. In particular, $K_{1,n}$ is called an n-**star** (or sometimes simply, a **star**). A graph containing no cycles is called **acyclic**. An acyclic graph is called a **forest** and if the acyclic graph is also connected, it is called a **tree**. If tree T is a subgraph of graph G and if $V(T) = V(G)$, we call T a **spanning tree** of G.

A set of points S in a connected graph G is a **cutset** if $G - S$ is not connected. A similar definition obtains for a set of lines. If S is a cutset of G consisting of a single point v, the point v is called a **cutpoint** of G, and if S contains a single line e, line e is a **cutline**, or **bridge**, of G. A connected graph containing no cutpoint is called a **non-separable** or **2-connected** graph, or simply a **block**.

If G is not a complete graph, the cardinality of a minimum cutset of points in graph G is called the **(point)-connectivity** of G and is denoted by $\kappa(G)$. If $G = K_n$, κ is defined to be $n-1$. Similarly, the size of a minimum cutset of lines in G is the **line connectivity** of G and is written $\lambda(G)$. A graph G is said to be k-**connected** if $k \leq \kappa(G)$ and to be k-**line-connected** if $k \leq \lambda(G)$. A maximal n-connected subgraph will be called an n-**connected component**, or simply an n-**component**.

A **point coloration** of graph G is an assignment of positive integers to the points of G so that no two points labelled with the same integer are adjacent. G is said to be n-**colorable** if G has a point coloration in n colors. The smallest integer k for which graph G has a coloration of its points in k colors is called the **chromatic number** of G and is denoted by $\chi(G)$. If we assign positive integers to the lines of G so that no two lines with the same integer label are adjacent, we have a **line coloration** of G. The smallest value of k for which G has a line coloration in k colors is called the **chromatic index** of G and is written $\chi_e(G)$.

The **genus** of a graph G, $\gamma(G)$, is the smallest genus of an orientable surface in which G may be embedded so that no two lines meet, except perhaps at their endpoints. Graphs of genus 0 are said to be **planar**.

If the lines of a graph have a direction assigned to them, we have what is known as a "directed graph". More precisely, a **directed graph**, or **digraph**, D consists of a set of **points** $V(D)$ and a set of *ordered* pairs of points $E(D)$ called **lines**. The number of lines having v as their second point is called the **indegree** of v and is denoted by $\deg^-(v)$. Similarly, the **outdegree** of point v is the number of lines having v has their first point and is written $\deg^+(v)$. The definitions of walk, trail, path and cycle must be modified somewhat in the case of directed graphs. In each of these alternating sequences of points and lines, we shall insist that each (directed) line join the point before it to the point after it *in the sequence*. An **acyclic** digraph is one containing no (directed) cycles. A digraph is **strongly connected** if given every ordered pair of points (u, v), there is a (directed) path from u to v.

A set of lines in a graph G is called **independent** or a **matching** if no two lines have a point in common. The size of any largest matching in G is called the **matching number** of G and is denoted by $\nu(G)$. Now suppose M is a fixed matching in graph G. A point v is said to be **covered**, **matched** or **saturated** by M if some line of M is incident with v. Unmatched points are also called **unsaturated**, **uncovered** or **exposed**. A path (or cycle) P is said to be M**-alternating** if the lines of P are alternately in and not in M. Note that an M-alternating path may begin with a line in M or with a line not in M. If, however, an M-alternating path P begins and ends with lines not in matching M, we call P an M**-augmenting** path. If the matching M is understood, we may simply refer to a path as being **alternating** or **augmenting**. A matching is **perfect** if it covers all of $V(G)$. A graph with a perfect matching is sometimes called a **factorizable** graph.

A **line cover** in a graph G is a set of lines collectively incident with each point of G. The cardinality of any smallest line cover in a graph G is called the **line covering number** of G and is denoted by $\rho(G)$.

A set of points in a graph G is said to be **independent** if no two of them are adjacent. The cardinality of any largest independent set of points in G is known variously as the **(point) independence number** of G, the **stability number** of G and the **vertex packing number** of G, and is written $\alpha(G)$. A set of points S of G is a **point cover** of G if each line of G has at least one endpoint in set S. The cardinality of any smallest point cover is denoted by $\tau(G)$ and is known as the **point covering number** of G.

Given a graph G, **the line graph** of G, $L(G)$, is constructed as follows. The point set of $L(G)$ is $E(G)$ and two points of $L(G)$ are adjacent if and only if they are adjacent as lines in G.

The operation of inserting a new point of degree two on a line of a graph is called **subdividing** the line. If the number of new points of degree two inserted is even, the subdivision is said to be **even**; otherwise it is **odd**.

In this book, the positive integers, the integers, the rationals and the real numbers will be denoted by Z_+, Z, Q and $\Re$, respectively. The finite field containing two elements will be denoted by $GF(2)$. The greatest integer not greater than real number x will be written as $\lfloor x \rfloor$ and the least integer not less than x as $\lceil x \rceil$. If k is a positive integer, the symbol $k!!$ denotes the product $k(k-2)(k-4)\cdots 4 \cdot 2$ if k is even, and $k(k-2)(k-4)\cdots 3 \cdot 1$ if k is odd.

There are a number of good books on graph theory currently available with terminology mostly consistent with that adopted here. In addition, they offer more extended discussion of these basic concepts as well as a plethora of examples. To mention but two of these books, we direct the reader to the volumes by Bondy and Murty (1976) and Bollobás (1978b).

This Page Intentionally Left Blank

1

Matchings in Bipartite Graphs

1.0. Introduction

A graph is **bipartite** if its set of points $V(G)$ can be partitioned into two sets A and B such that every line in $E(G)$ has one endpoint in A and the other in B. We shall write $G = (A, B)$ accordingly, but we shall often speak about bipartite graphs G and assume that sets A and B are understood. The sets A and B are often called the **color classes** of G and (A, B) a **bipartition** of G.

The reader may see immediately by scanning ahead in the book that many results for bipartite graphs which appear here in Chapter 1 have been extended to more general families of graphs. We wish to avoid scattering most of the results for bipartite graphs elsewhere and thus this first chapter deals with bipartite graphs in particular. Historically, to be sure, some of the most important theorems for bipartite graphs were proved directly and only later were these bipartite results obtained as corollaries of more general non-bipartite theorems. Moreover, we feel bipartite graphs deserve special treatment because it is still the case that the majority of real world applications of matching theory deal with graph models which are bipartite.

First, however, we present a few results which are true for all graphs, bipartite or not, but which are so basic to any study of matchings that we think the reader should have them in hand here at the beginning of the book.

Let G be any graph, bipartite or not, and let $\nu(G)$ denote the matching number of G, $\tau(G)$, the point covering number, $\alpha(G)$, the (point) independence number and $\rho(G)$, the line covering number. There are several results linking these numbers and we now present two of the most basic of these which we shall refer to as the **Gallai Identities**. (See Gallai (1959). In an historically interesting footnote, Gallai says that he proved these results back in 1932 and moreover he believes that König was already aware of them at that time.) We will prove only the second of these leaving the (easier) first result to the reader.

1.0.1. LEMMA. *For any graph G, $\alpha(G) + \tau(G) = |V(G)|$.* ■

1.0.2. LEMMA. *For any graph G having no isolated points, $\nu(G) + \rho(G) = |V(G)|$.*

PROOF. Let C be a line cover of G containing $\rho(G)$ lines. Let $\langle C \rangle$ denote the subgraph of G with line set C and point set $V(C)$. Then by the minimality of C, subgraph $\langle C \rangle$ consists of a union of point-disjoint stars. If n is the number of components in $\langle C \rangle$ we see that $n = p - \rho(G)$. On the other hand, each star in $\langle C \rangle$ contains at least one line, so taking one line from each star we get a matching M and hence $\nu(G) \geq n = p - \rho(G)$ or

$$\nu(G) + \rho(G) \geq p. \tag{1.0.1}$$

Now let M_0 be any maximum matching in G and U, the set of points not covered by M_0. Thus U is an independent set. Now G has no isolates, so for each of the $p - 2\nu(G)$ points of U, select a line covering it and call this collection of lines S. Then $M \cup S$ is a line cover for G and we have $\rho(G) \leq |M \cup S| = \nu(G) + p - 2\nu(G) = p - \nu(G)$. That is,

$$\nu(G) + \rho(G) \leq p. \tag{1.0.2}$$

Upon combining inequalities (1.0.1) and (1.0.2) the lemma is proved. ■

In proving Lemma 1.0.2 the reader undoubtedly discovered for himself a nice relationship between minimum point covers and independent sets of points, namely that if S is a minimum point cover then $V(G) - S$ is a maximum independent set of points. This complementarity relation clearly fails for line covers and matchings. Hence at this point the reader may feel that questions involving $\alpha(G)$ and $\tau(G)$ will be easier to handle somehow than problems involving $\nu(G)$ and $\rho(G)$. It is perhaps surprising, then, to learn that in most cases the opposite is true! In Chapters 1 and 9 the reader will find efficient matching algorithms whereas in Box 6A we will see that the determination of $\alpha(G)$ is one of the most difficult combinatorial problems.

Let us now present two results due to Norman and Rabin (1959) and independently to Gallai (1959) which together with Lemma 1.0.2 may help the reader's insight with respect to relationships between matchings and line covers.

1.0.3. EXERCISE. (a) A minimal line cover is minimum if and only if it contains a maximum matching.

(b) A maximal matching is maximum if and only if it is contained in a minimum line cover.

The fundamental inequality in the next exercise will repeatedly prove to be useful.

1.0.4. EXERCISE. For any graph G, $\nu(G) \leq \tau(G)$.

Now let us return to the study of bipartite graphs in particular. First a few words are in order about our method of attack. There are a number of results which are equivalent to the next result we present, the so-called "König Minimax Theorem". We will present some of these equivalent results, but we do not want to bog the reader down with a tedious and lengthy "circle of implications" showing all these to be equivalent. Working through such equivalences is helpful to the reader's understanding, of course, but usually does not contribute to the "flow" of the book. Therefore many of these tasks will be delegated to the list of Exercises for the reader. For a thorough treatment of these equivalences, we refer the reader to (collectively) Jacobs (1969), Ford and Fulkerson (1962), Hoffman (1960) and Robacker (1955).

We will emphasize strongly not only the statements of fundamental results (like König's Minimax Theorem and P. Hall's Theorem on distinct representatives), but also certain proof techniques which will be invaluable throughout the course of this book and beyond. In our treatment of the theorems of König and P. Hall, for example, the reader will find different proofs presented (or mentioned) which use the method of "critical graphs", induction on the size of the graph, alternating path techniques and (cf. Chapter 2) the theory of network flows. Such proof techniques will prove useful again and again in various areas of matching theory in general (and certainly not only in the special case of bipartite graphs!) and therefore all deserve charter membership in the investigator's arsenal. Several other techniques for proving these theorems — for example, linear programming duality and determinant theory — will be treated later in the book (cf. Chapters 7 and 8).

Let us add one more remark about names associated with these theorems. A number of early equivalent results were proved by several people. Historically, the first of these was Frobenius, but the names of König, Egerváry and P. Hall are more often assigned to various generalizations of his result. We shall adopt the point of view by which we shall name various versions after the person who to the best of our knowledge formulated them first.

1.1. The Theorems of König, P. Hall and Frobenius

In the special case of bipartite graphs another easily stated identity holds in addition to the basic identities of Gallai mentioned in the introduction to this chapter. Its simple statement, however, belies the fact that it, together with the equivalent versions of P. Hall and Frobenius, is probably the single most important result to date in all of matching theory. The "extremal graph" proof given here first may be found in Lovász (1975b).

1.1.1. THEOREM. (*König's Minimax Theorem*). *If G is bipartite, then* $\tau(G) = \nu(G)$.

PROOF. By Exercise 1.0.4 we have

$$\tau(G) \geq \nu(G). \tag{1.1.1}$$

We proceed to obtain the reverse inequality. To this end, remove lines from G as long as possible while keeping τ the same. Denote the resulting minimal graph by G'. Hence $\tau(G') = \tau(G)$, but $\tau(G'-e) < \tau(G)$ for every line e in G'. We claim that G' consists of independent lines.

Suppose not. Then there are two lines x and y adjacent to a point a in G'. Consider $G' - x$. By the minimality of G' there is a point cover S_x covering $G' - x$ with $|S_x| = \tau(G') - 1$. Of course neither endpoint of x lies in S_x. Similarly, there is a set S_y covering $G' - y$ containing neither endpoint of y and $|S_y| = |S_x|$.

Form the induced subgraph G'' of G', $G'' = G'\big[\{a\} \cup (S_x \oplus S_y)\big]$, where $\oplus$ denotes the symmetric difference operation. Let $t = |S_x \cap S_y|$. Then $|V(G'')| = 2\big(\tau(G') - 1 - t\big) + 1$ and since G'' is bipartite (being a subgraph of G), there is a set T (the smaller of the two color classes of G'') which covers G'' and $|T| \leq \tau(G') - 1 - t$.

But now $T' = T \cup (S_x \cap S_y)$ covers G'. For suppose z is any line in G'. If $z \neq x$ or y, then z is covered by both S_x and S_y; that is, either it is covered by $S_x \cap S_y$ or it connects $S_x - S_y$ to $S_y - S_x$. In this second case it is a line of G'' and hence covered by T. Finally, x and y are lines of G'' and are therefore covered by T.

So $\tau(G') \leq |T'| = |T \cup (S_x \cap S_y)| = |T| + |S_x \cap S_y| \leq \tau(G') - 1 - t + t = \tau(G') - 1$, a contradiction. Thus G' consists of independent lines as claimed. Hence

$$\tau(G) = \tau(G') = \nu(G') \leq \nu(G) \tag{1.1.2}$$

and combining this with inequality (1.1.1), we are done. ■

We want to emphasize before proceeding that Theorem 1.1.1 is an important example of a class of results known as "minimax theorems". The well-known Max-Flow Min-Cut Theorem of Section 2.1 is another important example in this class.

Clearly inequality (1.1.1) holds in any graph, bipartite or not. It is natural to wonder precisely when equality holds in general. For example, $\nu = \tau = 2$ in the non-bipartite graph resulting from K_4 upon deleting a line, but $\tau = 2 > 1 = \nu$ in K_3. We shall defer further discussion of this and related questions to Sections 6.3 and 12.3.

Although tangential to our plan of treating only finite graphs, it is worth-while to mention that Aharoni (1984) has recently proved the following extension of König's Minimax Theorem to infinite graphs first conjectured by Erdős.

1.1.2. THEOREM. *In any (possibly infinite) bipartite graph there exists a matching M and a point cover P such that every line in M contains exactly one point in P and every point in P is contained in exactly one line of M.* ■

This is, of course, immediate for finite graphs by Theorem 1.1.1.

Probably the most widely used version of all those statements equivalent to König's Minimax Theorem is due to Philip Hall (1935). If X is any set in $V(G)$, let $\Gamma(X)$ denote all points in $V(G)$ which are adjacent to at least one point of X.

1.1.3. THEOREM. (*P. Hall's Theorem*). *Let $G = (A, B)$ be a bipartite graph. Then G has a matching of A into B if and only if $|\Gamma(X)| \geq |X|$ for all $X \subseteq A$.*

Since trivially, G has a matching of A into B if and only if $\nu(G) = |A|$, this theorem could be obtained easily as a corollary of König's Theorem 1.1.1. Here we present an independent proof due to Halmos and Vaughn (1950).

A plethora of proofs of Hall's result have appeared over the years, starting with one by Maak published in the same year as that of Hall. For good historical summaries we refer the reader to Mirsky (1971, pg. 38) and to Jacobs (1969).

PROOF. First note that if there is a matching of A into B, the desired inequality above holds for any subset X of A.

Suppose $|\Gamma(X)| \geq |X|$ for all $X \subseteq A$. We proceed by induction on $|A|$. If $|A| = 0$ or 1, the result is immediate.

Next suppose for all $X \subset A$, $X \neq \emptyset$, $|X| < |\Gamma(X)|$ holds. Let a and b be adjacent points with $a \in A$. Let $G' = G - a - b$ and let X be any subset of $A-a$. If $X=\emptyset$, then $|X|=0=|\Gamma_{G'}(X)|$ so suppose $X\neq\emptyset$. Since $X\neq A$, $|X|<|\Gamma_{G'}(X)|$ by assumption and thus $|\Gamma_{G'}(X)|\geq|\Gamma_G(X)|-1\geq|X|$. Hence by the induction hypothesis, there is a matching M' of G' which covers all points of $A-a$. But then $M=M'\cup\{ab\}$ matches A and B as desired.

Now suppose, on the other hand, there is a set $A'\subset A$, $A'\neq\emptyset$ with $|\Gamma_G(A')|=|A'|$. We proceed to split G into two smaller subgraphs by letting G_1 be the subgraph induced by $A'\cup\Gamma(A')$ and $G_2=G-A'-\Gamma(A')$ and show that G_1 and G_2 satisfy the induction hypothesis separately. Suppose $X\subseteq A'$. Then $\Gamma_G(X)\subseteq\Gamma_G(A')$, so $\Gamma_{G_1}(X)=\Gamma_G(X)$ and hence $|\Gamma_{G_1}(X)|=|\Gamma_G(X)|\geq|X|$. Now in G_2 assume $X\subseteq A-A'$. Then $\Gamma_G(X\cup A')=\Gamma_{G_2}(X)\cup\Gamma_G(A')$ and therefore $|\Gamma_{G_2}(X)|=|\Gamma_G(X\cup A')|-|\Gamma_G(A')|\geq|X\cup A'|-|\Gamma_G(A')|=|X\cup A'|-|A'|=|X|$, since $X\cap A'=\emptyset$ and $\Gamma_{G_2}(X)\cup\Gamma(A')=\emptyset$.

By applying the induction hypothesis to both G_1 and G_2, we see that there must exist matchings M_1 of A' into (and therefore *onto*) $\Gamma_G(A')$ and M_2 of $A-A'$ into $B-\Gamma_G(A')$. The union $M=M_1\cup M_2$ is then the desired matching. ■

A **perfect matching** (or **1-factor**) is a matching which covers all points of G. The following result of Frobenius (1917), often called "the Marriage Theorem", characterizes bigraphs with a perfect matching.

1.1.4. COROLLARY. (*The Marriage Theorem*). *A bipartite graph $G=(A,B)$ has a perfect matching if and only if $|A|=|B|$ and for each $X\subseteq A$, $|X|\leq|\Gamma(X)|$.* ■

So we have seen that the theorem of Frobenius is a special case of that of P. Hall, which in turn may be viewed as a special case of König's Theorem. On the other hand, it is not difficult to derive König's Theorem from that of Frobenius (Exercise 1.1.5). For this reason, the Marriage Theorem is often said to be a *self-refining* result.

1.1.5. EXERCISE. Deduce König's Theorem from Frobenius' Theorem.

1.1.6. EXERCISE. Either find a perfect matching or a set X violating the condition in the Marriage Theorem for the graph shown in Figure 1.1.1.

We conclude this section with one more equivalent form of the bipartite matching theorems, also due to König (cf. Gallai (1958)). Recall that a **line cover** is a set of lines which collectively cover each point of G and that $\rho(G)$ is the minimum number of lines in any line cover. Also as before $\alpha(G)$ denotes the (point) independence number. Then the Gallai Identities (Lemmas 1.0.1 and 1.0.2), together with König's Theorem, immediately imply the following result. (Cast in the language of Section

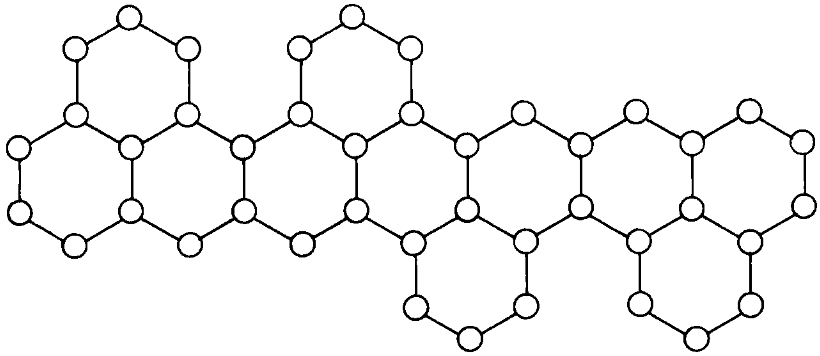


FIGURE 1.1.1.

12.2, this Corollary proves that the complement of a bipartite graph is perfect.)

1.1.7. COROLLARY. *If G is bipartite, $\rho(G) = \alpha(G)$.* ■

1.1.8. EXERCISE. Let $S = \{1, 2, \ldots, n\}$ and suppose $0 \leq k < k/2$. Let A (B) be the collection of all k-element $(k+1$-element) subsets of S. Construct a bipartite graph G on $V(G) = A \cup B$ by joining $X \in A$ to $y \in B$ if and only if $X \subseteq Y$. (a) Prove that G has a matching of A into B. (b) For each $X \in A$, define

$$\Lambda(X) = 1 + \max\{2t \mid |X \cap \{1, \ldots, 2t\}| \geq t\}.$$

Show that $X \to \Lambda(X)$ defines a matching of (all of) A into B in graph G. (c) Use part (a) to prove Sperner's Lemma (1928) which says that the maximum number of subsets of an n-element set such that none is contained in any other is $\binom{n}{\lfloor n/2 \rfloor}$.

BOX 1A. NP-properties, Good Characterizations and Minimax Theorems

Frobenius' theorem characterizes those bipartite graphs which have a perfect matching. Hall's Theorem characterizes those bipartite graphs which have a matching of A into B. König's Theorem gives a formula for the matching number of a bigraph. What is the meaning of these sentences? One is often faced with a problem starting with the phrase "characterize those ...". If we find a necessary and sufficient condition, how do we know it is not just restating the defining property, maybe in disguise? Both König's Theorem 1.1.1 and the Gallai Identity 1.0.2 provide formulas for $\nu(G)$. One feels, however, that König's theorem is the deeper result. Why?

These are extremely important questions. The fact that we are now able to answer them in a mathematically precise way has altered the whole of combinatorics. The idea of this answer occurs in a paper of Edmonds (1965c); its precise formulation is due to S. Cook (1971).

To make this simple — but somewhat sophisticated — idea clear, let us tell a story. In the court of King Arthur there dwelt 150 knights and 150 ladies-in-waiting. The king decided to marry them off, but the trouble was that some pairs hated each other so much that they would not even get married, let alone speak! King Arthur tried several times to pair them off but each time he ran into conflicts. So he summoned Merlin the Wizard and ordered him to find a pairing in which every pair was willing to marry. Now Merlin had supernatural powers and he saw immediately that none of the 150! possible pairings was feasible, and this he told the king. But Merlin was not only a great wizard, but a suspicious character as well, and King Arthur did not quite trust him. "Find a pairing or I shall sentence you to be imprisoned in a cave forever!" said Arthur. Fortunately for Merlin, he could use his supernatural powers to find the *reason* why such a pairing could not exist. He asked a certain 56 ladies to stand on one side of the king and 95 knights on the other side, and asked: "Is any one of you ladies, willing to marry any of these knights?", and when all said "No!", Merlin said: "O King, how can you command me to find a husband for each of these 56 ladies among the remaining 55 knights?" So the king, whose courtly education did include the pigeon-hole principle, saw that in this case Merlin had spoken the truth and he graciously dismissed him.

Some time elapsed and the king noticed that at the dinners served for the 150 knights at the famous round table, neighbors often quarrelled and even fought. Arthur found this bad for the digestion and so once again he summoned Merlin and ordered him to find a way to seat the 150 knights around the table so that each of them should sit between two friends. Again, using his supernatural powers Merlin saw immediately that none of the 149! seatings would do, and this he reported to the king. Again, the king bade him find one or explain why it was impossible. "Oh

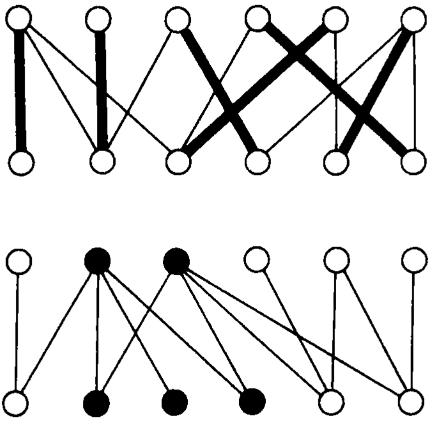


FIGURE 1A.1. A bigraph with a perfect matching and one without

I wish there were some simple reason I could give to you! With some luck there could be a knight having only one friend, and so you too could see immediately that what you demand from me is impossible. But alas!, there is no such simple reason here, and I cannot explain to you mortals why no such seating exists, unless you are ready to spend the rest of your life listening to my arguments!" The king was naturally unwilling to do that and so Merlin has lived imprisoned in a cave ever since. (A severe loss for applied mathematics!)

The moral of this tale is that there are properties of graphs which, when they hold, are easily proven to hold. If a graph has a perfect matching, or a Hamilton cycle, this can be "proved" easily by exhibiting one. If a bipartite graph does *not* have a perfect matching, then this can be "proved" by exhibiting a subset X of one color class which has fewer than $|X|$ neighbors in the other. The reader (and King Arthur!) are directed to Figure 1A.1 in which graph G has a perfect matching (indicated by the heavy lines), but graph H does not. To see the latter, consider the subgraph induced by the five black points.

If a graph is planar this can be "proved" by drawing it in the plane without intersections. (We put the word "prove" in quotation marks since, although these proofs are indeed proofs in the strict logical sense say, in Zermelo-Fraenkel set theory, they are in flavor quite different from mathematical proofs: they concern one special property of one special finite structure. The fact that such a proof exists is not a question here; it is the *length* of these proofs that matters. For this reason we shall prefer the word "exhibit" to the word "prove".)

A property of graphs which can be exhibited "easily" for each graph possessing it is called an *NP-property*. Here "NP" stands for *non-deterministic polynomial*. The phrase can be defined "easily" as follows: the number of steps in a formal proof that graph G has this property should not exceed a polynomial in the number of points and lines of G. This definition is somewhat arbitrary and its justification lies in the fact that the class of NP-properties defined this way is a very natural and useful class of properties. (In contrast, the letter "P" will be used to denote *deterministic polynomial.*)

The name "NP-property" itself refers to an alternative definition: a property is an **NP-property** if and only if it can be recognized by a *non*-deterministic Turing machine in polynomial time. Another nice way to define these properties is the following: they are precisely those properties of finite graphs which can be expressed by a second order formula containing one free second order variable ("adjacency") and an arbitrary number of existentially quantified second order variables. For the details of these descriptions of NP we refer the reader to Aho, Hopcroft and Ullman (1974), Fagin (1974) and Garey and Johnson (1979).

It is clear that NP-properties can be defined for structures other than graphs; for example, numbers, formulas, matrices, etc. The point here is that the length of the proof (exhibition) of the property should be polynomial in an appropriately defined "size" of the structure, which is usually just the number of digits in an appropriate encoding as a binary sequence. As a matter of fact, it is often more convenient to encode each structure as a binary sequence, and then define NP-properties for binary sequences only.

Most graph-theoretic properties which interest us are NP-properties or related to NP-properties. The two problems that Merlin had to face — the existence of a perfect matching and the existence of a Hamilton cycle — are clearly NP-properties. The 3-colorability of a graph, the existence of k independent points, connectivity, planarity and the isomorphism of two graphs are other NP-properties and there are many more (see Garey and Johnson (1979)). But NP-properties also appear quite frequently in other parts of mathematics. A very important NP-property of Boolean functions (mappings $\{0,1\}^n \to \{0,1\}$, which can be described by, say, polynomials of the operations $\wedge$, $\vee$ and $\neg$) is their *satisfiability*: the property that there is a choice for the variables for which the value of the function is 1. The existence of such a choice can be exhibited easily: just evaluate the function for this particular choice of variables. A very important NP-property of natural numbers is their **compositeness**: if a natural number is composite then this can be exhibited easily by showing a decomposition $n = ab$ $(a, b > 1)$.

The remarks we have made so far explain how Merlin might remain free if he is lucky and the task assigned to him by King Arthur has a solution. For instance, suppose he could find a good way to seat the knights. He could then convince King Arthur that his seating plan was "good" because the property of the corresponding friendship graph that it contains a Hamilton cycle is an NP-property. But how could he survive

Arthur's wrath in the case of the marriage problem and not in the case of the seating problem when these problems do *not* have solutions? What distinguishes the non-existence of a Hamilton cycle in a graph from the non-existence of a perfect matching in a bigraph? From our tale, we hope the answer is clear: *the non-existence of a perfect matching in a bigraph is also an NP-property* (this is a main implication of Frobenius' Theorem), while the non-existence of a Hamilton cycle in a graph is not! (To be precise, no proof of this latter fact is known, but there is strong evidence in favor of it. Cf. Box 6A).

So for certain NP-properties the negation of the property is again an NP-property. Such an NP-property is called **well-characterized** and a theorem asserting the equivalence of an NP-property with the negation of another NP-property is called a **good characterization**. There are famous good characterizations throughout graph theory and elsewhere. Kuratowski's theorem well-characterizes planarity of a graph; a theorem of König well-characterizes bipartite graphs; Euler's Theorem well-characterizes those graphs having an Eulerian trail; and numerous theorems in this book will well-characterize various other properties. In fact, seeking good characterizations wherever possible will be a leading motivational principle in our studies.

We shall say that a class of graphs is **in NP** if the property defining membership in this class is an NP-property. Similarly, a class is **in co-NP** if non-membership is defined by an NP-property.

We shall also be often concerned with various numerical functions defined on graphs (matching number, chromatic number, connectivity, etc.). It is sometimes convenient to extend the notion of a good characterization and say that an integer-valued graphical function $f(G)$ is **well-characterized**, if the property $f(G) \geq k$ is a well-characterized property of the pair (G, k). Thus König's Theorem gives a good characterization of the matching number of a bigraph: to exhibit $\nu(G) \geq k$, it suffices to exhibit k independent lines in G; to exhibit $\nu(G) \leq k$, it suffices to exhibit a point cover of G of size k. At this point the reader is invited to return once again to Figure 1A.1 and, using König's Theorem this time, try to convince the Good Sovereign of Camelot that graph G has a perfect matching, while graph H does not.

Now we can explain more precisely the difference in depth between König's Theorem and the Gallai Identity mentioned at the beginning of this box: The formula $\nu(G) = p - \rho(G)$ does not yield a good characterization of the matching number. In fact, to exhibit that $\nu(G) \leq k$ we would have to exhibit that $p - \rho(G) \leq k$, that is, $\rho(G) \geq p - k$. But this means that the points of G cannot be covered by fewer than $p - k$ lines and there is no immediate way to convince King Arthur of this.

Most graphical functions which shall concern us will be defined as the maximum or minimum of some value $v(G, p)$ over all values of the parameter p. Here p may range over a possibly exponentially large set, but usually it is the case that (a) it can be recognized which values of p belong to the parameter set, and (b) for every p in the parameter set, $v(G,p)$ can be computed in polynomial time. A **minimax theorem** (**minimax formula**) will be an identity of the form

$$\max_p v(G,p) = \min_q w(G,q),$$

where v and w satisfy the criteria (a) and (b) above. Just as König's Theorem does for the matching number, $\nu(G)$, every minimax theorem gives a good characterization of the common value of the two sides of this equation. A minimax theorem is the most common way to give a good characterization of a graphical function, and we shall meet minimax theorems throughout this book.

It is important to emphasize that the notion of a good characterization is a mathematical notion and not a judgement about the intrinsic merit of a given result!

1.2. A Bipartite Matching Algorithm: The Hungarian Method

A minimax result like König's Theorem 1.1.1 is both aesthetically pleasing and useful, but what if we want to actually find a maximum matching? Can we produce a good algorithm to do so? The idea of an "alternating path" provides us with an affirmative answer.

The concept of an alternating path, although quite simple, is one of the most important in all of matching theory. Indeed we shall return to it again and again throughout this book. In the present section we shall use alternating paths to provide a second proof of König's Minimax Theorem and then to formulate a good algorithm for finding a maximum matching in a bipartite graph. The more complicated problem of finding good matching algorithms for general (i.e., not necessarity bipartite) graphs will be the content of Chapter 9.

Let G be any (not necessarily bipartite) graph and M any matching in G. A path $P = v_1, v_2, \ldots, v_m$ is said to be an **alternating path with respect to** M or an M**-alternating path** if $v_iv_{i+1} \in M$ if and only if $v_{i+1}v_{i+2} \notin M$ for $1 \leq i \leq m-2$. A point v is **exposed** (**unmatched, unsaturated, not covered**) with respect to matching M if no line of M is incident with v. Clearly, if G contains any M-alternating path P joining two exposed points then M cannot be a maximum matching, for one readily obtains a larger matching M' by discarding the lines of $P \cap M$ and adding those of $P - M$. An M-alternating path joining two exposed points is called an M**-augmenting path** for this reason.

The preceding observation furnishes a proof of the "easy" half of the next theorem due to Berge (1957).

1.2.1. THEOREM. *Let M be a matching in a graph G. Then M is a maximum matching if and only if there exists no augmenting path in G relative to M.*

PROOF. By virtue of our remark prior to the theorem we will be finished if, given a matching M which admits no augmenting path, we can show it to be maximum. To this end, let M' be any maximum matching and form $M \oplus M'$. Then the components of $M \oplus M'$ must consist of even cycles and alternating paths. But no alternating path component may be M-augmenting by hypothesis and hence each is even in length. But then $|M| = |M'|$ and the proof is complete. ∎

At first glance it may appear — and in fact until the work of Edmonds (1965a) it was generally believed — that the above result furnishes an "adequate" algorithm for finding a maximum matching starting, say, with a single line. Of course it does indeed provide an algorithm, but not a *good* one! We shall have much more to say later about good and bad algorithms. Suffice it to say here that to use the above algorithm we must check, in the worst case, *all* matchings in G and we therefore find that we require the answer to get the answer! So from a computational point of view this theorem by itself is useless. (For further elaboration on "good" algorithms, see Boxes 1A and 1B in this chapter.) But the idea of an augmenting path is far from useless in matching theory. On the contrary, the reader will see many applications of this fundamental concept as he makes his way through this book.

Now let us return to the task at hand, namely to find a good bipartite matching algorithm. Let G be a bigraph with bipartition (A, B) and let M be any matching in G. Suppose A_1 and B_1 are the sets of exposed points in A and B respectively. We want to find an M-augmenting path, if any, connecting A_1 to B_1. To this end we consider the set U of points in A accessible from A_1 on an M-alternating path. We can construct set U as follows. Form a maximal forest F in G with the properties:

(i) each point b of F in B has degree 2 and one of the forest lines incident with b belongs to M and

(ii) each component of F contains a point of A_1.

Note that $A_1 \subseteq V(F)$ for any point in A_1, but not in F, can be added to F as a singleton component. Then it is easy to see that $U = V(F) \cap A$, but we shall not need this fact below.

1.2.2. LEMMA. *If G, M, A, B, A_1, B_1, and F are as above, then M is a maximum matching if and only if no point of B_1 is adjacent to any point of F.*

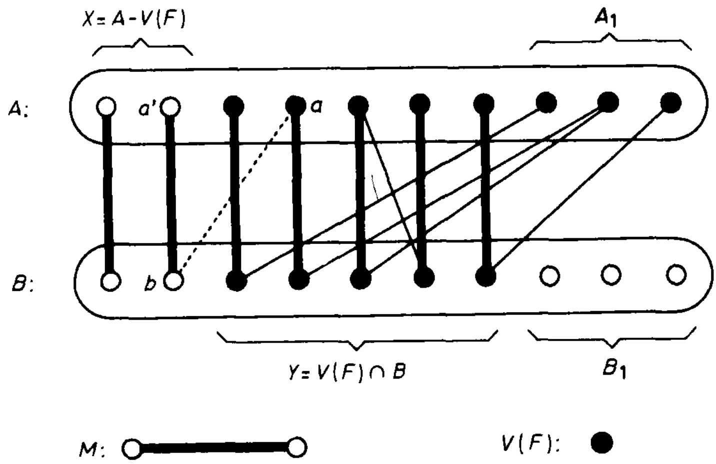


FIGURE 1.2.1.

PROOF. First suppose there is a point x of F incident with a point $b_1 \in B_1$. Then x is in A and by property (ii) and the maximality of F there is a path P joining x to a point $a \in A_1$. Hence $P + xb_1$ is an M-augmenting path and thus M is not maximum.

To prove the converse, suppose no line joins a point of F to one of B_1. Let $X = A - V(F)$ and $Y = V(F) \cap B$. We claim that $X \cup Y$ is a point cover and that $|X \cup Y| = |M|$.

Clearly M covers $X \cup Y$. On the other hand, no line of M covers two points of $X \cup Y$ for if $xy \in M$ and $y \in V(F) \cap B$, since the lines of M are independent it follows (from (i)) that $x \in V(F)$, a contradiction. Hence $|X \cup Y| = |M|$ and each line of M contains exactly one point of $X \cup Y$. See Figure 1.2.1.

It remains to show that $X \cup Y$ is a cover. Indirectly, suppose line ab is not covered by $X \cup Y$ where $a \in A$ and $b \in B$. Then $a \in V(F)$ and $b \notin V(F)$. Moreover, by hypothesis, $b \notin B_1$. Thus matching M covers b, say by line $a'b$. Moreover, $a \neq a'$ since if a is matched by M, it must be matched into Y. But now F can be extended to a larger forest F' containing path aba', contradicting maximality.

So $X \cup Y$ covers $E(G)$ and we have

$$\tau(G) \leq |X \cup Y| = |M|. \tag{1.2.1}$$

Thus M is a maximum matching. ■

1.2.3. EXERCISE. Derive König's Minimax Theorem from Lemma 1.2.2.

At this point it is easy to present an algorithm for finding a maximum matching in a bipartite graph based on the above.

Algorithm for bipartite matching:

1^0. Start with any matching M.

2^0. Form a maximal forest F having properties (i) and (ii) above.

3^0. If there is a line joining $V(F) \cap A$ to a point of B_1, we can obtain a new larger matching M' as in the proof of Lemma 1.2.1. Return to step 1.

4^0. If no line joins $V(F) \cap A$ to B_1, then again by Lemma 1.2.1, M is maximum. (Moreover, if $X = A - V(F)$ and $Y = V(F) \cap B$, $X \cup Y$ is a minimum cover for G.) End.

It is interesting to note that the approach used in this algorithm has come to be called the **Hungarian Method** since it seems to have first appeared in the work of König (1916a, 1916b, 1931, 1936) and of Egerváry (1931) who reduced problems with general non-negative weights on the lines to the case where line weights are 0 or 1 as treated by König. (See Section 7.1.) Also the forest present at the final stage of the algorithm, that is, when the matching is maximum, has come to be called a **Hungarian forest** for the same historical reasons.

The adjective "Hungarian" in this context seems to have been used first by Kuhn (1955, 1956) who cast the procedure in terms of a primal-dual linear program. The corresponding algorithm can be implemented so as to produce an optimal matching (in both the weighted and unweighted cases) in $O(q^2p)$ steps. For details see Lawler (1976). (We also recommend Lawler for a clear and thorough treatment of the implementation of matching algorithms in general.)

In the *un*weighted case, Hopcroft and Karp (1971, 1973) currently have the fastest bipartite matching algorithm. They proved termination in $O(p^{5/2})$ steps, but recently Galil (1983) has shown their algorithm to be even faster — $O(qp^{1/2})$ — especially if the graph is sparse. Their algorithm is based upon network flow techniques which will be discussed in Chapter 2.

The fastest *weighted* bipartite matching algorithm to date is composed of a primal-dual method together with some "fine tuning" of the data structure. It runs in $O(qp\log_{\lceil q/p+1\rceil} p)$ time and is due to Galil (1983).

BOX 1B. On Algorithms

A characteristic feature of recent development in combinatorics and graph theory is the increasingly important role of algorithms. The main reason for this is certainly the mushrooming use of computers, but probably in part it is also due to the intrinsic development of the subject. Just how far the "Theorem-Proof" style of mathematics we learned at school will shift to the "Algorithm-Analysis of algorithm" style is difficult to predict. It is certainly true that while writing this book we have been meeting more and more results and methods of combinatorial algorithms whose significance for the subject as a whole is growing. We shall in fact discuss algorithmic aspects of matching theory at several points. Yet we have chosen to hold to the traditional "Theorem-Proof" approach and derive the algorithms as consequences of the theorems and/or the proofs. This is not the only possible point of view. Most of the structure theory for matchings in graphs, which is the central theme of our book, could be developed as a consequence of a careful analysis of just *one* algorithm, the matching algorithm of Edmonds (see chapter 9).

In combinatorial situations the mere algorithmic solvability of a particular problem is not the question; the problems are such that they can be solved trivially by considering a finite number of cases. The point is that in a naive approach the number of cases to consider — though finite — is usually exponentially large in the size of the input. For example, we may have to consider all subsets of points and lines of a graph, while the size of the input is only the number of points and lines. So we are interested in finding faster algorithms (and also in algorithms with small storage requirements etc., but for the sake of simplicity let us consider running time only). It is customary to distinguish as particularly nice those algorithms whose running time is bounded by a polynomial in the size of the input (the so-called "good" or "polynomial" algorithms). This method of categorizing is, of course, rather arbitrary and is often criticized for not reflecting the practical value of the algorithms accurately. Indeed, an algorithm which takes $1000n^7$ steps for every input of size n is clearly inferior to one which makes 2^n steps for a very few "unlucky" inputs and many fewer steps on the average.

There are, however, some good arguments for distinguishing the class of polynomial algorithms. First, as problems grow in size the computation time of a polynomial algorithm grows much more slowly than that of an exponential one, so refinement of implementation and improvement in computer technology may balance the increase of input size much longer than is the case for an exponential time algorithm. In the case of exponential time algorithms an abrupt jump, the "exponential explosion", renders them infeasible even for a small increase in problem size. There is a syntactical difference between polynomial and non-polynomial algorithms: roughly speaking, a polynomial algorithm can be programmed using a given number of cycles but no jumps, while a slower algorithm cannot. The notion of polynomiality has proved very useful in theoretical investigations. Finally, there is a (completely subjective) feeling that it is the polynomial algorithms which "go to the heart of the problem".

Thus when we discuss algorithms we focus on their polynomiality. Obviously, the mere polynomiality of an algorithm does not automatically render it practical, and in most cases a considerable amount of further work is needed to actually implement these algorithms. We analyse our algorithms only up to the point where this analysis involves graph theoretical ideas. Further refinements based on programming, data manipulation etc. are beyond the scope of our book and we will only give references. Also, we will not discuss certain algorithms which are practical, but whose graph-theoretic content does not go beyond the enumeration of all cases.

1.3. Deficiency, Surplus and a Glimpse of Matroid Theory

The following notion was first studied by Ore (1955). Let G be a bipartite graph with bipartition (A, B). For $X \subseteq A$, define

$$\mathrm{def}_G(X) = \mathrm{def}(X) = |X| - |\Gamma(X)|,$$

and call this number the **deficiency** of the set X. The maximum deficiency of all subsets of A will be called the A-**deficiency** of G. If A is understood, we shall simply call this number the **deficiency** of G, and denote it by $\mathrm{def}(G)$. Note that since $\mathrm{def}(\emptyset) = 0$, we have $\mathrm{def}(G) \geq 0$. The following theorem is an easy consequence of the König or P. Hall Theorems of Section 1.1 and hence these two classical results are said to be *self-refining* just as is the Marriage Theorem. We will see other self-refining theorems later in our studies, perhaps the most notable being Tutte's Theorem 3.1.1.

1.3.1. THEOREM. *The matching number of a bipartite graph G is $|A| - \mathrm{def}(G)$.* ∎

We will find it worth-while to study the properties of the deficiency function $\operatorname{def}(X)$. In these studies we shall not make use of Theorem 1.3.1. In fact, a proof of this theorem could easily be obtained from Lemmas 1.3.2 and 1.3.3 below (see Exercise 1.3.4).

1.3.2. LEMMA. *For every two subsets $X, Y \subseteq A$ we have*

$$\operatorname{def}(X \cup Y) + \operatorname{def}(X \cap Y) \geq \operatorname{def}(X) + \operatorname{def}(Y). \tag{1.3.1}$$

PROOF. By a trivial computation we have

$$|X \cup Y| + |X \cap Y| = |X| + |Y|. \tag{1.3.2}$$

It is also trivial that

$$\Gamma(X \cup Y) = \Gamma(X) \cup \Gamma(Y)$$

and

$$\Gamma(X \cap Y) \subseteq \Gamma(X) \cap \Gamma(Y).$$

Hence

$$|\Gamma(X \cup Y)| + |\Gamma(X \cap Y)| \leq |\Gamma(X)| + |\Gamma(Y)|. \tag{1.3.3}$$

Upon subtracting (1.3.3) from (1.3.2), the lemma follows. ■

A subset $X \subseteq A$ is called **tight** if $\operatorname{def}(X) = \operatorname{def}(G)$.

1.3.3. LEMMA. *The intersection and union of tight sets are tight.*

PROOF. Suppose that X and Y are tight. Then the right hand side of inequality (1.3.1) is $2\operatorname{def}(G)$. On the other hand, the left hand side is at most $2\operatorname{def}(G)$ by maximality in the definition of $\operatorname{def}(G)$. So we must have equality and $X \cup Y$ and $X \cap Y$ must also be tight. ■

1.3.4. EXERCISE. Let G be a bipartite graph such that deleting any line from G, $\operatorname{def}(G)$ increases. Prove that every point of G has degree at most 1. Obtain a new proof of Theorem 1.3.1 this way.

Many of these ideas carry over if we are interested in the deficiency of *non-empty* sets only. The exclusion of the empty set from our considerations has an effect only if every non-empty subset of A has negative deficiency. It is therefore more convenient to change sign and define the **surplus of a set** $X \subseteq A$ by

$$\sigma_G(X) = \sigma(X) = |\Gamma(X)| - |X| = -\operatorname{def}(X).$$

The **surplus of the bigraph** G is the minimum surplus of *non-empty* subsets of A, and will be denoted by $\sigma(G)$. Note that $\operatorname{def}(G) = \max\{0, -\sigma(G)\}$. Lemma 1.3.2 implies immediately that

$$\sigma(X \cup Y) + \sigma(X \cap Y) \le \sigma(X) + \sigma(Y). \tag{1.3.4}$$

Call a set $X \subseteq A$ **σ-tight** if $\sigma(X) = \sigma(G)$. Then the proof of Lemma 1.3.3 carries over to yield the following result.

1.3.5. LEMMA. *If two σ-tight sets have a non-empty intersection, then their intersection and union are also σ-tight.* ∎

The following theorem, together with P. Hall's Theorem, yields a good characterization of the number $\sigma(G)$, since if $\operatorname{def}(G) > 0$ then $\sigma(G) = -\operatorname{def}(G)$ and $\operatorname{def}(G)$ is known to be well-characterized. Thus it suffices to consider the case when $\operatorname{def}(G) = 0$.

1.3.6. THEOREM. *For a bigraph G, with* $\operatorname{def}(G) = 0$, *$\sigma(G)$ is the largest integer s satisfying the following property for every point $x \in A$: if we add s new points to A and connect them to the points in $\Gamma(x)$, the resulting bigraph has non-negative surplus.*

PROOF. First we show that $s = \sigma(G)$ has the property formulated in the theorem. Let $x \in A$, and let $x_1, \ldots, x_s$ be new points, connected to the points in $\Gamma(x)$. We have to show that in the resulting bigraph G', $|\Gamma_{G'}(X)| \ge |X|$ holds for every $X \subseteq A \cup \{x_1, \ldots, x_s\}$, $X \ne \emptyset$. If $X \subseteq A$ then $|\Gamma_{G'}(X)| = |\Gamma_G(X)| \ge |X| + s \ge |X|$. If $X \not\subseteq A$ then $|\Gamma_{G'}(X)| = |\Gamma_G((X \cap (A - x)) \cup \{x\})| \ge |(X \cap (A - x)) \cup \{x\}| + s \ge |X \cap A| + s \ge |X|$.

Second, assume that s is a number with the property formulated in the theorem. We show that $s \le \sigma(G)$; this will complete the proof of the theorem. Let $X \subseteq A$, $X \ne \emptyset$. Select any $x \in X$, and add s new points $x_1, \ldots, x_s$ to the graph, connecting them to $\Gamma(x)$. By hypothesis, the bigraph G' arising in this way has non-negative surplus. But then

$$|\Gamma_G(X)| = |\Gamma_{G'}(X \cup \{x_1, \ldots, x_s\})| \ge |X \cup \{x_1, \ldots, x_s\}| = |X| + s.$$ ∎

To see how this theorem provides a good characterization of the surplus of a bigraph consider the following: we can prove that $\sigma(G) \ge s$ by exhibiting $|A| + s$ independent lines in each of $|A|$ bigraphs obtained from G by adding s new points and connecting them to the points of $\Gamma(x)$ for some $x \in A$. On the other hand, we can prove that $\sigma(G) \le s$ by exhibiting a non-empty set $X \subseteq A$, such that $|\Gamma(X)| \le |X| + s$.

Even though Theorem 1.3.6 gives a good characterization of the surplus, the condition given in it is not too transparent. In the case $\sigma(G)=1$ we can give a much nicer condition. First we prove the following result of Las Vergnas (1970) and Lovász (1970c).

1.3.7. LEMMA. *Let G be a bigraph with $\sigma(G)=s>0$ such that deleting any line from G, $\sigma(G)$ decreases. Then every point in A has degree $s+1$.*

PROOF. Let $x \in A$, and $\Gamma(x)=\{y_1,\cdots,y_d\}$. We want to show that $d=s+1$. By definition of s, $d \geq s+1$. Deleting the line xy_i from G, the remaining graph G' will have surplus less than s, and thus it will contain a set $X_i \subseteq A$, $X_i \neq \emptyset$ with $\sigma_{G'}(X_i)<s$. Since $\sigma_G(X_i) \geq s$, it follows that X_i must be a σ-tight subset in G, $x \in X_i$, and moreover, y_i must be adjacent to a single point of X_i, namely x. Then by Lemma 1.3.5, the set $X_0 = X_1 \cap \ldots \cap X_d$ is also σ-tight, and no line connects $\Gamma(x)=\{y_1,\ldots,y_d\}$ to $X_0 - x$. If $X_0=\{x_0\}$ we are finished so suppose, to the contrary, that $X_0 - x_0 \neq \emptyset$. Then

$$|\Gamma(X_0)| = |\Gamma(x)| + |\Gamma(X_0 - x_0)| \geq (1+s) + (|X_0 - x| + s) = |X_0| + 2s,$$

which is a contradiction since X_0 is σ-tight. ■

So now we have a characterization of all *minimal* (with respect to line deletion) positive surplus bigraphs. It follows from the next theorem that they are just the forests in which each point in A has degree two.

1.3.8. THEOREM. *A bigraph G has positive surplus if and only if it contains a forest F such that every point in A has degree 2 in F.*

PROOF. It is trivial that if G contains a forest F such that every point of A has degree 2 in F then G has positive surplus. Conversely, assume that G has positive surplus, and let F be a minimal subgraph of G which contains all points of A and has positive surplus. By Lemma 1.3.7, every point of A has degree 2 in F. Hence it also follows that F is a forest. In fact if C were a cycle in F then $V(C) \cap A$ would have surplus 0. ■

Graphs with positive surplus will play an important role as building blocks in the structure theory of graph matching (cf. Chapter 3).

Expressions like (1.3.1), (1.3.2) and (1.3.3) play an important role in most branches of combinatorics. A real-valued set function, i.e., a mapping $f: 2^S \to \Re$, is called **submodular** if it satisfies

$$f(X \cup Y) + f(X \cap Y) \leq f(X) + f(Y) \qquad (X, Y \subseteq S),$$

it is called **supermodular** if

$$f(X \cup Y) + f(X \cap Y) \geq f(X) + f(Y) \qquad (X, Y \subseteq S),$$

and **modular** if it is both submodular and supermodular, that is, if it satisfies both of the above inequalities with equality. Supermodular and submodular functions often enable us to formulate and prove theorems with great unifying power.

Thus, for example, surplus is a submodular, deficiency is a supermodular and cardinality is a modular set function on the point set of a bigraph. The reader is invited to show by example that if the graph under consideration is not bipartite, then deficiency and surplus are neither submodular nor supermodular in general.

We are going to mention some examples of super- and submodular set functions. In some of these examples the verification of (super or sub)-modularity is not quite trivial. The following lemma will be useful here as well as later on. (The proof is left to the reader.)

1.3.9. LEMMA. *A set function f defined on the subsets of a finite set S is submodular if and only if the set function $f(X \cup \{a\}) - f(X)$ is monotone decreasing on the subsets of $S - a$, for every $a \in S$.* ■

1.3.10. EXAMPLE. Let ϕ be any function defined on the elements of any finite set S, let c be any real number and let

$$f(X) = c + \sum_{a \in X} \phi(a).$$

Then f is a modular set function. Conversely, every modular set function arises this way.

1.3.11. EXAMPLE. Let D be a digraph and let $\rho(X)$ denote the number of lines of D entering the set $X \subseteq V(D)$. Then ρ is a submodular function on the subsets of $V(D)$. In fact, it is straightforward to verify that every line counted in $\rho(X \cup Y) + \rho(X \cap Y)$ is also counted in $\rho(X) + \rho(Y)$ and every line counted in both terms of the first sum is also counted in both terms of the second. (The lines between $X - Y$ and $Y - X$ are only counted in the second sum. This is why we cannot assert modularity, but only submodularity.)

1.3.12. EXAMPLE. Let G be a graph, and for $X \subseteq E(G)$, let $r(X) = |V(G)| - c\big(V(G), X\big)$. (Here $\big(V(G), X\big)$ is the graph with point set $V(G)$ and line set X, and $c\big(V(G), X\big)$ is the number of connected components of this graph.) Then, $c(X) = c\big(V(G), X\big)$ is a supermodular function and $r(X)$ is a submodular function. This follows easily from Lemma 1.3.9.

1.3.13. EXAMPLE. Let G be a graph and $A \subseteq V(G)$ an independent set of points. Define $f(X) = c(G - X)$ for $X \subseteq A$. Then this set function is supermodular. To verify this, let $X' = E(G - X)$ for $X \subseteq A$, and note that $(X \cap Y)' = X' \cup Y'$ and $(X \cup Y)' = X' \cap Y'$. (Here we have used the fact that A is an independent set, or more exactly, that no line connects $X - Y$ to $Y - X$.) Further, if r is as in the preceding example,

$$f(X) = c(G - X) = c\big(V(G), X'\big) - |X| = |V(G)| - |X| - r(X'),$$

and so using the submodularity of r, the assertion follows.

1.3.14. EXAMPLE. Let A be a matrix and let S denote the set of columns of A. For $X \subseteq S$, let $r(X)$ denote the rank of the matrix formed by the columns in X. Then r is submodular. The proof follows by Lemma 1.3.9.

1.3.15. EXAMPLE. Our last example is more closely related to bipartite matching theory. Let G be a bipartite graph with bipartition (A, B) and for $X \subseteq A$, let $r(X)$ denote the maximum number of points in X which can be matched with points in B; i.e., let $r(X) = \nu\big(G - (A - X)\big)$. Then by Theorem 1.3.1,

$$r(X) = |X| + \min\{\sigma(Y) \mid Y \subseteq X\}.$$

It is easy to see that $r(X)$ is a submodular set function. In fact, if f is any submodular set function then the set function f_1 defined by

$$f_1(X) = \min\{f(Y) \mid Y \subseteq X\}$$

is also submodular.

Let us remark that the set functions in Examples 1.3.12, 1.3.14 and 1.3.15 have some additional properties: $r(\emptyset) = 0$, r is monotone increasing, and $r\big(\{x\}\big) \leq 1$ for every $x \in S$. A finite set S endowed with such a submodular function is called a **matroid** and r is called the **rank function** of the matroid.

To be more precise, let S be any set and r a function from $2^S \to Z_+$. Then r is called a **matroid rank function** on S if the following three conditions hold:

(a) if $X \subseteq S$ then $r(X) \le |X|$,
(b) if $X \subseteq Y \subseteq S$ then $r(X) \le r(Y)$,
(c) if $X, Y \subseteq S$ then $r(X \cup Y) + r(X \cap Y) \le r(X) + r(Y)$.

The pair (S, r) is called a **matroid** if r is a matroid rank function on S.

Matroids are of great importance in combinatorics and graph theory, and even though the scope of this book does not permit a systematic development of matroid theory, we shall encounter matroid theoretic ideas repeatedly in matching theory. An important feature of matroids is that they have a number of alternate definitions, of which we discuss only one — the definition in terms of independent set. For a comprehensive treatment see Welsh (1976) or von Randow (1975).

Let (S, r) be a matroid. Call a subset $A \subseteq S$ **independent** if $r(A) = |A|$. Then we can show that the following properties hold:

(I–1) $\emptyset$ is independent.

(I–2) If A is independent and $B \subseteq A$ then B is independent.

(I–3) If A and B are independent and $|A| > |B|$ then there exists an element $a \in A - B$ such that $B \cup \{a\}$ is independent.

Conversely, if $\mathcal{M}$ is a family of subsets of S, called independent, and axioms (I–1), (I–2), (I–3) are satisfied, then $\mathcal{M}$ is the set of independent sets of a unique matroid. The **rank** of set X is then the size of its maximum independent subset.

We shall not go into many particulars about matroids. There are, however, a few basic notions which we must introduce. A set which is not independent is called **dependent**, and a minimal (with respect to inclusion) dependent set is called a **circuit.** It is easy to derive from the submodularity of the rank function that every set X has a unique largest superset with the same rank as X; this superset is called the **closure** of X. A set X is called **closed** if it coincides with its own closure; that is, if the addition of any new element to X increases its rank. Closed sets are also called **flats**.

A trivial matroid in which every set is independent (or equivalently, in which the rank of any set is equal to its cardinality) is called a **free** matroid. Less trivial examples can be found in Box 1C, which also contains a brief historical survey of matroids. If the reader finds some of the above names for matroid concepts a bit peculiar, we hope this box will also help explain these appellations.

Submodular functions and matroids give rise to a large number of important minimax theorems, which often generalize fundamental minimax theorems in graph theory, such as the theorems of König, Menger, and others.

Let us state one rather general result which gives rise to many minimax theorems. This result is due to Frank (1982). By rather complicated arguments it can be shown that the result is equivalent to other theorems on submodular set functions, due to Edmonds (1970), Edmonds-Giles (1977) and Martel (1981), but Frank's result is probably the easiest to

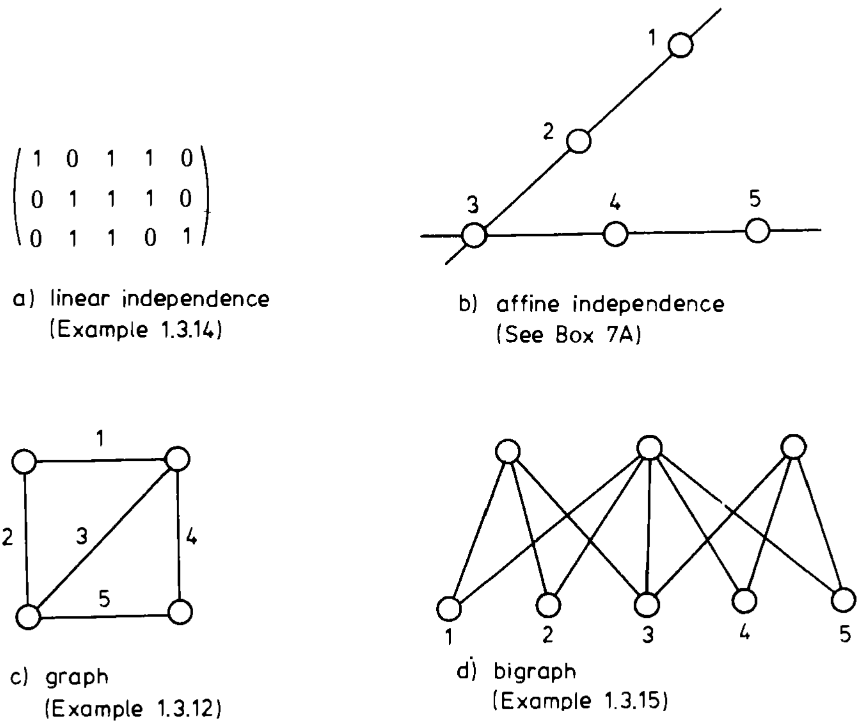


FIGURE 1.3.1. Different representations of the same matroid

state. This theorem shows an interesting analogy between submodular-supermodular and convex-concave. This analogy can be further exploited and explained (see Lovász (1983a)).

1.3.16. THEOREM. *Let f be a submodular and g a supermodular set function defined on the same set S and assume that $f \geq g$. Then there* *exists a modular set function h such that $f \geq h \geq g$. If, in addition, f and g are integer-valued then h may be chosen integer-valued.*

PROOF. We give a proof of the second statement, which is the interesting one from the point of view of combinatorial applications and from which the first statement follows as well. The proof will use an idea similar to the Halmos-Vaughn proof of P. Hall's Theorem.

We use induction on $|S|$. For $|S| = 1$ the assertion is obvious. Assume first that there exists an integer-valued submodular set function f_1 such that $f \geq f_1 \geq g$, $f \neq f_1$. Then it would suffice to separate f_1 from g by a modular function. Proceeding similarly we see that it suffices to prove the theorem for the case when no integer-valued submodular set function separating f from g exists (other than f itself). We may assume without loss of generality that $f(\emptyset) = 0$.

Next we claim that if

$$f(S-a) + f(a) = f(S) \tag{1.3.5}$$

holds for every $a \in S$ then f is, in fact, modular and hence we are done.

To prove this it suffices to use the monotonicity property of the difference function $f(X) - f(X-a)$. If $a \in X$, then

$$f(a) = f(a) - f(\emptyset) \geq f(X) - f(X-a) \geq f(S) - f(S-a) = f(a),$$

whence

$$f(X) = f(X-a) + f(a),$$

and hence by induction

$$f(X) = \sum_{a \in X} f(a),$$

which is indeed a modular set function. So we may assume that there exists an element $a \in S$ such that $f(S-a) + f(a) > f(S)$. Define a new function

$$f_1(X) = \begin{cases} \min\bigl(f(X),\, f(X-a) + f(a) - 1\bigr) & \text{if } a \in X, \\ f(X) & \text{if } a \notin X. \end{cases}$$

It is straightforward to verify that f_1 is submodular. Trivially, $f \geq f_1$ and $f \neq f_1$ since $f_1(a) = f(a) - 1$. So by assumption, $f_1 \not\geq g$, that is, there exists a set A such that $f_1(A) < g(A)$. Since also obviously $f_1(A) \geq f(A) - 1$, we have that $f_1(A) = f(A) - 1 = g(A) - 1$. It is immediate that $A \neq \emptyset$ and $A \neq S$.

Thus there exists a set $A \neq \emptyset, S$ such that $f(A) = g(A)$. Now consider these two set functions restricted to A. By the induction hypothesis there exists a modular set function h_1 on A separating f and g. Also consider $f(A \cup X)$ as a set function on $S - A$. This is trivially submodular. Similarly, the set function $g(A \cup X)$ is supermodular on $S - A$. Hence again by the induction hypothesis there exists a modular set function h_2 on $S - A$ separating them. We claim that the set function

$$h(X) = h_1(X \cap A) + h_2(X - A) - f(A)$$

defined on all subsets of S, is modular and separates f and g. The first assertion is obvious. The second follows by the computation

$$\begin{aligned} h(X) &= h_1(X \cap A) + h_2(X - A) - f(A) \leq f(X \cap A) + f(X \cup A) - f(A) \\ &\leq f(X), \end{aligned}$$

by the submodularity of f. The fact that $h \geq g$ follows similarly. ∎

By specializing, Theorem 1.3.16 can be used to obtain many deep results in combinatorics, such as the theorems of König and Menger, the results on orientations of graphs (cf. Section 2.4) and others. Here we discuss only a very few of these consequences.

First we derive the famous "Matroid Intersection Theorem" of Edmonds (1970).

1.3.17. THEOREM. (*The Matroid Intersection Theorem*). *Suppose (S, r_1) and (S, r_2) are two matroids on the same underlying set. Then the maximum size of a common independent set of the two matroids is equal to the minimum of $r_1(X) + r_2(S - X)$ over all $X \subseteq S$.*

PROOF. Let $k = \min\{r_1(X) + r_2(S - X) \mid X \subseteq S\}$ and set

$$\begin{aligned} f(X) &= \min(k, r_1(X)), \\ g(X) &= \max(0, k - r_2(S - X)). \end{aligned}$$

Then clearly f is submodular, g is supermodular, and by the definition of k, $f \geq g$. Thus by Theorem 1.3.16, there exists an integer-valued modular function h separating f and g. Since $f(\emptyset) = g(\emptyset) = 0$, we have $h(\emptyset) = 0$ and since $0 \leq g(a) \leq f(a) \leq 1$ for every $a \in S$, we have $h(a) = 0$ or 1. Let $A = \{a \in S \mid h(a) = 1\}$. Then $h(X) = |X \cap A|$ for every X.

Now $h(A) = |A| \leq f(A)$ implies that A is independent in (S, r_1) and that $|A| \leq k$. On the other hand, $h(S - A) = 0 \geq g(S - A)$ implies that $r_2(A) \geq k$. This implies that $r_2(A) = |A| = k$, that is, A is a common independent set of the two matroids of size k.

That no common independent set B of the two matroids can be of larger cardinality than k follows easily for if X is a subset minimizing $r_1(X)+r_2(S-X)$ then

$$\begin{aligned}|B| &= |B\cap X| + |B\cap(S-X)| = r_1(B\cap X) + r_2\big(B\cap(S-X)\big)\\ &\le r_1(X)+r_2(S-X) = k.\end{aligned}$$ ∎

Let us state another version of the Matroid Intersection Theorem which shows clearly that it is a generalization of König's Theorem (Aigner and Dowling (1971)).

1.3.18. THEOREM. *Let G be a bigraph with bipartition (A,B) and let (A,r_A) and (B,r_B) be two matroids. Then the maximum size of a matching in G which meets an independent set in both matroids is equal to the minimum rank of a point cover, i.e., the minimum of $r_A(T\cap A)+r_B(T\cap B)$, over all point covers T.*

PROOF. Let $S=E(G)$, $r_1(X)=r_A\big(V(X)\cap A\big)$, $r_2(X)=r_B\big(V(X)\cap B\big)$, and apply the Matroid Intersection Theorem. ∎

Clearly König's Theorem is now obtained by setting $r_A(X)=|X|$ and $r_B(Y)=|Y|$ for every $X\subseteq A$, $Y\subseteq B$.

BOX 1C. Matroids

Matroids were introduced in the early thirties by Birkhoff, van der Waerden and Whitney, who arrived at this important notion from different directions. In his book "Moderne Algebra" (1931), van der Waerden points out that linear independence of sets of vectors over a field and algebraic independence of sets of field elements over a subfield have similar combinatorial properties, namely the properties (I-1) — (I-3) for independent sets of matroids mentioned before Theorem 1.3.16. Matroids arising from linear independence of vectors are nowadays called **linear** or **coordinatizable**.

Whitney (1935) introduced the name **matroid** and laid the foundations of matroid theory. One of his motivations for doing so was the study of planar graphs and duality. A planar graph, when embedded in the plane in different ways, may have non-isomorphic duals, but all duals of the same graph have many properties in common. There is a natural bijection between the lines of any two duals of the same graph, which maps cycles onto cycles and, consequently, forests onto forests. So what is common to two duals of the same graph is the system of those subsets of lines which form forests. Whitney noticed that these subsets form the independent sets of a matroid. Matroids arising this way are usually called **graphic**. Thus matroids may be viewed as generalizations of graphs. One of the nice features of this generalization is that a "dual" of every matroid can be defined in a way which generalizes duality for planar graphs. The dual of the matroid of a non-planar graph turns out to be a matroid which does not arise from a graph; this shows the advantage of introducing this more general notion.

Birkhoff (1967) has considered semimodular, atomic lattices; he called them **geometric** because the lattice of subspaces of various geometries (projective, affine, hyperbolic etc.) is geometric. As it turns out, finite geometric lattices are essentially equivalent to matroids. This approach, however, shall not concern us here.

A further development in the theory of matroids, which is of great relevance to us, is found in the work of Rado (1942). He proved a minimax theorem which is essentially equivalent to the special case of Theorem 1.3.18 where $r_B(X) = |X|$. (It should be noted, however, that Theorem 1.3.18 could be reduced to this special case.) He also introduced the matroids in Example 1.3.15 which are called **transversal matroids**. The independent sets in these matroids are precisely those subsets of A which can be matched.

Rado's theorem already suggests the importance of matroids in combinatorial optimization, but the most striking development in this direction is due, independently, to Rado (1957), Edmonds (1971), Gale (1968) and Welsh (1968). Let S be a finite set and $\mathcal{A}$, a non-empty collection of subsets of S such that if $A \in \mathcal{A}$ then every subset of A also belongs to $\mathcal{A}$. (In brief, $\mathcal{A}$ is a **hereditary hypergraph**, **independence system** or **simplicial complex**). A very common and general problem in combinatorial optimization is the following. Assume that there is a non-negative weight $w(a)$ associated with every $a \in S$. Find a subset $A \in \mathcal{A}$ for which $\sum_{a\in A} w(a)$ is maximum.

A trivial approach to this problem is the following. Pick an element $a_1 \in A$ such that $\{a_1\} \in \mathcal{A}$ and $w(a_1)$ is maximum. Pick $a_2 \in S$ such that $\{a_1, a_2\} \in \mathcal{A}$ and $w(a_2)$ is maximum etc. Stop if no further element can be added to the set previously selected in such a way that we get a subset in $\mathcal{A}$. This procedure results in a subset of S which belongs to $\mathcal{A}$ and for which $w(A)$ can be (heuristically) expected to be "large", although of course not necessarily maximum! This procedure is called the **Greedy Algorithm**. The four authors mentioned above observed that the Greedy Algorithm produces the true optimum for every weighting of the elements of S if and only if $\mathcal{A}$ is the set of independent sets of a matroid.

It is impossible to survey even the basic results of matroid theory in a side remark in a book such as this, so we only refer to the textbooks and monographs by Tutte (1965, 1966, 1971), Crapo and Rota (1968), von Randow (1975), Aigner (1975, 1976, 1979) and Welsh (1976). Let us note at least, however, that there are two (interrelated, but distinguishable) main lines of research. One is concerned with minimax theorems and the application of matroids in combinatorial optimization. The Matroid

Intersection Theorem 1.3.17 is a characteristic example of this kind of result. This trend in matroid theory is experiencing repeated encounters with matching theory; a recent one will be described in Chapter 11.

The other main branch in matroid theory is the study of representation of matroids. That is, which matroids are representable by the columns of matrix over a given (or, over some) field (Example 1.3.14), or by a graph (Examples 1.3.12, 1.3.15). Whitney (1935) characterized those matroids representable by a matrix over the field GF(2), Tutte (1965) characterized those representable by a matrix over every field, and many important results in this direction are currently being obtained (see, for example, Seymour (1980)). However, a complete solution of the representation problem is still not known.

1.4. Some Consequences of Bipartite Matching Theorems

Philip Hall (1935) first proved his version of the bipartite matching theorem, not in the language of graph theory, but in a set-theoretic context. Let $S_1, \dots, S_n$ be (not necessarily distinct) subsets of a finite set S. When can we find a set of n *distinct* elements $s_1, s_2, \dots, s_n$ with $s_i \in S_i$? Such a set of elements is called a **system of distinct representatives** or simply an **SDR**. (Let us note in passing that a finite set S together with a collection of its subsets is today often called a **hypergraph**.) Hall originally formulated his famous result as follows.

1.4.1. THEOREM. *A collection of sets $\{S_1, \dots, S_n\}$ has an* SDR *if and only if for every $k, 0 \le k \le n$, the union of any k of the sets $S_1, \dots, S_n$ has cardinality at least k.*

PROOF. The left to right implication is trivial. Suppose, conversely, that for every k, $0 \le k \le n$, and for every subset $\{i_1, \dots, i_k\}$ of $\{1, \dots, n\}$, we have $|\bigcup_{j=1}^{k} S_{i_j}| \ge k$. Then form a bipartite graph $G = (A, B)$ as follows. Let the points of A be the sets $S_1, \dots, S_n$ and the points of B be the points of $\bigcup_{i=1}^{n} S_i$. Join point S_i of A to point s_j of B if and only if $s_j \in S_i$. Now choose any $A' \subseteq A$. By definition of G we have $|\Gamma(A')| = |\bigcup_{S_{i_j} \in A'} S_{i_j}| \ge |A'|$. But then by Theorem 1.1.3, A can be matched into B and it follows that the endpoints in B of this matching must be an SDR. ■

1.4.2. EXERCISE. Complete the demonstration of the equivalence of Theorems 1.4.1 and 1.1.3 by showing that Theorem 1.4.1 implies Theorem 1.1.3.

There are an astonishing number of variations and generalizations of P. Hall's Theorem, many of which can be proved by reduction to one of the two basic versions, Theorems 1.1.3 and 1.4.1. We shy from the abyss and direct the interested (and industrious!) reader to the books of Mirsky (1971) and Welsh (1976) as well as the survey article of Brualdi (1975).

However, let us give one example from a class of such results known as "marginal element versions". (If an SDR is required to contain a prescribed set M of elements, these elements are called **marginal**.)

This next result, although simple to state and prove, serves well as a signpost for our purposes. In one direction, it points to linear programming (see Chapter 7), for marginal element versions of P. Hall's Theorem were first dealt with by Hoffman and Kuhn (1956a, 1956b) using LP techniques. In another direction it points to matroids. Indeed this result, due in its present form to Mendelsohn and Dulmage (1958), has an elegant proof due to Kundu and Lawler (1973) who in the same paper present a matroid generalization.

1.4.3. EXERCISE. Let $G=(A,B)$ be a bipartite graph. Suppose $S \subseteq A, T \subseteq B$ and that there is a matching of S into B and one of T into A. Show that there is then a matching in G covering both S and T.

We can further exploit the connection between set systems and matroids. Let $\mathcal{F} = \{S_1, \ldots, S_n\}$ be a family of (not necessarily distinct) subsets of a finite set S. **A partial SDR (PSDR)** is a subset of S which is an SDR for some subfamily of $\mathcal{F}$.

1.4.4. LEMMA. *The partial* SDR*'s of $\mathcal{F}$ form a matroid of S.*

PROOF. By the same construction as in the proof of Theorem 1.4.1 we obtain the matroid in Example 1.3.15. (These matroids are called **transversal** in Box 1C. In fact, SDR's themselves are often called **transversals**, and hence the name of these matroids.)

This construction and Theorem 1.3.1 also gives us the rank function of this matroid: for $X \subseteq S$,

$$r(X) = n - \max_{\mathcal{G} \subseteq \mathcal{F}} \{|\mathcal{G}| - |\bigcup_{A \in \mathcal{G}} A \cap X|\}. \tag{1.4.1}$$

∎

We can also apply the Matroid Intersection Theorem to get a necessary and sufficient condition for the existence of a so-called "common system of representatives". Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two families of sets, $|\mathcal{F}_1| =$

$|\mathcal{F}_2| = n$. A set U of n elements is a **common system of distinct representatives (CSDR)** for $\mathcal{F}_1$ and $\mathcal{F}_2$ (or perhaps more aptly a **simultaneous system of distinct representatives** for $\mathcal{F}_1$ and $\mathcal{F}_2$) if U is an SDR for $\mathcal{F}_1$ and for $\mathcal{F}_2$. For example, let $\mathcal{F}_1 = \{\{1,2\},\{2,3,4\},\{4,5\}\}$, $\mathcal{F}_2 = \{\{1,3,4\},\{2\},\{5\}\}$ and $\mathcal{F}_3 = \{\{1,2,3,5\},\{4\},\{1,4\}\}$. Then $\mathcal{F}_1$ and $\mathcal{F}_2$ have a CSDR, for example, $\{1,2,5\}$, $\mathcal{F}_1$ and $\mathcal{F}_3$ have a CSDR, for example $\{1,4,5\}$, but $\mathcal{F}_2$ and $\mathcal{F}_3$ have no CSDR.

The following theorem was first proved by Ford and Fulkerson (1958, 1962) who used the techniques of network flows. (See Chapter 2.)

1.4.5. THEOREM. *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two families of subsets of a finite set S and suppose that $|\mathcal{F}_1| = |\mathcal{F}_2| = n$. Then $\mathcal{F}_1$ and $\mathcal{F}_2$ have a* CSDR *if and only if for any two subfamilies $\mathcal{G}_1 \subseteq \mathcal{F}_1$ and $\mathcal{G}_2 \subseteq \mathcal{F}_2$, we have*

$$|(\bigcup_{A\in\mathcal{G}_1} A)\cap(\bigcup_{A\in\mathcal{G}_2} A)| \geq |\mathcal{G}_1| + |\mathcal{G}_2| - n.$$

PROOF. Let (S, r_i) be the transversal matroid formed by the partial SDR's of $\mathcal{F}_i$ $(i = 1,2)$. Then a CSDR is just a common independent set of these two matroids of size n. By the Matroid Intersection Theorem 1.3.17, such a common independent set exists if and only if

$$r_1(X) + r_2(S - X) \geq n \tag{1.4.2}$$

for all $X \subseteq S$.

Substitute for $r_1(X)$ and $r_2(S - X)$ from equation (1.4.1). Then inequality (1.4.2) is equivalent to

$$n \geq \max_{G_1\subseteq\mathcal{F}_1} \{|G_1| - |\cup_{A\in G_1} A\cap X|\} + \max_{\mathcal{G}_2\subseteq\mathcal{F}_2} \{|\mathcal{G}_2| - |\cup_{A\in\mathcal{G}_2} A\cap(S - X)|\}.$$

So $\mathcal{F}_1$ and $\mathcal{F}_2$ have a CSDR if and only if for all $X \subseteq S$, $\mathcal{G}_1 \subseteq \mathcal{F}_1$ and $\mathcal{G}_2 \subseteq \mathcal{F}_2$, we have

$$|(\cup_{A\in\mathcal{G}_1} A)\cap X| + |(\cup_{A\in\mathcal{G}_2} A)\cap(S - X)| \geq |\mathcal{G}_1| + |\mathcal{G}_2| - n.$$

The set X which minimizes the left hand side is clearly $X = S - \cup_{A\in\mathcal{G}_1} A$. For this choice of X, the left hand side is just $|(\cup_{A\in\mathcal{G}_1} A)\cap(\cup_{A\in\mathcal{G}_2} A)|$, which proves the theorem. ∎

We remark in passing that it is easy to construct an example of three families of sets which have no CSDR, but any two of which do have a CSDR. The theory of the existence of CSDR's for three or more families of sets will be ignored here for the very good reason that it does not exist! In fact, the problem of finding a CSDR for three (or more) families of sets belongs to the "hardest" combinatorial problems, the so-called NP-complete problems (see Box 6A).

Let us give an application of Theorem 1.3.8 to hypergraphs which is analogous to P. Hall's result on systems of distinct representives. Let H be a hypergraph. We say that H has a **forest-representative system**, if it is possible to choose a pair of points from every line of H such that the pairs chosen form the lines of a forest; that is, if we join each such pair with a line, we get a forest.

1.4.6. COROLLARY. *A hypergraph H has a forest-representative system, if and only if for every $k > 0$ the union of any k lines of H has cardinality more than k.*

PROOF. The proof is straightforward by the same construction as in the proof of 1.4.1. (Use Theorem 1.3.8.) ∎

1.4.7. EXERCISE. A hypergraph H is called **2-colorable**, if it is possible to 2-color the points so that no line is monochromatic. Prove that if H is a hypergraph such that for every $k > 0$ the union of any k lines of H has cardinality $> k$, then H is 2-colorable.

One of the most fundamental results in the theory of partially ordered sets is due to Dilworth (1950). We now proceed to show that this theorem too is equivalent to König's Theorem. Let $P = \{a_1, \ldots, a_n\}$ be any finite set. Any relation $R \subseteq P \times P$ is said to be a **partial ordering of** P if R is reflexive, anti-symmetric and transitive. A set P endowed with such a relation is called a **partially ordered set** or **poset**. We shall denote R by "$\leq$"; that is, $(a_i, a_j) \in R$ if and only if $a_i \leq a_j$. This notation will be unambiguous for our purposes since we shall have no need to change a partial order in any of our discussions to follow. We shall also write $a_i < a_j$ if and only if $a_i \leq a_j$ and $a_i \neq a_j$. A subset $\{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$ of P is **totally ordered** or is a **chain** if there is some permutation π of the subscripts $i_1, \ldots, i_k$ such that $a_{\pi(i_1)} \leq a_{\pi(i_2)} \leq \cdots \leq a_{\pi(i_k)}$. Any two elements a_i and a_j are called **comparable** if they belong to some chain and are **incomparable** otherwise. Any set of pair-wise incomparable elements is called an **antichain**.

A collection $\mathcal{P} = \{C_1, C_2, \ldots, C_r\}$ of chains [antichains] in P is called a **chain [antichain] partition of** P if the C_i's are mutually disjoint and their union is P. Observe that any partially ordered set can be represented as a union of disjoint chains, for the one-element chains

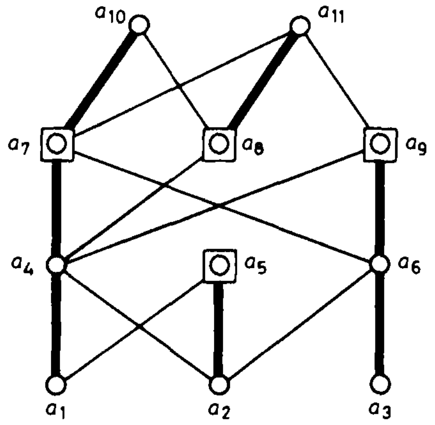


FIGURE 1.4.1.

$\{a_1\}, \{a_2\}, \ldots, \{a_n\}$ trivially suffice. Thus it makes sense to talk about chain partitions of P of smallest cardinality. On the other hand, since the cardinality of an antichain cannot exceed $|P|$, we may speak of antichains of largest cardinality. One of the fundamental results of the theory of partially ordered sets relates these two concepts and is due to Dilworth (1950).

1.4.8. THEOREM. (*Dilworth's Theorem*). *In any finite partially ordered set the cardinality of any largest antichain equals the cardinality of any smallest chain partition.*

In Figure 1.4.1 we illustrate Dilworth's Theorem as applied to an eleven-element poset. (As is customary in the diagram of a poset, we suppress all lines implied by transitivity and loops implied by reflexivity.) One easily sees in this example that the cardinality of a largest antichain = the cardinality of a smallest chain partition = 4.

The set $\{a_5, a_7, a_8, a_9\}$ (shown in boxes) is a largest antichain and the "heavy" paths correspond to the four disjoint chains $C_1 = \{a_1, a_4, a_7, a_{10}\}$, $C_2 = \{a_2, a_5\}$, $C_3 = \{a_8, a_{11}\}$ and $C_4 = \{a_3, a_6, a_7\}$.

Although not even couched in the language of bipartite graphs, it is interesting that Dilworth's Theorem is equivalent to König's Theorem. To show this we follow the treatment of Fulkerson (1956) and Ford and Fulkerson (1962). The most important idea in this proof is how to build a suitable bigraph model when given the poset.

We shall prove Dilworth's Theorem, assuming König's Theorem to be true. The converse will be left as an exercise for the reader. Let $P = \{x_1, \ldots, x_n\}$ denote a poset. Build a bigraph $G = (A, B)$ where

$A = \{a_1, a_2, \ldots, a_n\}$, $B = \{b_1, b_2, \ldots, b_n\}$ and $a_i b_j \in E(G)$ if and only if $x_i < x_j$. We shall need the following two lemmas.

1.4.9. LEMMA. *Let M be a set of independent lines in G. Then there is a chain partition $\mathcal{P}$ of P such that $|M| + |\mathcal{P}| = n$.*

PROOF. Suppose $M = \{a_{i_1} b_{i_2}, a_{i_3} b_{i_4}, \ldots, a_{i_{2k-1}} b_{i_{2k}}\}$. So $x_{i_1} < x_{i_2}, x_{i_3} < x_{i_4}, \ldots, x_{i_{2k-1}} < x_{i_{2k}}$ and among the distinct elements of these $2k$ x_{i_j}'s we have a set $\mathcal{P}'$ of chains each of length at least two and they are pairwise disjoint since M is independent. We extend $\mathcal{P}'$ to a chain partition $\mathcal{P}$ of P in a trivial way, namely by appending all one-element chains not already present in $\mathcal{P}'$. Now let $n_j =$ the number of elements in the jth chain of $\mathcal{P}$. We then have

$$n = \sum_{j=1}^{|\mathcal{P}|} n_j = \sum_{j=1}^{|\mathcal{P}|} (n_j - 1) + |\mathcal{P}| = |M| + |\mathcal{P}|$$

since there are precisely $n_j - 1$ lines in the jth chain of $\mathcal{P}$. ■

1.4.10. LEMMA. *If $C_G \subseteq A \cup B$ is a minimum cover of bigraph $G = (A, B)$ of poset P, then there is an antichain U contained in P with $|C_G| + |U| \geq n$.*

PROOF. Let $C_G = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}, b_{j_1}, b_{j_2}, \ldots, b_{j_m}\}$ be a minimum cover of G.

Clearly C_G corresponds to a subset Q of P with $|Q| \leq |C_G|$. Let $U = P - Q$. Then the elements of U are pairwise unrelated in P since C_G is a cover in G. Moreover, by construction,

$$|C_G| + |U| \geq |Q| + |U| = |Q| + |P - Q| = |P| = n.$$

■

We are now prepared to prove Dilworth's Theorem.

PROOF (of Theorem 1.4.8). Let P be any partially ordered set and $G = (A, B)$, the bigraph corresponding to P. Let L_G be a maximum independent set of lines in G and $C_G \subseteq A \cup B$, a minimum cover of G. By Lemma 1.4.9 there is a chain partition $\mathcal{P}$ of P such that $|L_G| + |\mathcal{P}| = n$. By Lemma 1.4.10 there is an antichain $U \subseteq P$ with $|C_G| + |U| \geq n$. But by König's Minimax Theorem, $|L_G| = |C_G|$ and hence $|\mathcal{P}| \leq |U|$.

On the other hand, it is an elementary observation to note that for any arbitrary chain partition $\mathcal{P}'$ of P and any arbitrary antichain U' of P, we must have $|U'| \leq |\mathcal{P}'|$ since no two elements of U' can lie in the same chain of $\mathcal{P}'$. Thus $|\mathcal{P}| = |U|$ and hence it follows that $\mathcal{P}$ is a minimum chain partition of P and U, a maximum antichain. ■

Note that this proof of Dilworth's Theorem also yields a polynomial algorithm to find a minimum chain partition and a maximal antichain in a poset.

1.4.11. EXERCISE. Prove König's Theorem, assuming Dilworth's Theorem to be true.

Dilworth's theorem has a "dual", which is much easier to prove and which is included here for the sake of later reference when studying perfect graphs in Chapter 12.

1.4.12. THEOREM. *In a finite partially ordered set the cardinality of any longest chain equals the cardinality of any smallest antichain partition.*

PROOF. It is clear that every antichain partition contains at least as many antichains as the cardinality m of a longest chain. It suffices to construct an antichain partition containing m antichains.

For each $a \in P$, let $\phi(a)$ denote the maximum length of a chain with smallest element a. Obviously, $1 \leq \phi(a) \leq m$. Further, $a < b$ implies that $\phi(a) > \phi(b)$. Hence the sets $A_i = \{a \in P \mid \phi(a) = i\}$, for $i = 1, \ldots, m$, are antichains and thus they form a partition of the desired type. ∎

As we have already pointed out in our historical notes on early theorems related to bipartite matching, most were couched in terms of matrices. The next theorem due to König (1916a, 1916b) is such a result. We present it not only for its own intrinsic interest, but also because it yields as an immediate corollary a more recent and more well-known result on doubly stochastic matrices. A non-negative $n \times n$ matrix A is **doubly stochastic** if all row sums and all column sums are equal to 1. A **permutation matrix** P is a square matrix with exactly one "1" in each row and in each column (the rest of the entries being zero). A **permutation set** of a square matrix is a set of entries containing one entry from each row and from each column.

1.4.13. THEOREM. *Every doubly stochastic matrix A has a permutation set of non-zero entries.*

PROOF. Define a bigraph $G = (R, C)$ with $R = \{r_1, \ldots, r_n\}$ the set of rows of A and $C = \{c_1, \ldots, c_n\}$ the set of columns of A. Then define (r_i, c_j) to be a line of G if and only if entry $a_{ij} \neq 0$ in A.

We proceed to show G has a perfect matching. Suppose not. Then by P. Hall's Theorem there is a set of k rows in R which have non-zero entries only in some s columns with $s < k$. Now summing the entries in these i rows row by row we get k, but summing the same entries column by column, we get a sum not exceeding s and hence $k \leq s$, a contradiction. So G has a perfect matching as claimed. ■

A famous result due independently to Birkhoff (1946) and von Neumann (1953) now follows easily. This result has also been called the "Birkhoff-König Theorem" — and sometimes just the "Birkhoff Theorem" — elsewhere in the literature.

1.4.14. COROLLARY. (*The Birkhoff–von Neumann Theorem*). *If A is doubly stochastic then A can be written as a convex combination of permutation matrices; i.e., $A = c_1P_1 + \cdots + c_kP_k$ where $\sum_{i=1}^{k} c_i = 1$, $P_1, \ldots, P_k$ are permutation matrices and each $c_i \geq 0$.* ■

We shall prove the following, clearly equivalent, form of this result.

1.4.15. COROLLARY. *If A is a non-negative matrix with column sums and row sums all equal to a common value, then A can be written as a non-negative linear combination of permutation matrices.*

PROOF. By Theorem 1.4.13, A has a permutation set with all entries positive. Let P_1 denote the corresponding permutation matrix and let c_1 be the minimum entry in this permutation set. Then $A - c_1P_1$ is non-negative and has equal row and column sums. Denote this "remainder" matrix by R_1. We can clearly repeat this argument on R_1 and in some k steps we can write $A = c_1P_1 + \cdots + c_kP_k$ where each P_i is a permutation matrix. It is easy to see that this decomposition terminates after a finite number of steps (i.e., k is finite), for in each iteration the number of zeros in R_i increases by at least one. Moreover, since there are no more than n^2 non-zero entries in A to begin with, clearly $k \leq n^2 - n + 1$. ■

The procedure outlined in the above proof for finding the permutation matrices to express the doubly stochastic matrix has come to be called the "Birkhoff Algorithm". It is interesting to note that the upper bound on the number of permutation matrices — $n^2 - n + 1$ — obtained in the proof is not best possible. The reader may note at this point that the set of doubly stochastic matrices is just the convex hull of permutation matrices when all are considered as points in n^2 dimensional space. (Cf. Chapter 7.) Approaching from this polyhedral point of view, Hoffman and Wielandt (1953) devised a different proof of Corollary 1.4.13. Their proof shows that, in fact, we have the better upper bound of $n^2 - 2n + 2$ for the number of permutation matrices necessary to express the doubly stochastic matrix. On the other hand, the Hoffman-Wielandt proof is not constructive. So for a time it remained to determine just how good the Birkhoff Algorithm really was. Johnson, Dulmage and Mendelsohn (1960) showed that Birkhoff's Algorithm actually gives the decomposition using no more than $n^2 - 2n + 2$ permutation matrices — the Hoffman and Wielandt bound. For details of this work, as well as information on other results involving this upper bound, the reader is referred to Mirsky (1963) and to Brualdi (1982).

The Birkhoff–von Neumann Theorem relates to another result of König (1916a, 1916b).

1.4.16. LEMMA. *If G is a regular bipartite graph, then G has a perfect matching.*

PROOF. Immediate from Theorem 1.1.3. ■

This innocent-sounding result can lead one immediately in the direction of line coloration, another area of considerable research in graph theory. A **valid line coloration** of any graph (bipartite or not) is an assignment of integers (the "colors") to the lines of G in such a way that no two lines having the same color are adjacent. The smallest number of colors required to produce a valid line coloration is called the **chromatic index** (or **line chromatic number**) of G and is denoted $\chi_e(G)$. Of course each color class of lines in a valid line coloration is, in particular, a matching.

Let $\Delta(G)$ denote the maximum degree of the graph G. It is immediate that

$$\chi_e(G) \geq \Delta(G). \tag{1.4.3}$$

1.4.17. LEMMA. *If G is a bipartite graph regular of degree Δ, then $\chi_e(G) = \Delta$.*

PROOF. This is obvious from Lemma 1.4.16 by successively deleting Δ perfect matchings. ■

Actually, this lemma may be used to obtain the following more general result due to König (1916a, 1916b) which gives a good characterization for the chromatic index of a bigraph. We shall discuss the problem of chromatic index of non-bipartite graphs in Section 7.4.

1.4.18. THEOREM. *(König's Line Coloring Theorem). For every bipartite graph G, $\chi_e(G) = \Delta(G)$.*

PROOF. We shall embed G into a bipartite graph G_1 which is regular of degree $\Delta = \Delta(G)$. Then the result will follow from Lemma 1.4.17 and inequality (1.4.3).

First, by adding isolated points if necessary, we embed G into a bigraph G' such that $\Delta(G') = \Delta$ and such that the bipartition (A', B') of G' has $|A'| = |B'|$. Now connect A' to B' by new lines so that we do not increase the maximum degree. Let G_1 be the maximal graph so obtained. We claim that G_1 is regular of degree Δ. Assume indirectly that G_1 is not Δ-regular. Then there is a point $a_0 \in A'$ with $\deg_{G_1}(a_0) < \Delta$. Then

$$\Delta \cdot |A'| > \sum_{a \in A'} \deg_{G_1}(a) = \sum_{b \in B'} \deg_{G_1}(b)$$

and so there exists a point $b_0 \in B'$ with $\Delta_{G_1}(b_0) < \Delta$. We may then join a_0 and b_0 with a new line without increasing the maximum degree, a contradiction. ∎

Let us remark that Lemma 1.4.17 can also be formulated as follows.

1.4.19. COROLLARY. *If A is a non-negative integral matrix with equal row and column sums, then A can be written as a sum of permutation matrices.*

The similarity to Corollary 1.4.15 is apparent. In fact, Corollary 1.4.15 can be deduced from Corollary 1.4.19. (No such deduction in the other direction is known. (Cf. Section 12.4.))

1.4.20. EXERCISE. Deduce Corollary 1.4.15 from Corollary 1.4.19.

At this point we hasten to point out that the relationship between matching theory and linear algebra is much deeper and more extensive than the above connection between bigraphs and doubly stochastic matrices. This will become more and more apparent as we proceed. In particular, the reader is referred to Chapter 8.

We mentioned earlier that the Marriage Theorem is an example of a so-called "self-refining result". König's Line Coloring Theorem is yet another. Consider the following two corollaries.

1.4.21. COROLLARY. (*de Werra* (1971, 1975)). *Let G be a bigraph and let k be any integer ≥ 1. Then G is the union of k line-disjoint spanning subgraphs $G_1, \ldots, G_k$ such that for each $v \in V(G)$:*

$$\left\lfloor \frac{\deg(v)}{k} \right\rfloor \leq \deg_{G_i}(v) \leq \left\lceil \frac{\deg(v)}{k} \right\rceil.$$

PROOF. We build a new graph G' by splitting each point v of G into $\lfloor \deg(v)/k \rfloor$ points of degree k and, if necessary, one more point of degree $\deg(v) - k\lfloor \deg(v)/k \rfloor$. (The partition of the lines incident with v in G is arbitrary with respect to which lines go with each new point; only the cardinality of the bundle at each new point is important.)

Note that G' remains bipartite and it has maximum degree k. But then by Theorem 1.4.18, $E(G')$ is the union of k line-disjoint matchings $M_1, \ldots, M_k$. Now "collapse" G' back to G by identifying the points of G' which correspond to each point of G. For each $i = 1, \ldots, k$ let M_i' denote the subgraph of G resulting from M_i after collapsing. Since M_i is a matching there is at most one line of M_i incident with each of the $\lceil \deg(v)/k \rceil$ points of G' corresponding to $v \in V(G)$; i.e., $\deg_{M_i'}(v) \leq \lceil \deg(v)/k \rceil$.

On the other hand, there are $\lfloor \deg(v)/k \rfloor$ points of G' corresponding to v in G each of which has degree k. But then each of these has — for each i — one line of M_i incident with it. Hence

$$\deg_{M_i'}(v) \geq \left\lfloor \frac{\deg(v)}{k} \right\rfloor$$

and the theorem is proved. ∎

Thus König's Line Coloring Theorem implies the above result. On the other hand, let $k = \Delta$, the maximum degree of G. Then the inequality of the preceding corollary becomes

$$0 \leq \left\lfloor \frac{\deg(v)}{\Delta} \right\rfloor \leq \deg_{G_i}(v) \leq \left\lceil \frac{\deg(v)}{\Delta} \right\rceil = 1$$

and thus each G_i consists of disjoint lines and isolates. But this just means that G is the union of Δ disjoint matchings; that is, $\chi_e(G) = \Delta$ and we have König's Line Coloring Theorem again! Hence our remark about "self-refining".

de Werra's result, in turn, implies the following result due to Gupta (1967) on decompositions of bipartite graphs into disjoint line covers. (Recall that a **line cover** is any set of lines which collectively touch each point of G.)

1.4.22. COROLLARY. *Let G be a bipartite graph with minimum degree r. Then G is the union of r line-disjoint line covers.*

PROOF. Let $k = r$ in Corollary 1.4.21. Then $E(G)$ is the union $G_1 \cup G_2 \cup \cdots \cup G_r$ of line-disjoint subgraphs such that for each $v \in V(G)$,

$$\left\lfloor \frac{\deg(v)}{r} \right\rfloor \leq \deg_{G_i}(v) \leq \left\lceil \frac{\deg(v)}{r} \right\rceil.$$

But by definition of r, $\lfloor \deg(v)/r \rfloor \geq 1$ and hence $\deg_{G_i}(v) \geq 1$ for all $i = 1, \ldots, r$. ∎

It is difficult to decide where to stop in any treatment of the theorems of König and P. Hall. Entire books can be written on their ramifications. (In fact, one has! See Mirsky (1971)). This is not our goal, however, and so we call a halt at this point and direct the reader to the references already given.

2

Flow Theory

2.0. Introduction

Flow theory is best understood as a model for transportation problems. Suppose we have a pipeline network and would like to provide a steady flow of oil between two specified locations. Various sections of pipeline have various upper bounds on the amount of oil which can flow per unit time (the so-called "capacities"). Another quite natural demand says that no amount of oil may be bought and stored, nor may any be drained off, at any intermediate point of the network. In other words, the amount of oil entering a pipeline network junction must be equal to the amount leaving the junction. This property is generally called "conservation of flow". In other real-world situations, we find different terminology; for example, in electrical networks this conservation property is called "Kirchhoff's Current Law". The archetypal "flow problem" is then to assign some amount of oil (perhaps none) to each of the lines of the network consistent with these constraints in such a way that a maximum amount is delivered from the source of supply to the terminal point.

It takes precious little imagination to see that this so-called "max-flow" problem can be used to model a host of practical situations. In addition to many applications to "two terminal" networks, as exemplified by the oil shipment problem described above, other less obviously related problems can be reduced to this archetypal form. For example, we can compute vulnerability in a communications network, solve multiterminal (single-commodity) flow problems, and decide whether there is any flow at all in a network which has minimum (as well as maximum) capacities on its lines.

The max-flow problem and its efficient algorithmic solution form one of the true historical cornerstones of the new and thriving branch of mathematics called "combinatorial optimization". Indeed, since the solution of the max-flow problem in the 1950's, flow theory has swiftly proceeded to become one of the most well-developed areas of this young discipline.

But why include the study of flows in a book on matching? Because, as we shall soon see, flow theory often provides rather quick proofs to some of the basic theorems in matching theory when the matching problem at hand is suitably reformulated. This is especially true in the case of bipartite matching.

In Section 2.1 we formulate and prove the celebrated Max-Flow Min-Cut Theorem. It is our second primary example — after König's Minimax Theorem in Chapter 1 — from the class of results known as "minimax theorems". The reader will note our intention to emphasize such results. (See Box 1A and elsewhere in this book.) In fact, the Max-Flow Min-Cut Theorem is in a sense equivalent to König's Minimax Theorem.

In Section 2.2 we present the essentials of several efficient maximum flow algorithms. Although these algorithms have been further "fine-tuned" by clever manipulations of data structures, we shall not make any excursions in this direction, but prefer to direct the interested reader to the relevant references.

In Section 2.3 we treat a result — not as widely known as it should be — which says roughly that all the information about flows in a graph can be stored effectively in a certain tree polynomially derivable from the parent graph. This is the so-called "flow-equivalent tree". We shall use such trees later in Chapter 6.

In Section 2.4 we apply flow theory to obtain Menger's theorems on the connectivity of graphs and digraphs, and to obtain several degree-constrained subgraph theorems for bigraphs and digraphs.

In the final section of this chapter we show how some of the results obtained earlier can be used in mathematical analysis as we steal quietly into the camp of the measure theorist!

Altogether in this chapter only a few results from flow theory are treated, namely those which we need for the matching results mentioned. Other important concepts such as minimum cost flows, multiterminal flows, multicommodity flows and the "Out-of-Kilter Method" will not be covered. For excellent coverage of these topics the reader is urged to consult the books of Ford and Fulkerson (1962), Hu (1969) and Lawler (1976).

2.1. The Max-Flow Min-Cut Theorem

Our main objective in this section is to state and prove the fundamental theorem on network flows called variously the "Ford-Fulkerson" or "Max-Flow Min-Cut" theorem. An important point regarding this theorem which is sometimes glossed over, is that proving existence of a flow of maximum possible value and actually exhibiting such a flow (by a finite number of iterations of a suitable algorithm) are two different problems in general. Only in cases where all capacities are integral will the algorithm, as often presented, necessarily terminate in a maximum flow after a finite number of iterations.

Let D be a digraph with point set V and two distinguished points s (the **source**) and t (the **sink**). Let $c : E(D) \to \Re_+$ be a function which associates with each line (u, v) of D a non-negative real number $c(u, v)$ called the **capacity** of the line. The quadruple (D, c, s, t) will be called a **network**.

Any function $f : E(D) \to \Re_+$ is called a **flow** in (D, c, s, t) (or briefly, in D) if it satisfies the following properties:

(i) $\sum_u f(u, v) = \sum_w f(v, w)$ for all $v, w \in V(D) - \{s, t\}$ (conservation of flow) and,

(ii) $f(u, v) \leq c(u, v)$ for all (u, v) in $E(D)$.

It is sometimes convenient to extend the functions f and c to all of $V(D) \times V(D)$, by writing $f(u, v) = c(u, v) = 0$, if both (u, v) and $(v, u) \notin E(D)$.

Problems in which each line has a "minimum" capacity as well as a maximum capacity — although important — will not be useful for our purposes and hence will not be treated. The interested reader may consult Even (1979) or Liu (1968).

$\sum_u f(s, u) - \sum_u f(u, s)$ defines the **value** of the flow f, denoted by $\mathrm{val}(f)$. It is easy to see that this value is equal to $\sum_u f(u, t) - \sum_u f(t, u)$.

If A is a set of points in a digraph D, let $\nabla^+(A)$ denote the set of all lines joining points of A to points of $V(D) - A$ (and directed out of A). $\nabla^+(A)$ will be called the **directed cut out of** A. An **s–t separator** (or just a **separator** when s and t are understood) in D is any set $A \subseteq V(D)$ such that $s \in A$, but $t \in V(D) - A$. The **s–t cut** determined by separator A (or by cut $\nabla^+(A)$) is the set of lines $\nabla^+(A) \subseteq E(D)$. The **capacity** of cut $\nabla^+(A)$ or of separator A, $\mathrm{cap}(A) = \sum_{u \in A, v \in V(D) - A} c(u, v)$. It may be helpful to point out immediately that since D is finite there will be only a finite number of cuts in D and hence the existence of cuts of minimum capacity (so-called **minimum cuts** or **min-cuts**) is clear.

That flows of maximum value (so-called **maximum flows** or **max-flows**) exist is not so immediate and the existence of an algorithm to determine one is even less transparent.

2.1.1. LEMMA. *In every network (D, c, s, t) a maximum flow exists.*

PROOF. We may view each flow as a point in the q-dimensional space $\Re^q$, where q is the number of lines in D. Observe that the set of flows is a compact set of points in $\Re^q$ and also that the function $\text{val}(f)$ is a continuous — in fact, linear — function of q variables. Hence it attains its maximum on the set by the Weierstrass Theorem. ■

2.1.2. LEMMA. *If f is any flow in D and C is any $s-t$ cut, then* $\text{val}(f) \le \text{cap}(C)$.

PROOF. Let f and $C = \nabla^+(A)$ denote an arbitrary $s-t$ flow and an $s-t$ cut in D respectively. Then

$$\begin{aligned}
\text{val}(f) &= \sum_u f(s,u) - \sum_u f(u,s) \\
&= \sum_u f(s,u) - \sum_u f(u,s) + \sum_{a\in A-s} \Big(\sum_w f(a,w) - \sum_v f(v,a)\Big) \\
&= \sum_{a\in A} \Big(\sum_w f(a,w) - \sum_v f(v,a)\Big) \\
&= \sum_{a\in A}\sum_w f(a,w) - \sum_{a\in A}\sum_v f(v,a) \\
&= \Big(\sum_{\substack{a\in A\\ w\in A}} f(a,w) + \sum_{\substack{a\in A\\ w\in V-A}} f(a,w)\Big) - \Big(\sum_{\substack{a\in A\\ v\in A}} f(v,a) + \sum_{\substack{a\in A\\ v\in V-A}} f(v,a)\Big)
\end{aligned}$$

Noting that the first and third terms cancel we have

$$\text{val}(f) = \sum_{\substack{a\in A\\ w\in V-A}} f(a,w) - \sum_{\substack{a\in A\\ v\in V-A}} f(v,a).$$

But by definition of flow, $\sum_{a\in A, v\in V-A} f(v,a) \ge 0$, so

$$\text{val}(f) \le \sum_{\substack{a\in A\\ w\in V-A}} f(a,w) \le \sum_{\substack{a\in A\\ w\in V-A}} c(a,w) \le \text{cap}(A).$$

■

Note that equality holds if and only if $f(a,v) = c(a,v)$ and $f(v,a) = 0$ for every $a \in A$ and $v \in V-A$. In this case f must be a maximum flow and C, a minimum cut. But note that we do not yet know whether a flow and a cut exist for which equality holds in Lemma 2.1.2.

Let f be any flow in D. A path $P = u_0u_1\cdots u_k$ (in the underlying undirected graph G) is an **f-augmenting path to u_k** if

(a) $u_0 = s$,

and for each line $l = u_iu_{i+1}$ of P we have

(b) $f(u_i, u_{i+1}) < c(u_i, u_{i+1})$ when l is directed from u_i to u_{i+1} in D and

(c) $f(u_{i+1}, u_i) > 0$ when $l = u_iu_{i+1}$ is directed from u_{i+1} to u_i.

If $u_k = t$ we simply call P an **f-augmenting path.**

2.1.3. LEMMA. *A flow f is maximum if and only if there are no f-augmenting paths.*

PROOF. Let $A_f = \{u_k \in D \mid \exists$ an f-augmenting path P to $u_k\} \cup \{s\}$. If $t \in A_f$, this means there is an f-augmenting path P. In this case, let $\epsilon_1(P) = \min\{c(u,v) - f(u,v) \mid uv$ is a line of P and (u,v) is directed away from s in $D\}$, let $\epsilon_2(P) = \min\{f(u,v) \mid uv$ is a line of P and (u,v) is directed toward s in $D\}$, and finally let $\epsilon(P) = \min\{\epsilon_1(P), \epsilon_2(P)\}$. Since $t \in A_f$, $\epsilon(P) > 0$ and there is a flow f' with $\mathrm{val}(f') = \mathrm{val}(f) + \epsilon(P) > \mathrm{val}(f)$ obtained by increasing f by $\epsilon(P)$ along all lines of P directed away from s and decreasing f by the same amount on all lines of P directed toward s. Thus when $t \in A_f$, f is not maximum.

On the other hand, suppose $t \notin A_f$. Thus A_f is a separator in D and for each line (u,v), $u \in A_f$, $v \in \bar{A}_f$, $f(u,v) = c(u,v)$, whereas for every line (u,v), $u \in \bar{A}_f$, $v \in A_f$, $c(u,v) = 0$. Hence as remarked after Lemma 2.1.2, $\mathrm{val}(f) = c(A)$ and f is a maximum flow. ■

We are now prepared to prove the following result, proved by Ford and Fulkerson (1956) and independently by Elias, Feinstein and Shannon (1956).

2.1.4. THEOREM. (*Max-Flow Min-Cut Theorem*). *If D is a digraph with source s and sink t, the maximum value of any $s-t$ flow equals the minimum capacity of any $s-t$ cut.*

PROOF. Let f be any maximum flow (f must exist by Lemma 2.1.1). Let γ be the minimum capacity over all $s-t$ cuts. Since f is maximum there is no f-augmenting path and so as we saw in the proof of Lemma 2.1.3, $\mathrm{val}(f) = \mathrm{cap}(A_f) \geq \gamma$. Since $\mathrm{val}(f) \leq \gamma$ by Lemma 2.1.2, the theorem is proved. ■

We shall discuss some of the applications of the Max-Flow Min-Cut Theorem in Section 2.4. Let us formulate here just one particularly simple corollary, which has the following interpretation in economics.

Assume that there are k suppliers $s_1, \ldots, s_k$ of some commodity, and in m consumers $t_1, \ldots, t_m$. It is known which suppliers can ship to which consumers, and also the supply $a(s_i) \geq 0$ of supplier s_i and the demand $b(t_j) \geq 0$ of consumer t_j. Under what conditions can the suppliers satisfy the consumers?

2.1.5. COROLLARY. (*The Supply-Demand Theorem*). *The following condition is necessary and sufficient for the satisfiability of the consumers:*

For every subset T of consumers,

$$\sum_{t\in T} b(t) \leq \sum_{s\in \Gamma(T)} a(t),$$

where $\Gamma(T)$ denotes the set of suppliers servicing at least one consumer in T. ■

This result is part of a family of theorems collectively called "supply-demand theorems" (see Ford-Fulkerson (1962, Chapter 2)).

2.2. Flow Algorithms

Note that we have proved the existence of a maximum flow by a non-constructive argument. But can we effectively compute such a flow? The answer, fortunately, is "yes"! We proceed to present an algorithm to do this. The algorithm is based upon the important observation already made that a flow f is maximum if and only if there are no f-augmenting paths.

At first sight, in order to produce a maximum flow it seems quite reasonable to proceed as follows: begin with any flow f_0 (the zero flow will do nicely) and look for f_0-augmenting paths. If there are none, we know f_0 is maximum and we are finished. If, on the other hand, we find such an augmenting path P, then augment along P to get a flow f_1 of value larger than that of f_0. Now look for f_1-augmenting paths and repeat this process. Terminate if f_k is a maximum flow. We will call this procedure the Ford-Fulkerson Algorithm.

Consider first the case when all capacities are integral. Start with flow $f_0 = 0$. Notice that the amount by which we augment is integral and hence f_1 is integral on every line. Continuing, we get a sequence of flows $f_0, f_1, f_2, \ldots$ each of which is integral on every line. So the sequence $\mathrm{val}(f_0) < \mathrm{val}(f_1) < \mathrm{val}(f_2) < \cdots$ consists of integers and by Lemma 2.1.2, it is bounded. So it must terminate with a flow f_k for which no f_k-augmenting path exists. Thus f_k is a maximum flow.

This argument has an important consequence.

2.2.1. THEOREM. (*Flow Integrality Theorem*). *If the capacities of a network are integers, then there exists a maximum flow which is integral on every line.* ■

It also follows that if all capacities are integers then the algorithm described above terminates in a finite number of steps (the number of

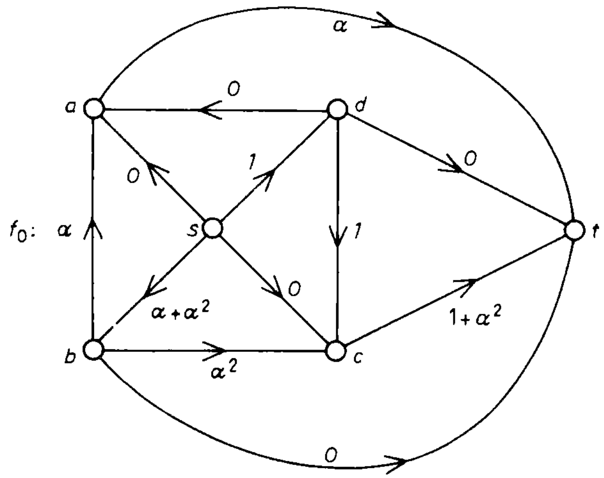


FIGURE 2.2.1.

augmentations is no larger than the value of a maximum flow). It is clear that the same algorithm will terminate if the capacities are rational, since we may "scale" by multiplying all capacities by their least common denominator (see, however, Box 2A). It is something of a shock, however, to learn that for general real capacities the algorithm described above may not produce a maximum flow at all! In fact, $\mathrm{val}(f_k)$ may tend to a number much less than the optimum value of a flow.

Another strange phenomenon is that if we do not start with the zero flow, then even if all capacities are infinite, the sequence $\mathrm{val}(f_i)$ may converge to a finite value. We now describe an example of this.

2.2.2. EXAMPLE. Consider the network shown in Figure 2.2.1.

The capacities on all lines are infinite (or if the reader prefers, say, 1000). Let α denote the positive root of $x^3 + x - 1 = 0$. (Clearly $1/2 < \alpha < 1$.) Let the initial flow f_0 be as shown in Figure 2.2.1. Note that the flow values on the lines from s and those to t will always increase during any augmentation procedure and they really need never be considered.

We claim that by employing a diabolically chosen sequence of four augmenting paths over and over, we can find an infinite sequence of flows, the values of which are monotone increasing and which converge to a limiting value not exceeding 16! In particular, for any $m \geq 0$ let flow f_{4m} be as shown in Figure 2.2.2. Augment along path $scdabt$ by an amount α^{4m+1} to produce flow f_{4m+1}, augment along path $scbadt$ by an amount α^{4m+2} to produce f_{4m+2}, augment along $sabcdt$ by α^{4m+3} to get f_{4m+3}, and finally augment along $sadcbt$ by α^{4m+4} to get f_{4m+4}. We show the first of these four augmentations in Figure 2.2.2.

Continue this sequence of augmentations. Then for $r \geq 1$, the augmentation from f_{r-1} to f_r is α^r and hence

$$\begin{aligned}
\mathrm{val}(f_r) &= \mathrm{val}(f_0) + \alpha + \alpha^2 + \cdots + \alpha^r \\
&= (1+\alpha+\alpha^2) + \alpha + \alpha^2 + \cdots + \alpha^r \\
&= (1/\alpha)(\alpha + \alpha^2 + \alpha^3 + \alpha^2 + \alpha^3 + \cdots + \alpha^{r+1}) \\
&= (1/\alpha)(\alpha + \alpha^2 + (1-\alpha) + \alpha^2 + \alpha^3 + \cdots + \alpha^{r+1}) \\
&= (1/\alpha)(1 + 2\alpha^2 + \alpha^3 + \cdots + \alpha^{r+1}) \\
&< (1/\alpha)(1 + \alpha + \alpha^2 + \alpha^3 + \cdots + \alpha^{r+1}) \\
&< \frac{1}{\alpha - \alpha^2} = \alpha^{-4} < 16.
\end{aligned}$$

From this construction it is also easy to get an example of a network where some capacities are irrational such that even starting with $f_0 = 0$, the value of the flow f_r is less than, say, 16 while the value of a maximum flow is arbitrarily large. We leave this construction to the reader as an exercise.

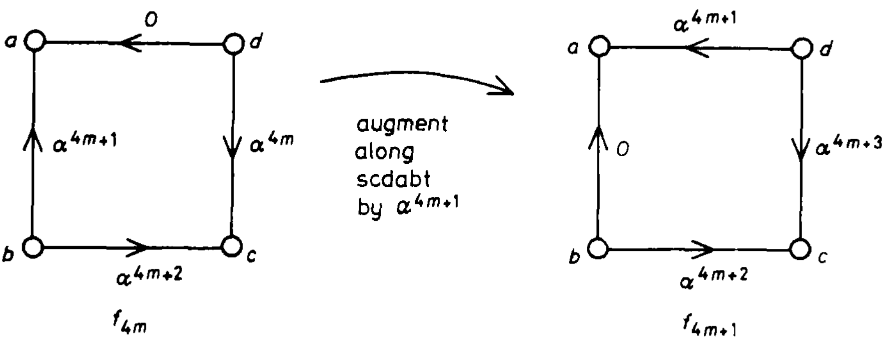


FIGURE 2.2.2.

Fortunately, all is not lost! Edmonds and Karp (1970, 1972) and Dinic (1970) have found flow algorithms which terminate for arbitrary capacities. (We may well ask why bother since in practice all capacities will have to be represented as rational numbers anyway. See Box 2A for a discussion of this.) Of these two algorithms, the one due to Edmonds-Karp is closer to the original idea and we discuss it first. Their algorithm is based upon the following theorem.

2.2.3. THEOREM. *If at each step, flow f_{i+1} is obtained from f_i by augmenting along an augmenting path with a minimum number of lines, then the algorithm terminates with a maximum flow in less than $p^3/2$ steps.*

But let us defer the proof for a moment. How can we find a *shortest* f_i-augmenting path? Fortunately, there is a simple procedure to do this. In fact, if we design an algorithm to find *any* f_i-augmenting path, we are likely to find a shortest one! So quite some time before the work of Edmonds-Karp, flow algorithms were running which were convergent (in fact, polynomial) but their polynomiality had not been verified, since the significance of the fact that they picked a shortest augmenting path had not been noticed.

The procedure for finding a shortest augmenting path is a version of a method called "breadth-first search".

Given a network D and a flow f, define a new directed graph D_1 on $V(D)$ by connecting u to v if and only if

(1) $(u,v) \in E(D)$ and $f(u,v) < c(u,v)$, or

(2) $(v,u) \in E(D)$ and $f(v,u) > 0$.

Observe now that an f-augmenting path in D corresponds to a directed $s-t$ path in D_1 and vice versa. So the task of finding a shortest f-augmenting path in D is tantamount to finding a shortest directed $s-t$ path in D_1.

There is a very natural and easy algorithm for finding a shortest directed $s-t$ path in digraph D_1. Label s. In the next step, label all points which are accessible from s on a line (i.e., on a directed path of length one). Also mark one line for each point labelled, said line directed from s to this labelled point. In the i^{th} step, label all points which are not yet labelled and which are accessible from a labelled point via a line and also mark such a line for every newly labelled point. When t gets labelled, stop the labelling procedure. The marked lines form a tree which contains a unique $s-t$ path. The proof of the simple and well-known fact that this path is an $s-t$ path of minimum length is left to the reader.

Before confronting the proof of Theorem 2.2.3, we need several lemmas.

2.2.4. LEMMA. *Let D be a digraph with distinguished points s and t such that*

(i) $\deg^+(v) = \deg^-(v)$ *for all* $v \in V(D) - \{s,t\}$, *and*

(ii) $\deg^+(s) - \deg^-(s) = k > 0$.

Then D contains k line-disjoint $s-t$ dipaths.

PROOF. First we show that there exists at least one $s-t$ dipath. To this end, let X be the set of all points accessible from s on a dipath and assume that $t \notin X$. Then in the subgraph D' induced by X,

$$\deg^+_{D'}(v) \geq \deg^-_{D'}(v) \text{ for all } v \neq s, \text{ and}$$
$$\deg^+_{D'}(s) > \deg^-_{D'}(s).$$

But this is a contradiction since the sum of outdegrees in the graph must be equal to the sum of indegrees.

So let P be an $s-t$ dipath. Deleting the lines of P from D, we are left with a digraph D_1 which satisfies the conditions of the hypotheses of this lemma with k replaced by $k-1$. So by induction on k we find $k-1$ line-disjoint $s-t$ dipaths in D_1. Together with P, these form k line-disjoint $s-t$ dipaths in D. ∎

2.2.5. LEMMA. *Suppose D is a network with distinguished points s and t, initial $s-t$ flow f_0, and suppose $P_1, P_2, \ldots, P_i, \ldots$ is a sequence of shortest augmenting $s-t$ paths. In particular, suppose P_i is f_{i-1}-augmenting, for all i, $1 \leq i \leq m$. Then for each i, $|E(P_i)| \leq |E(P_j)|$ for all $j > i$.*

PROOF. Consider any k for which shortest augmenting paths P_k and P_{k+1} exist. We form a new digraph H from $P_k \cup P_{k+1}$ as follows: reorient all lines of P_k and P_{k+1} toward t and form the union H of the resulting dipaths $\hat{P}_k$ and $\hat{P}_{k+1}$, but remove those pairs (u,v) and (v,u) for which $(u,v) \in \hat{P}_k$ and $(v,u) \in \hat{P}_{k+1}$. Those common lines of $\hat{P}_k$ and $\hat{P}_{k+1}$ which are passed in the same sense (in $\hat{P}_k \cup \hat{P}_{k+1}$) are, however, taken twice in H. Clearly then, H has equal in and outdegrees at all points except s and t. (Of course H is not simple in general.)

We claim there are two line-disjoint directed $s-t$ paths Q'_1 and Q'_2 in H. But this is clear upon recalling that $\deg^+(s) = 2$ in H and applying the preceding lemma.

Now let Q_1 and Q_2 be the $s-t$ paths in D corresponding to Q'_1 and Q'_2 in H respectively. Then, $E(Q_1) \cup E(Q_2) \subseteq E(P_k) \cup E(P_{k+1})$ and, when thought of as a path from s to t, each line of Q_1 is traversed in the same direction as it is when traversed along $P_k \cup P_{k+1}$ from s to t. Similarly for Q_2. Moreover, $E(Q_1) \cap E(Q_2) \subseteq E(P_k) \cap E(P_{k+1})$.

We now claim that Q_1 and Q_2 are both f_{k-1}-augmenting paths. We need only show this for one of the Q_i's, say Q_1. Let $e \in E(Q_1)$. First suppose e is traversed along Q_1 from tail to head. If $e \in P_k$, then since both Q_1 and P_k traverse e from tail to head, and since P_k is f_{k-1}-augmenting, $f_{k-1}(e) < c(e)$. If $e \in P_{k+1} - P_k$, then $f_k(e) < c(e)$, but

$f_{k-1}(e) = f_k(e)$ since $e \notin P_k$, so $f_{k-1}(e) < c(e)$. The argument is similar if e is traversed along Q_1 from head to tail.

Thus each of Q_1 and Q_2 is f_{k-1}-augmenting and each, therefore, is considered along with P_k for a shortest augmenting path in the k^{th} step of the algorithm. But by the minimality of $|E(P_k)|$, $|E(P_k)| \leq |E(Q_i)|$, for $i = 1, 2$. Thus

$$\begin{aligned}
2|E(P_k)| &\leq |E(Q_1)| + |E(Q_2)| \\
&= |E(Q_1) \cup E(Q_2)| + |E(Q_1) \cap E(Q_2)| \\
&\leq |E(P_k) \cup E(P_{k+1})| + |E(P_k) \cap E(P_{k+1})| \\
&= |E(P_k)| + |E(P_{k+1})|
\end{aligned}$$

and therefore, $|E(P_k)| \leq |E(P_{k+1})|$ and the proof of the lemma is complete. ∎

REMARK. Note that if $\hat{P}_k$ and $\hat{P}_{k+1}$ traverse some line x in opposite directions, then x will not be in $E(Q_1) \cup E(Q_2)$ and moreover we have $|E(Q_1)| + |E(Q_2)| = |E(Q_1')| + |E(Q_2')| = |E(P_k)| + |E(P_{k+1})| - 2$ and hence $2|E(P_k)| \leq |E(P_k)| + |E(P_{k+1})| - 2$, or $|E(P_k)| \leq |E(P_{k+1})| - 2$.

2.2.6. LEMMA. *Let $P_1, \ldots, P_k, \ldots, P_{k+l}, \ldots$ be a sequence of shortest augmenting paths, $l \geq 1$ and suppose $\hat{P}_k$ and $\hat{P}_{k+l}$ use a line x in opposite directions. Then $|E(P_{k+l})| \geq |E(P_k)| + 2$.*

PROOF. We know from Lemma 2.2.5 that $|E(P_k)| \leq |E(P_{k+1})| \leq \cdots \leq |E(P_{k+l})|$. Thus it suffices to prove the desired inequality when P_{k+l} and P_k use a line oppositely, but no pair P_i, P_j use a line oppositely for any i, j such that $k \leq i < j \leq k + l$ and $j - 1 < l$.

As in Lemma 2.2.5 form a new digraph H from $P_k \cup P_{k+l}$ by orienting all lines toward t, forming the union of the resulting dipaths $\hat{P}_k \cup \hat{P}_{k+l}$, taking twice any common lines of $\hat{P}_k$ and $\hat{P}_{k+l}$ passed in the same sense from s to t, but removing those common lines which are passed in the opposite sense. We know one such pair exists by hypothesis. Just as before, let Q_1 and Q_2 be two line-disjoint paths from s to t in D where $E(Q_1) \cup E(Q_2) \subseteq E(P_k) \cup E(P_{k+l})$.

It remains to show that each Q_i is f_{k-1}-augmenting. We show this for Q_1. Let $e \in E(Q_1)$ and first suppose e is traversed along Q_1 from tail to head. If $e \in P_k$ and since Q_1 and P_k traverse e in the same direction, $f_{k-1}(e) < f_k(e)$ by definition of P_k and hence $f_{k-1}(e) < c(e)$. If $e \in P_{k+l} - P_k$ we know e appears in none of $P_{k+1}, \ldots, P_{k+l-1}$, so it follows that $f_{k-1}(e) = f_k(e) = \cdots = f_{k+l-1}(e) < c(e)$ by definition of P_{k+l}

and thus $f_{k-1}(e) < c(e)$. The argument is similar if e is traversed tail to head along Q_1.

So each Q_i is f_{k-1}-augmenting and hence each is considered together with P_k for the choice of a shortest f_{k-1}-augmenting path. So $|E(Q_i)| \geq |E(P_k)|$, $i = 1, 2$. Moreover, by assumption there is at least one line used by $\hat{P}_k$ and $\hat{P}_{k+l}$ in opposite directions, so as in the preceding remark, we have $2|E(P_k)| \leq |E(Q_1)| + |E(Q_2)| \leq |E(P_k)| + |E(P_{k+l})| - 2$ and thus $|E(P_k)| \leq |E(P_{k+l})| - 2$ and the proof of Lemma 2.2.6 is complete. ∎

Let P be any f-augmenting path. When the flow f is increased along P by amount $\epsilon(P)$ there will be at least one line x which becomes "saturated"; that is, if x is directed toward t, $f(x) + \epsilon(P) = c(x)$ or if x is directed away from t, $f(x) - \epsilon(P) = 0$. Any such line x is called a **bottleneck** relative to P.

We are now prepared to prove Theorem 2.2.3.

PROOF (of Theorem 2.2.3). Let $P_1, P_2, ..$ be any sequence of shortest augmenting paths. Note that each P_i contains at least one bottleneck line. Let x be any line. How many different times, that is, in how many different P_i's, can x be a bottleneck? Let $P_{i_1}, P_{i_2}, \ldots$ be the subsequence of all P_i's in each of which x is a bottleneck. The crucial observation is that as we traverse successive augmenting paths in this subsequence, since the flow value in x oscillates between 0 and $c(x)$, it follows that the direction of traversal of the P_{i_j}'s oscillates between "with" the direction of x in D and "against" this direction. But it then follows that we successively oscillate between "with" and "against" in the corresponding $\hat{P}_{i_j}$'s. But then by Lemma 2.2.6, $1 \leq |E(P_{i_1})| \leq |E(P_{i_2})| - 2 \leq \cdots \leq |E(P_{i_j})| - 2(j-1) \leq \cdots$. Hence for all j, $1 \leq |E(P_{i_j})| - 2(j-1)$ or $2j - 1 \leq |E(P_{i_j})|$.

On the other hand, for each j, $|E(P_{i_j})| \leq p - 1$. Hence $2j - 1 \leq p - 1$ or $j \leq p/2$, so at most $p/2$ of the P_i's can contain x as a bottleneck. There are $q \leq p(p-1)$ different lines in D, and hence there are $\leq qp/2 \leq (p^3 - p^2)/2$ P_i's in our original sequence of augmenting paths $P_1, P_2, \ldots$. This proves Theorem 2.2.3. ∎

Note that an important feature of the bound on the number of steps in the Edmonds-Karp algorithm obtained in Theorem 2.2.3 is that *it is independent of the capacities.* This fact is important even in the case when the capacities are integers, since it may be that the Ford-Fulkerson Algorithm (with a very unlucky choice of the augmenting paths) runs much longer than the Edmonds-Karp version.

2.2.7. EXAMPLE. Let G be the graph in Figure 2.2.3, with the capacities as marked, N being a large positive integer. Starting with the 0 flow and augmenting alternatingly along the paths *sabt* and *sbat*, we arrive at a maximum flow after $2N$ iterations. On the other hand, the Edmonds-Karp procedure takes at most $4^3/2 = 32$ steps by Theorem 2.2.3 (in fact, it takes only two steps) regardless of the value of N.

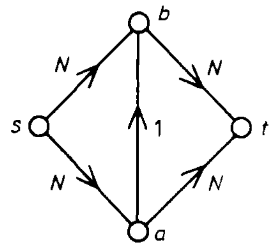


FIGURE 2.2.3.

Several interesting questions have not been addressed yet. For example, although Edmonds and Karp always select a *shortest* augmenting path, we may adopt other criteria for augmenting path selection. Two alternatives which have been studied are:

(1) select an augmenting path from among those with a minimum number of "reverse" lines, i.e., lines directed toward s or

(2) select a path which gives a largest augmentation.

For discussion of each of these, see Edmonds-Karp (1970, 1972).

BOX 2A. Searching a Graph

In this section we have described a procedure to find a shortest augmenting path. In fact, if we get rid of the unnecessary details, then it turns out that this algorithm finds a shortest path from a given source to every other point in the auxiliary graph obtained from network D by taking all directed lines (u, v) with $f(u, v) < c(u, v)$, as well as the reverse of all directed lines (u, v) with $f(u, v) > 0$.

To find a path between two given points in a graph or digraph has been an extremely important algorithmic problem ever since Theseus almost became lost in the Labyrinth of the Minotaur. Several strategies had already been proposed for Theseus before the notion of an algorithm was made precise (see the algorithm of Trémaux in Lucas (1882) and the solution in Tarry (1895)). To make this task, as well as the differences among the algorithms, more precise, let us define a **search of a digraph** D **from root** r as the following algorithm. We construct an **arborescence** A **rooted at** r (i.e., a spanning tree which is oriented so that every point can be reached from r) as follows. We start with $V(A) = \{r\}$ and $E(A) = \emptyset$. If we already have a certain A, then we choose any line connecting $V(A)$ to $V(D) - V(A)$ and add this line to A. Stop if no such line exists.

It is easy to see that we end up with an arborescence A which contains exactly those points of D which can be reached from r on a directed path. In this sense any search of the graph D reaches all points it can possibly reach.

We have considerable freedom in choosing the line to be added to A at each step. By being more specific, we get special versions of the general search algorithm which have various important additional properties. To discuss these, let us first rephrase the search algorithm as follows. Let the points of the current arborescence be called **labelled**. While searching for a line going from a labelled point to an unlabelled one, we may already have convinced ourselves that some labelled point is not connected to any of the unlabelled points by a directed line. Let us call such a labelled point **scanned**. Note that if a point is scanned at a given step, then no line directed out of it will be added to the arborescence at any later step either, and so it remains scanned. We can now describe the search procedure as follows.

(1) To begin, let r be labelled, and no point scanned.
(2) Choose any point which is labelled but not scanned. If no such point exists, stop. Otherwise, let v be such a point and go to (3).
(3) Choose any line which connects v to any unlabelled point. If no such line exists, call v "scanned" and return to (2). Otherwise, let u be the head of this line. Call u "labelled" and return to (2).

We obtain two important special cases of this algorithm by restricting the choice of v in (2) in two different ways. These special selection rules are often called the "first labelled first scanned" and the "last labelled first scanned" rules. In the first version, we choose in step (2) the point v (labelled but not scanned), which was labelled first. The procedure obtained this way is called **breadth-first search**, because the arborescence A obtained this way grows in height (or depth, depending upon how you prefer to draw arborescences!) only if it cannot grow in breadth any more.

The breadth-first search arborescence A has a number of important properties. Most significant of these is the fact that for every point v of A, the unique path in A from r to v is a shortest rv-path in the whole digraph D. Thus breadth-first search provides a shortest path from the root to every other point.

Another important property of the breadth-first search arborescence is that no line of $E(D) - E(A)$ goes from any point of A to any of its "descendants". In the case of undirected graphs (which from the point of view of search procedures may be viewed as directed graphs with the

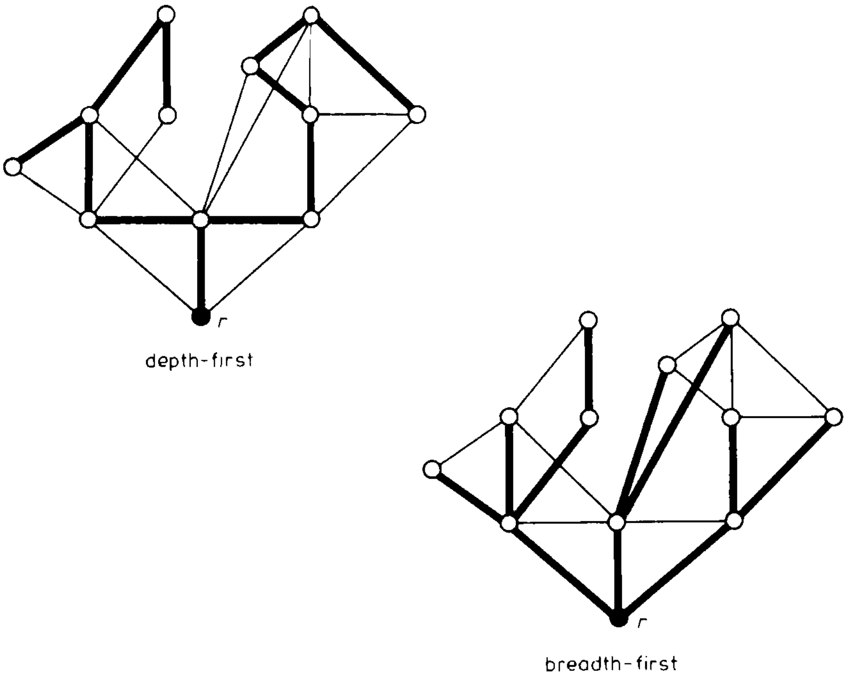


FIGURE 2A.1. The two basic types of tree search

property that every line is bi-directed), the breadth-first search tree has the property that every line of the undirected graph G which is not a line of A connects two points which belong to different branches of A.

The other important selection rule for (2) is obtained if we choose the point v which was labelled *last.* This version is called **depth-first search**. The arborescence A obtained by depth-first search has the property that every line of D connects a point to one of its descendants, predecessors or to a point of a branch which was grown earlier. In the special case when the graph is undirected this last possibility does not occur; every line of the undirected graph G connects two points of the tree, one of which is the predecessor of the other. While we do not explicitly use depth-first search in this book, it is as an algorithmic tool just as important as breadth-first search. It is widely used, for example, in algorithms for finding the 2-connected components of a graph, testing a graph for planarity, etc. See Tarjan (1972), Hopcroft and Tarjan (1973b, 1974), Lempel, Even and Cederbaum (1967), Booth and Lueker (1976), Even and Tarjan (1975, 1976) and Chapters 3 and 8 of Even (1979).

Figure 2A.1 provides an example of each of the two types of search in a graph starting from the same root point r.

The result, then, of Edmonds and Karp can also be formulated as follows:

If in the maximum flow algorithm the augmenting path is found by breadth-first search, then the algorithm terminates with a maximum flow in $O(p^3)$ augmentations.

Another interesting problem is encountered when we not only estimate the number of augmentations needed to get a maximum flow, but also the time bound for producing an augmentation. If we call the product of these two time bounds "total time", then it turns out that Edmonds-Karp will terminate in $O(p^3 \cdot p^2) = O(p^5)$ total time.

A more involved algorithm, the idea of which is due to Dinic (1970) and which has been improved by Karzanov (1974), gives a better total time bound of $O(p^3)$. The main idea is to make more efficient use of the breadth-first search labelling procedure, which we applied to construct an augmenting path, and construct an "augmenting flow" instead.

Let $(D, c, s, t) = (D, c)$ be a network as before. First let us agree to construct a "reasonably good" flow (rather than a maximum flow). In fact, assume that we have an algorithm to construct this "reasonably good" flow in every network. (For the moment we shall not worry about what "reasonably good" means precisely!) Then we may carry out the following procedure. Start with the flow $f_0 = 0$. After some number of iterations, assume that we have a flow f_i. Define a new network (D_i, c_i), called the i^{th} **layered network** by Even (1976, 1979), where $V(D_i) = V(D)$ and $(u, v) \in E(D_i)$ if either $(u, v) \in E(D)$ and $f_i(u, v) < c(u, v)$ or $(v, u) \in E(D)$ and $f_i(v, u) > 0$ (this is the same construction used above to find a shortest augmenting path). Also define $c_i(u, v) = c(u, v) - f_i(u, v)$ or $c_i(u, v) = f_i(v, u)$ according to which of these two cases holds. Find a "reasonably good" flow Φ_i in the network (D_i, c_i). (If no $s-t$ flow exists in (D_i, c_i), stop.) Use this flow Φ_i to augment f_i as follows:

$$f_{i+1}(u, v) = f_i(u, v) + \Phi_i(u, v) \text{ if } f_i(u, v) = 0,$$
$$f_{i+1}(u, v) = f_i(u, v) - \Phi_i(v, u) \text{ if } f_i(u, v) = c(u, v), \text{ and}$$
$$f_{i+1}(u, v) = f_i(u, v) + \Phi_i(u, v) - \Phi_i(v, u) \text{ if } 0 < f_i(u, v) < c(u, v).$$

It is easy to see that f_{i+1} is a flow in (D, c) and that

$$\text{val}(f_{i+1}) = \text{val}(f_i) + \text{val}(\Phi_i).$$

So we get a sequence $f_1, f_2, \ldots$ of flows in (D, c) with increasing values.

The problem is to interpret the notion of "reasonably good" in such a way that:

(1) it should be easy to find a "reasonably good" flow in any network in which a non-trivial flow exists at all, and
(2) the sequence $f_0, f_1, \ldots$ should terminate after a small number of iterations with a maximum flow.

The reader may note at this point that the Ford-Fulkerson Algorithm is a special case of the above procedure when "reasonably good" means a flow along an $s-t$ dipath which saturates at least one line of the path. The Edmonds-Karp algorithm corresponds to the special case when "reasonably good" means a flow along a shortest $s-t$ dipath. Dinic (1970) introduced a more efficient notion of "reasonably good". Let a flow f be called a **greedy flow** if every line e with $f(e) > 0$ belongs to a shortest $s-t$ dipath and every shortest $s-t$ dipath contains a line e with $f(e) = c(e)$. Thus a flow f is "greedy" in two ways: every atom of the transported material travels as short a route as possible and the flow is maxim*al* (although not necessarily maxim*um*) with respect to this property.

2.2.8. LEMMA. *If Φ_i is a greedy flow in (D_i, c_i) then the length of a shortest $s-t$ dipath is strictly longer in D_{i+1} than in D_i.*

PROOF. The proof is left to the reader. (See Even (1976, 1979)). ■

Using this lemma, it follows immediately that in the sequence $f_0, f_1, f_2, \ldots$ a maximum flow is reached after at most $p-1$ iterations. So property (2) is nicely satisfied.

To find a greedy flow Dinic gave an algorithm which takes $O(pq)$ steps. The algorithm is a version of depth-first search and we will not go into the details. Karzanov improved this by finding a greedy flow in $O(p^2)$ steps. For the details the reader is referred to Karzanov (1974) and to Even (1976, 1979).

A number of further improvements in the total time bound have been obtained since the formulation of the algorithms discussed above. These improvements are due to ever more clever data manipulation schemes rather than fundamental improvements in the algorithm itself and hence we will be satisfied to direct the interested reader to the chronological table of results and accompanying references below.

So what is the best we can hope for in total time for a flow algorithm? Galil (1978) remarks that $O(pq)$ is perhaps attainable, but he conjectures that fundamentally new ideas will be necessary. By way of contrast, Lawler (1976) has shown, using linear programming techniques that, in fact, a maximum flow can always be obtained from an arbitrary starting

Reference	Total Time Bound
1. Ford and Fulkerson (1956)	—
2. Edmonds and Karp (1970, 1972)	pq^2
3. Dinic (1970)	p^2q
4. Karzanov (1974)	p^3
5. Čerkasskiĭ(1977)	$p^2\sqrt{q}$
6. Malhotra, et. al.* (1978)	p^3
7. Galil (1978, 1980)	$p^{5/3}q^{2/3}$
8. Shiloach (1978)	$pq\log^2 p$
9. Galil and Naamad (1979, 1980)	$pq\log^2 p$
10. Sleator (1980)	$pq\log p$
Sleator and Tarjan (1983)	

*(Does not improve the bounds of algorithms 4, 5 or 7, but is simpler.)

TABLE 2.2.1.

flow in no more than q steps, but unfortunately no algorithm is known to find this optimal sequence of augmentations!

Finally, let us return for a moment to the example above where the Ford-Fulkerson flow algorithm did not terminate. What precisely was so "diabolical" about the choice of augmenting paths? The problem turns out to hinge upon the fact that in successive iterations labelled points were scanned in different orders. Let us call a labelling procedure consistent if whenever we have a set of labelled, but unscanned, points we always pick the points to be scanned in the same order. A. Tucker (1977) has proved that if a consistent labelling procedure is used with the Ford-Fulkerson flow algorithm, the algorithm will terminate in a maximum flow. Fulkerson himself (unpublished) formulated the following conjecture.

2.2.9. CONJECTURE. *Any consistent labelling procedure results in a maximum flow in a polynomial number of steps.*

BOX 2B. Numbers in Algorithms

Numbers are basic objects and tools of study throughout mathematics and of course they arise in many combinatorial problems as they have in this chapter on capacities and flows. It can be rather difficult, however, to handle numbers from the point of view of the theory of algorithms.

Trouble begins to appear even if the numbers in question are integers. For example, consider the case when all flow capacities are integers. We have pointed out that the Ford-Fulkerson Algorithm, starting with the 0 flow, always yields a maximum flow in at most t iterations, where t is the maximum value of a flow. Is this algorithm polynomial? Since trivially $t \leq \sum_{u,v} c(u,v)$, it is certainly polynomial if the capacities are all 1's (and in combinatorial applications this is often the case). But do we still consider it polynomial if the $c(u,v)$ are large?

The crux of the difficulty here is the definition of input size, when the input contains integers. There are at least three, non-equivalent ways to handle this problem.

(1) We may say that a natural number t contributes t to the input size. In this case we say that the problem is in **unary encoding**.

(2) We may say that a natural number t contributes $1 + \lfloor \log_2 t \rfloor$ to the input size (this is the number of digits in the binary expansion of t). In this case we say that the problem is in **binary encoding**. (Considering bases other than 2 would make no essential difference here.)

(3) We may say that a natural number t contributes 1 to the input size. In this case we say that the problem is in **arithmetic encoding**.

To each of these notions of input size there corresponds a natural way to define the number of steps of an algorithm. In cases (1) and (2) every arithmetic operation must be carried out bitwise, and so the multiplication of two k-digit numbers by the usual procedure will take about $2k^2$ steps. In case (3), however, it is better to count every arithmetic operation (addition, subtraction, multiplication, comparison) as only one step. Note that these arithmetic operations are polynomial-time in any of these three encodings. (This is not the case with exponentiation, since there the size of the output is not polynomially bounded).

Since the length of input in binary encoding is not larger than its size in unary encoding, it may happen that a certain algorithm is polynomial in unary encoding, but not polynomial in binary. For example, we have seen that the Ford-Fulkerson Algorithm is polynomial in unary encoding (the number of iterations is bounded by the problem size and one iteration is easily seen to be polynomial). However, Example 2.2.7 shows that it is not polynomial in the binary encoding. On the other hand, an important conceptual development was initiated by an observation of Edmonds and Karp, who proved that their algorithm is polynomial *even in the binary sense*. Similarly the Dinic Algorithm (and its improvement by Karzanov) is polynomial in the binary sense also.

The Edmonds-Karp and Dinic-Karzanov Algorithms are also polynomial in arithmetic encoding: in this encoding the input size is somewhere between p and p^2 for a graph with p points (depending on the number of lines and on how the graph itself is encoded), and these algorithms take $O(p^4)$ and $O(p^3)$ steps, respectively, if arithmetic operations are counted as one step.

There is no immediate connection between polynomiality of an algorithm in the binary and arithmetic sense. Consider the following trivial algorithm: the input is a finite sequence $S = (1, 1, \ldots, 1)$ of 1's, and the algorithm computes $2^{2^{|S|}}$ as follows: We start with $x = 2$. We square x and cancel the last entry of S. We repeat this until S becomes void. In the arithmetic sense, this algorithm takes $O(|S|)$ steps and so it is polynomial. On the other hand, in the binary sense just to print out $2^{2^{|S|}}$ takes exponential time!

There are many algorithms which are polynomial in the binary sense but not in the arithmetic sense. A trivial example is the Euclidean algorithm to find the greatest common divisor of two integers. The well-known "Out-of-Kilter Method" to find a minimum cost flow also has this feature and indeed it is not known if *any* algorithm for this problem can be polynomial in the arithmetic sense. A more recent example is the algorithm of Hačijan for linear programming. (See Box 7B.)

A further question remains to be clarified here. Since we know that the Ford-Fulkerson Algorithm is polynomial if all capacities are 1, we may try to reduce the general case of integral capacities to this case by replacing a line having capacity c by c parallel lines with capacity 1. This reduction works fine as far as proving theorems is concerned; for example, the Max-Flow Min-Cut Theorem could be obtained from Menger's Theorem this way. However, *algorithmically* it is a polynomial time reduction only for the *unary* encoding of the capacities. Since reductions of this kind will often be used in the sequel, it is important to bear in mind that they are *not* polynomial in the stricter — *binary* — sense.

Now we come to the question of non-integral numbers. Rationals do not cause any trouble, since they may be handled as pairs of integers. But what can we say about an irrational number?

One point of view is that since we cannot store infinite decimals in a computer, we may as well ignore irrational numbers; they will never occur as real-world data. But this is just sweeping dirt under the rug! It does seem that Example 2.2.2 is trying to tell us something. Also, some mathematical models may lead to problems where the numbers stem from the model (not from a physical measurement), and hence they may be genuinely irrational.

It appears that no satisfactory theory of computation with real numbers has been developed so far. One idea is to represent a real number as an algorithm which produces arbitrarily good approximations of it.

In the spirit of polynomiality in binary encoding, we may say that the algorithm is efficient if it takes $O(k^{const})$ steps to get an approximation with an error less than 2^{-k}. Addition, multiplication and subtraction do not cause any trouble, but *comparison* does. If many approximations turn out to be zero, can we conclude that the number is zero?

Thus we must be careful with conditional steps. For example, a step like "If $x > 0$, go to 5; if $x \leq 0$, go to 6" should mean that for some $\epsilon > 0$ appropriately specified, "If $x > \epsilon$, go to 5; if $x < -\epsilon$, go to 6; if $-\epsilon \leq x \leq \epsilon$, go to either 5 or 6".

It is natural that if the output of an algorithm is a real number that this means that the algorithm is able to compute approximations within an arbitrarily prescribed error. For a discussion of some of these ideas and problems see Bishop (1967) and Lovász (1984). It may be shown that suitable modifications of the Edmonds-Karp and Dinic-Karzanov Algorithms will be polynomial in this sense, while the Ford-Fulkerson Algorithm fails even to converge.

There are some real number fields other than the rational field in which exact arithmetical computations can be carried out. For example if α is a real algebraic number of order n, then the numbers in the field $\mathcal{Q}[\alpha]$ can be uniquely written in the form

$$r_0 + r_1\alpha + \cdots + r_{n-1}\alpha^{n-1}, \qquad (r_i \in \mathcal{Q}),$$

and so they can be encoded by sequences $(r_0, \ldots, r_{n-1})$ of rational numbers. In this encoding arithmetic operations can be carried out in polynomial time (n is fixed throughout!).

The Edmonds-Karp and Dinic-Karzanov Algorithms are polynomial for these "exact" real number computations, while the Ford-Fulkerson Algorithm is not even convergent.

2.3. Flow-equivalent Trees

Suppose we are given a capacitated digraph G in which we are interested in determining the maximum flow between *every* pair of points. Of course, we can simply run the flow algorithm $\binom{p}{2}$ times. But we can do better, at least in the case when G is "undirected". (More precisely, when every line of G is bi-directed with the same capacity in both directions.) In this case, Gomory and Hu (1961) have shown that $p-1$ maximum flow computations suffice. It also turns out that there is always a *tree* T with capacities on its lines, with $V(T) = V(G)$ and such that the maximum flow between any two points in T is the same as it is in G. Moreover, we can construct this T and determine all the maximum flow values between pairs of points in total time less than that required to do the $\binom{p}{2}$ computations in the original graph G. In addition, this "flow-equivalent tree T" can be constructed so as to possess certain nice relationships between its minimum cuts and those of the parent graph G.

The actual algorithmic procedure for constructing such a tree is nicely treated in Ford and Fulkerson (1962) and in Hu (1969). Here we shall focus on the existence and properties of the flow-equivalent tree, but our treatment can be turned into a polynomial time algorithm. We shall make use of the flow-equivalent tree in treating the work of Padberg and Rao (1982) in Section 6.6.

Let G be any undirected graph and let $c : E(G) \to \Re_+$ be any assignment of non-negative capacities to the lines of G. Next let T be any tree with $V(T) = V(G)$, but, we emphasize, T need not be a subgraph of G. Nevertheless, any line e of T is a cut of T and hence induces a cut of G. Let us denote this induced cut of G by C_e. Such a tree T is called a **cut-equivalent tree** (relative to G) if for each line $e = xy \in E(T)$, the associated cut C_e induced in G is a minimum capacity $x - y$ cut in G.

Suppose we also have capacities $c' : E(T) \to \Re_+$ assigned to the lines of T. We say that T is a **flow-equivalent tree** if for all $x, y \in V(G)$ a maximum $x - y$ flow in G has the same value as the maximum $x - y$ flow in T. (Note that the maximum $x - y$ flow in T is trivial to determine.) We shall see later that if T is a cut-equivalent tree and we define $c'(e) = \text{cap}(C_e)$, for $e \in E(T)$ then T will also be flow-equivalent. The reader may easily convince himself by an example that the converse is not true.

2.3.1. THEOREM. (*The Cut-equivalent Tree Theorem*). *Any undirected graph G possesses a cut-equivalent tree.*

Before proceeding to the proof of this theorem we need to introduce the idea of two cuts which "cross". Let $\nabla(X)$ and $\nabla(Y)$ ($\emptyset \neq X, Y \neq V(G)$) be two cuts in G. We say these cuts are **crossing** if each of the sets $X \cap Y$, $X - Y$, $Y - X$ and $V(G) - X - Y$ is non-empty.

PROOF (of Theorem 2.3.1). Choose distinct points x_0 and y_0 in G such that among all such pairs, the maximum flow between x_0 and y_0 is as small as possible. Let C_0 be any minimum $x_0 - y_0$ cut in G. We proceed inductively to construct a sequence of non-crossing cuts as follows. Suppose we have found distinct cuts $C_0, C_1, \ldots, C_{i-1}$ which are non-crossing and each C_j is a minimum $x_j - y_j$ cut for some pair of points x_j, y_j. Choose a pair of points x_i, y_i not separated by any of $C_0, \ldots, C_{i-1}$ and such that the maximum flow between x_i and y_i is as small as possible. (If no such pair x_i, y_i exists, then we have a family of $p-1$ non-crossing cuts and we stop.)

Claim 1. Given such a pair of points x_i, y_i there exists a minimum $x_i - y_i$ cut C_i such that C_i does not cross $C_0, C_1, \ldots, C_{i-1}$.

To prove this claim, let C_i^* be a minimum $x_i - y_i$ cut. We may suppose there is a cut C_j, $0 \le j \le i-1$, which crosses C_i^*, or else we are done.

Without loss of generality we may assume $C_i^* = \nabla(X_i^*)$, where $x_i \in X_i^*$ and $y_i \notin X_i^*$. Furthermore, we may also assume that $C_j = \nabla(X_j)$, $x_j \in X_j$, $y_j \notin X_j$ and that one of the five situations portrayed in Figure 2.3.1 must obtain. Let $Y_i^* = V(G) - X_i^*$, $Y_j = V(G) - X_j$.

For example, if as in (a), $x_i, x_j \in X_i^* \cap X_j$, $y_i \in (V(G)-X_i^*) \cap X_j$, $y_j \in (V(G)-X_j) \cap X_i^*$ then let $C_i' = \nabla((V(G)-X_i^*)\cap X_i)$ and $C_j' = \nabla((V(G)-X_j)\cap X_i)$. In the first four cases, by submodularity (cf. Section 1.2) or simply by a straightforward line count, $c(C_i') + c(C_j') \le c(C_i^*) + c(C_j)$. But then by the minimality of cuts C_i^* and C_j, it follows that C_i' is a minimum $x_i - y_i$ cut (and C_j', a minimum $x_j - y_j$ cut). In the fifth case, we have trivially $c(C_i^*) = c(C_j)$ and so by submodularity we obtain again that $C_i')$ is a minimum $x_i - y_i$ cut. So we simply replace C_i^* with C_i'.

Repeat this process, if necessary, for all C_j which cross C_i', uncrossing a pair at a time until the C_i' at hand does not cross $C_0, C_1, \ldots$, or C_{i-1}. Then let $C_i = C_i'$. (Note that uncrossing one pair cannot introduce a new crossing elsewhere.) This completes the proof of Claim 1.

Thus we obtain a family $C_0, \ldots, C_m$ of non-crossing cuts and since there remain no pairs of points unseparated by these cuts, we must have $m = p-2$.

Now define $E(T)$ on point set $V(G)$ as the set of lines uv such that exactly one C_i separates u and v. We proceed to show T to be a tree.

Claim 2. Each cut C_i separates the endpoints of at most one line $uv \in E(T)$.

Suppose, on the contrary, that cut C_i separates pairs u_1v_1 and u_2v_2. Assume $u_1 \ne u_2$. (See Figure 2.3.2.) Now by Claim 1 there must be a cut C_j in the sequence $C_0, \ldots, C_{p-2}$ which separates u_1 and u_2 and, moreover, does not cross C_i. Thus two cuts in sequence $C_0, \ldots, C_{p-2}$ separate u_1 and v_1 and therefore $u_1v_1 \notin E(T)$, a contradiction. This proves Claim 2.

Claim 3. For every C_i there is a pair u, v not separated by any other cut of the family $C_0, \ldots, C_{p-2}$, except C_i.

To see this, fix C_i. We prove by induction on j that for each j, $i+1 \le j \le p-2$, there is a pair of points u and v separated by C_i,

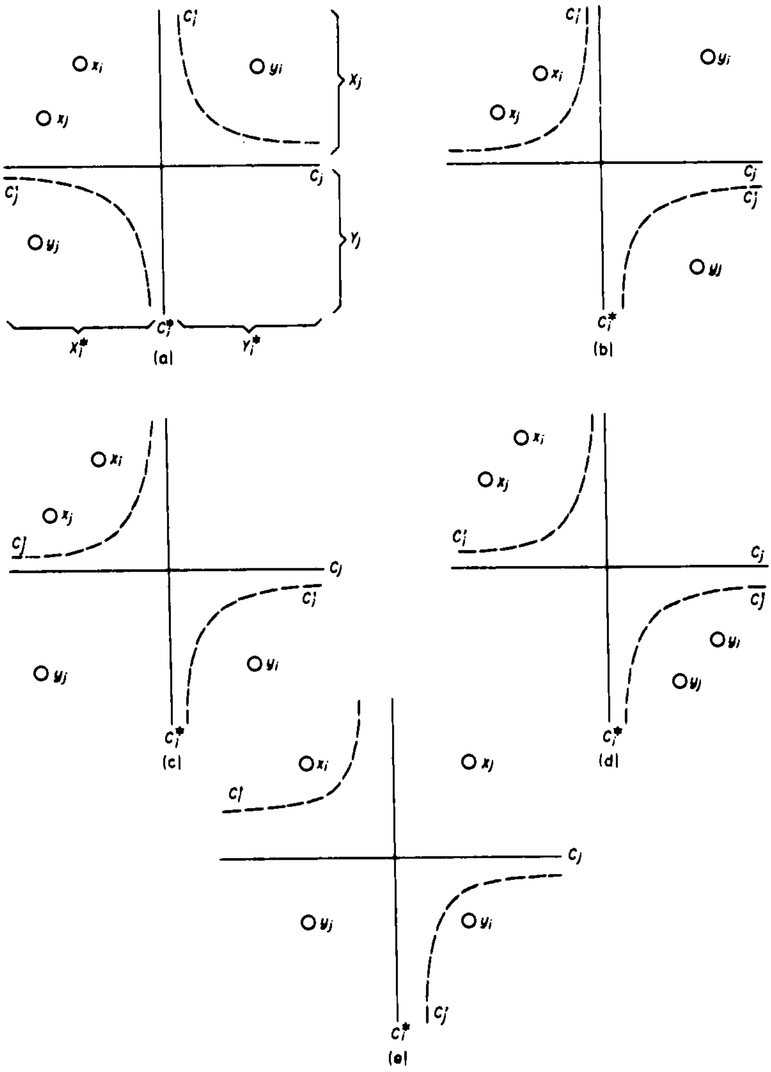


FIGURE 2.3.1. Uncrossing cuts

but not by any of $C_0, \ldots, C_{i-1}, C_{i+1}, \ldots, C_{j-1}$. This is certainly true for $j = i + 1$ when $u = x_i$ and $v = y_i$ satisfy this requirement. So suppose $j \geq i + 2$. If C_j does not separate u and v we are done, so suppose C_j does separate them. Without loss of generality, let us suppose u and x_j are on the same side of C_j. (Cf. Figure 2.3.3.)

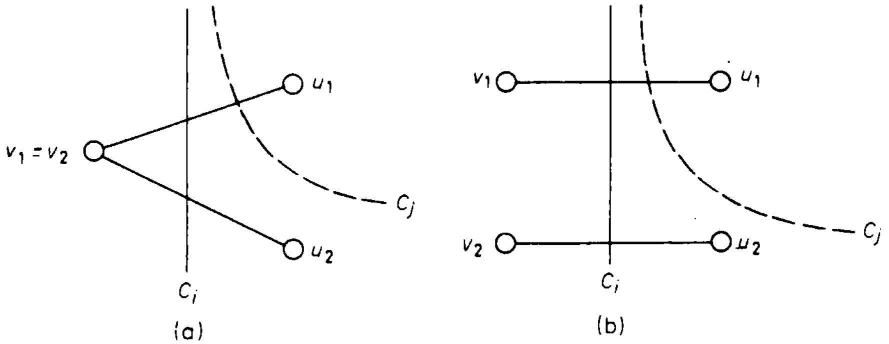


FIGURE 2.3.2.

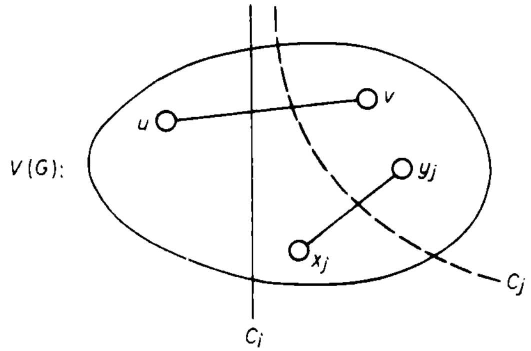


FIGURE 2.3.3.

We claim that u and x_j are not separated by $C_0, \ldots, C_{i-1}, C_{i+1}, \ldots,$ C_{j-1}, for suppose C_k from this list does, in fact, separate them. Cut C_k does not cross C_j and hence either separates u from v or x_j from y_j. But C_k cannot separate u from v by the induction hypothesis, so it separates x_j from y_j. But this is a contradition of the definition of C_j.

Hence u and x_j are not separated by $C_0, \ldots, C_{i-1}, C_{i+1}, \ldots, C_{j-1}$ (nor by C_j!). So let $u = u$ and $v = x_j$ and the proof of Claim 3 is complete.

So by Claims 2 and 3 each C_i contains exactly one line of T and by the definition of T, every line of T is contained in exactly one of the

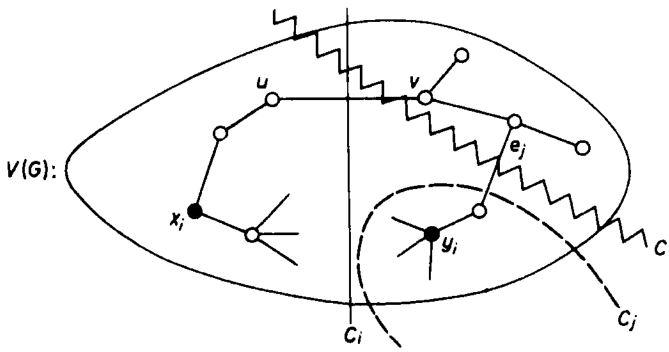


FIGURE 2.3.4.

C_i's. Thus T has exactly $p-1$ lines and every line of T is a cutline of T. Hence T is a tree.

Finally we claim that for all i, $0 \leq i \leq p-2$, if $e_i = uv$ is the unique line of T lying in cut C_i, then C_i is a minimum capacity $u-v$ cut in G.

Suppose, to the contrary, that this is false for some i. Let i be the largest index for which the statement is false. Then there exists a cut C separating u and v with $c(C) < c(C_i)$. Note that C cannot separate x_i and y_i by the minimality of C_i. On the other hand, C must separate the endpoints of some line e_j in T. (Cf. Figure 2.3.4.)

Consider cut C_j associated with line e_j. It separates x_i and y_i, since it does not cross C_i, and hence $j > i$. But then by the maximality of i, $c(C) \geq c(C_j)$. Now neither C_i nor C_j crosses any of $C_0, C_1, \ldots, C_{i-1}$ and since both separate x_i and y_i, it must be the case that $c(C_j) \geq c(C_i)$ (or else C_j would have been chosen as C_i!). Thus $c(C) \geq c(C_i)$, contradicting the definition of cut C. This completes the proof of the theorem. ∎

The "uncrossing procedure", which has been central to this proof, is a powerful tool which we shall use repeatedly later in this book.

Given a graph with line capacities, (G, c) and its corresponding cut-equivalent tree T, define a capacity c' for each line e of T by $c'(e) = \text{cap}(C_e)$ where, as before, C_e is the cut in G induced by line e.

2.3.2. THEOREM. (*The Flow-equivalent Tree Theorem*). *If* (G, c) *and* (T, c') *are as given in the preceding paragraph, then* (T, c') *is a flow-equivalent tree.*

PROOF. Let x and y be any two points in G. Now in T there is a unique $x-y$ path; denote it by $e_1 \cdots e_k$ where e_1 is incident with x, e_k with y. So $\max_T \text{val}(f_{xy}) = \min\{c'(e_1), \ldots, c'(e_k)\}$ where the maximum is taken over all $x-y$ flows in T. Suppose this value is $c'(e_r)$, $e_r \in \{e_1, \ldots, e_k\}$. Thus

$$\begin{aligned}\max{}_T \operatorname{val}(f_{xy}) = c'(e_r) &= \operatorname{cap}(C_{e_r}) \\ &\geq \max{}_G \operatorname{val}(f_{xy}). \qquad (2.3.1)\end{aligned}$$

On the other hand, let C be any minimum $x-y$ cut in G. Then there must exist a line e_s along the (unique) $x-y$ path in T such that $e_s \in C$. (Cf. Figure 2.3.5.)

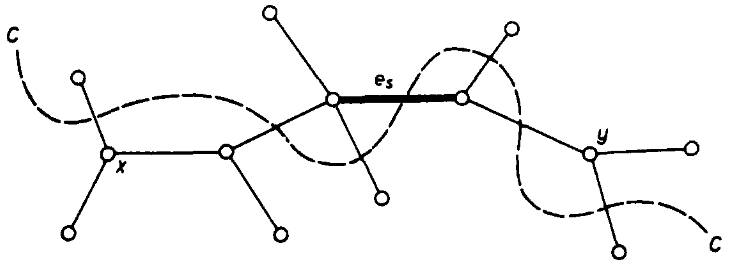


FIGURE 2.3.5.

Now by the definition of cut-equivalent tree, $c'(e_s) = \min\{\operatorname{cap}(C^*)\}$ where the minimum is taken over all cuts C^* in G separating the endpoints of e_s. Thus in particular, since C separates the endpoints of e_s, $c'(e_s) \leq \operatorname{cap}(C)$. Thus

$$\begin{aligned}\max{}_G \operatorname{val}(f_{xy}) &= \operatorname{cap}(C) \\ &\geq c'(e_s) \\ &\geq c'(e_r) \\ &= \max{}_T \operatorname{val}(f_{xy}). \qquad (2.3.2)\end{aligned}$$

Upon combining inequalities (2.3.1) and (2.3.2) the proof is complete. ■

2.3.3. COROLLARY. *Let (G, c) be an undirected graph with line capacity function c and associated cut-equivalent tree T. Then for all $x, y \in V(G)$, there is a line $e \in E(T)$ such that the cut C_e in G associated with line e is a minimum $x-y$ cut in (G, c).*

PROOF. Let θ denote the minimum value of all $x-y$ cuts in G. Then $\theta = \max_G \mathrm{val}(f_{xy}) = \max_T \mathrm{val}(f_{xy}) = \min\{c'(e_1), \ldots, c'(e_k)\} = c'(e_r) = \mathrm{cap}(C_{e_r})$, where e_r is as in the proof of the preceding theorem. But then e_r is the line we seek. ∎

Note that the result of this corollary "sharpens" the property in the definition of cut-equivalent tree in the sense that in this corollary x and y are no longer necessarily adjacent, whereas in the definition of cut-equivalent tree they were.

2.4. Applications of Flow Theory to Matching Theory

Having suffered through the above material on flows, the reader may well ask why we drag this material into a book on matching theory! There exists a large family of interesting problems for graphs and digraphs usually known as "degree-constrained subgraph" problems. In such problems we are often given non-negative integers (or integer pairs) assigned to the points of a graph and then we are asked to decide when a subgraph exists having degrees equal to the prescribed integers, bounded by the given integers, etc. Indeed the existence of a perfect matching is such a problem; namely, how do we decide when a given graph has a spanning subgraph with all degrees = 1 at every point. The statement of the f-matching problem is a straightforward generalization of the maximum matching problem although the *solution* is not so straightforward. (The reader's attention is directed to Chapter 10 for a more detailed study.)

In the next several theorems we present some sample problems of this kind which arise for bigraphs and digraphs. Our selection is motivated here by the fact that these particular problems can be easily solved by building a suitable network based on the graph at hand and then applying the Max-Flow Min-Cut Theorem (i.e., Theorem 2.1.4).

Among the most important results in the study of connectivity of graphs and digraphs are several which are usually lumped together and collectively called "Menger's Theorem" (1927). The theory of flows developed above will provide short proofs.

2.4.1. THEOREM. (*Menger's Theorem*).

(a) (*undirected line version*)

Let G be an undirected graph with two distinguished points s and t. Then the maximum number of line-disjoint $s-t$ paths in G is equal to the minimum size of any $s-t$ separating set of lines.

(b) (*undirected point version*)

If G is as in part (a), and s is not adjacent to t, the maximum number of *openly disjoint $s-t$ paths is equal to the minimum number of points the deletion of which destroys all $s-t$ paths.*

(c) (*directed line version*)

If D is a digraph with distinguished points s and t, then the maximum number of line-disjoint (directed) $s-t$ paths is equal to the minimum number of lines the deletion of which destroys all directed $s-t$ paths.

(d) (*directed point version*)

Let D be as in part (c) and suppose $(s,t) \notin E(D)$. Then the maximum number of openly disjoint directed $s-t$ paths is equal to the minimum number of points the deletion of which destroys all directed $s-t$ paths.

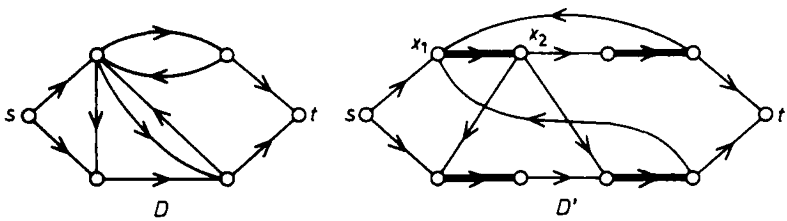


FIGURE 2.4.1.

PROOF. Let us consider version (c) first. Assign each line of D a capacity of 1. The result then follows immediately from the Max-Flow Min-Cut Theorem 2.1.4 and the Flow Integrality Theorem 2.2.1.

Now consider (d). Construct a new digraph D' as follows: split each point $x \neq s,t$ into two points x_1 and x_2 and join x_1 to x_2 by a directed line. Moreover, for each line of D into x, insert a line in D' into x_1 and a line of D' out of x_2 (cf. Figure 2.4.1); that is, x_2 is joined to y_1 in D' if and only if x is joined to y in D.

The reader is invited to complete the proof of (d) using part (c). The undirected versions follow from their directed counterparts by replacing every undirected line uv by two directed lines (u,v) and (v,u). ■

Let us next apply the Max-Flow Min-Cut Theorem to get a quick proof of König's Minimax Theorem. Let $G = (A,B)$ be any bigraph. Build a network G' by directing all lines of G from A to B, adding a new point s (the "source") joined to all points of A and a new point t (the "sink") to which all points of B are joined, and then assigning capacity ∞ to all lines of G and capacity 1 to all new lines of G'. The proof is then complete upon realizing that a maximum matching in G corresponds to a maximum flow in G' and a minimum cover in G corresponds to a minimum $s-t$ cut in G'.

An extremely interesting historical note is in order here. Menger's proof of his famous minimax theorem (1927) contained a gap. In fact, he failed to consider the case in which the graph involved was bipartite. So König was really the first to formulate and prove Menger's Theorem for this omitted case by means of his Minimax Theorem for bigraphs (Theorem 1.1.1)! König first noted in print in his book (1936) that his Minimax Theorem repaired this gap in Menger's proof.

Next let f be a non-negative integer-valued function defined on $V(G)$ where G is any graph. An **f-factor** is a spanning subgraph G' of G such that $\deg_{G'}(v) = f(v)$ for all $v \in V(G)$. We will show that flow theory provides an efficient tool for obtaining the following characterization of f-factors in bipartite graphs. In the following let $q(X,Y)$ denote the number of lines having one endpoint in set X, the other in Y.

2.4.2. THEOREM. *Let $G = (A, B)$ be a bigraph and let $f(v)$ be a non-negative integer-valued function on $V(G)$. Then G has an f-factor if and only if*

(i) $\sum_{v\in A} f(v) = \sum_{w\in B} f(w)$ *and*

(ii) *for all $X \subseteq A$ and $Y \subseteq B$, we have*

$$\sum_{x\in X} f(x) \le q(X,Y) + \sum_{y\in B-Y} f(y).$$

PROOF. ($\Rightarrow$). First suppose G has an f-factor F. Then if $|F|$ denotes the number of lines in F, clearly,

$$\sum_{v\in A} f(v) = |F| = \sum_{w\in B} f(w),$$

that is, (i) is satisfied.

Now choose $X \subseteq A$ and $Y \subseteq B$. There are $q(X,Y)$ lines of G joining X to Y and hence at most $q(X,Y)$ of F joining these two sets. Similarly, there are $\sum_{y\in B-Y} f(y)$ lines of F joining A to $B-Y$ and hence at most $\sum_{y\in B-Y} f(y)$ lines of F joining X to $B-Y$. But there are exactly $\sum_{x\in X} f(x)$ lines of F joining X to B and so (ii) is satisfied.

($\Leftarrow$). Now suppose (i) and (ii) hold. Orient all lines of G from A to B. Add a source point s joined to all points of A and a sink t to which every point of B is joined. Assign capacities to all lines of the resulting digraph as follows: $c(e) = f(x)$ if e joins s to x or x to t and $c(e) = 1$ on all other lines.

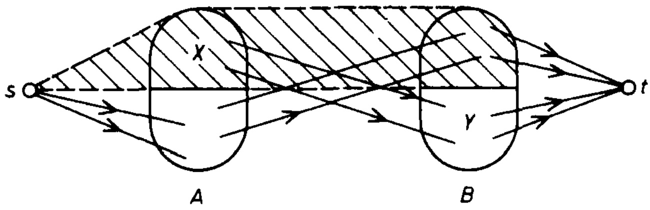


FIGURE 2.4.2.

Clearly, G has an f-factor if and only if the new digraph D has an (integral) flow from s to t of value $\sum_{v\in A} f(v) = \sum_{w\in B} f(w)$.

We claim that if S is any $s-t$ separator, where $s \in S \subseteq A \cup B$, the capacity of the separator will be at least $\sum_{v\in A} f(v)$. To see this, let $X = S \cap A$ and $Y = B - S$. Then we have (see Figure 2.4.2)

$$\begin{aligned}
\operatorname{cap}(S) &= \sum_{x\in A-X} f(x) + q(X,Y) + \sum_{y\in B-Y} f(y) \\
&\geq \sum_{x\in A-X} f(x) + \sum_{x\in X} f(x), \\
&= \sum_{x\in A} f(x).
\end{aligned}$$

Since S was an arbitrary separator, we have by the Max-Flow Min-Cut Theorem that D has a maximum $s-t$ flow F_0 such that $|F_0| \geq \sum_{x\in A} f(x)$. Moreover, since all the capacities in D are integral, F_0 may be chosen to be integral and hence D has an integral flow of value $\sum_{x\in A} f(x)$. ∎

2.4.3. EXERCISE. (a) Show that Theorem 2.4.2 contains the Marriage Theorem as a special case.

Complete a circle of four implications by showing that (b) the Marriage Theorem implies the directed line version of Menger's Theorem, and

(c) that the directed line version of Menger's Theorem implies the Max-Flow Min-Cut Theorem.

Suppose next that $G = (A,B)$ is again a bigraph and f, a non-negative real function on $V(G)$. A collection M_f of lines of G (with repetition allowed) is an f**-matching** if

$$\sum_{x\in E} w(E) \leq f(x) \qquad (2.4.1)$$

for each line E and each point x in G. Here $w(E)$ is the number of times E appears in M_f. If equality holds in inequality (2.4.1), M_f is said to be a perfect f-matching. Finally, note that if no line is repeated in a perfect f-matching M_f, then M_f is just an f-factor. We point out to the reader that the special case of 2-matchings is studied in considerable detail in Chapter 6.

The following necessary and sufficient conditions for the existence of a perfect f-matching are easily verified using flow theory. In fact, we need only change some capacities on the network model in the previous proof.

2.4.4. THEOREM. *Let $G = (A, B)$ be a bigraph with a non-negative integer function f defined on $V(G)$. Then G has a perfect f-matching if and only if*
(i) $\sum_{x\in A} f(x) = \sum_{y\in B} f(y)$ *and*
(ii) $\sum_{x\in X} f(x) \le \sum_{x\in \Gamma(X)} f(x)$, *for all* $X \subseteq A$.

PROOF. The proof of this theorem is analogous to that of Theorem 2.4.2. ■

It may be surprising that although we must postpone the discussion of the f-factor problem for non-bipartite graphs until Chapter 10, the corresponding problem for arbitrary *di*graphs can be easily solved at this point. Let D be a digraph. When does D have a subdigraph with in- and outdegrees equal to prescribed integer values?

In the following theorem let $\vec{m}_D(S,T)$ denote the number of lines in D with tail in S and head in T.

2.4.5. THEOREM. *Let D be a digraph with integer pairs $\big(f(x), g(x)\big)$ assigned to each point x. Then D has a subdigraph D' with* $\deg^-(x) = f(x)$ *and* $\deg^+(x) = g(x)$ *if and only if*
(i) $\sum_{x\in V(G)} f(x) = \sum_{x\in V(G)} g(x)$ *and*
(ii) *for all* X, $Y \subseteq V(D)$,

$$\sum_{x\in X} g(x) \le \sum_{y\in Y} f(y) + \vec{m}_D(X, V(G) - Y)$$

PROOF. We will construct an appropriate bigraph H to which we will apply Theorem 2.4.2. Let $V(H') = (A, B)$ where A and B are two disjoint copies of $V(D)$; that is, let $A = \{x' \mid x \in V(D)\}$ and $B = \{x'' \mid x \in V(D)\}$. Then let $E(H') = \{(x', x'') \mid (x,y) \in E(D)\}$. Finally, let H be the undirected bigraph arising from H' by surpressing the line orientations (cf. Figure 2.4.3).

Now weight the points of H with function h defined by:

$$h(y) = \begin{cases} g(x), & \text{when } y = x' \\ f(x), & \text{when } y = x''. \end{cases}$$

Now observe that D has a subdigraph D' with the desired properties if and only if H has an h-factor. But by Theorem 2.4.2 such an h-factor exists if and only if (i) and (ii) hold. ■

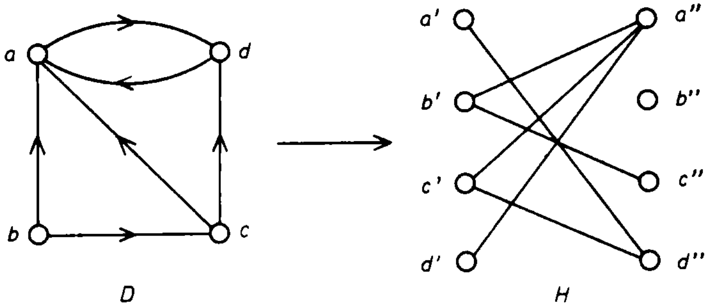


FIGURE 2.4.3.

Suppose next that we have an undirected graph G with non-negative integer weights on each point. When can G be oriented in such a way that the resulting digraph has outdegree at each point equal to the given integer?

2.4.6. THEOREM. *Let G be a graph and $f(x)$ a non-negative integer assigned to each point of G. Then there is an orientation D of G such that $\deg_D^+(x) = f(x)$ for every $x \in V(G)$ if and only if*

(i) $\sum_{x \in V(G)} f(x) = |E(G)|$ *and*

(ii) $|Z(X)| \geq \sum_{x \in X} f(x)$, *for each $X \subseteq V(G)$ where $Z(X)$ denotes the set of lines incident with X.*

PROOF. Build a bigraph H from G by inserting one new point on each line of G and extend f to H by setting $f = 1$ for each of these "new" points.

Then G has an orientation D of the type desired if and only if H has an f-factor. Translating the conditions of Theorem 2.4.2 from the graph H to the graph G we obtain the conditions in the statement of the theorem. ■

The preceding theorem is just one of a number of interesting results dealing with the task of orienting undirected graphs so that the resulting digraph satisfies certain prescribed conditions. One of the classical results of the type is due to Robbins (1939) who proved that every 2-line-connected graph possesses a strongly connected orientation. Frank and Gyárfás (1978) obtained the variation on Theorem 2.4.6 in which the orientation is required to be strongly connected.

A common generalization of these problems is obtained when we are looking for orientations where the in- and outdegrees of all subsets of points — not just single points — are prescribed. The reader is referred to Nash-Williams (1960), Hakimi (1965) and Frank (1980).

We mention one final example of the use of flow theory in bigraphs. It may be used to obtain the following result of Folkman and Fulkerson (1969) on the number of line-disjoint matchings of a given size in a bigraph. Let $q(X,Y)$ be as above.

2.4.7. THEOREM. *Let $G = (A,B)$ be a bigraph. Then G has h line-disjoint matchings of size t if and only if, for all $X \subseteq A$ and $Y \subseteq B$, $q(X,Y) \geq h(t - |A-X| - |B-Y|)$.*

PROOF. A challenging exercise for the reader! (Hint: Build a new bipartite graph H by adjoining $|A| - t$ new points to B and $|B| - t$ new points to A. Join each new point to each point in the opposite color class by many multiple lines. Then G has h line-disjoint matchings of size t if and only if H has an h-factor.) ∎

2.5. Matchings, Flows and Measures

This section contains two applications of matching theory to measure theory. This connection is somewhat surprising since matching theory is a very "discrete" branch of mathematics while measure theory belongs to the foundations of "continuous" mathematics. It seems, however, that in continuous mathematics one more and more frequently encounters situations where standard techniques of analysis reduce the problem to a non-trivial combinatorial situation and it is quite natural that matching theory often serves as a tool for solving these combinatorial problems.

Let us formulate a measure-theoretic version of the bipartite matching problem. For a general reference on measure theory see Halmos (1950). Let $(S_i, \mathcal{A}_i, P_i)(i = 1,2)$ be probability spaces and let $(S_1 \times S_2, \mathcal{A}, P)$ be their product. We say that a measure Q on the σ-algebra $(S_1 \times S_2, \mathcal{A})$

has **marginals** P_1 and P_2 if

$$Q(X \times S_2) = P_1(X) \quad \text{for every } X \in \mathcal{A}_1, \text{ and}$$
$$Q(S_1 \times Y) = P_2(Y) \quad \text{for every } Y \in \mathcal{A}_2.$$

It is clear that $Q = P$ has marginals P_1 and P_2, but in general there will be many more such probability measures.

Assume now that a set $E \in \mathcal{A}$ is also given. We say that a probability measure Q on $(S_1 \times S_2, \mathcal{A})$ is **concentrated on** E, if $Q(S_1 \times S_2 - E) = 0$.

Now consider the following problem. Given $(S_i, \mathcal{A}_i, P_i)$ $(i = 1,2)$ and $E \in \mathcal{A}$, when does a measure Q exist on $(S_1 \times S_2, \mathcal{A})$ which is concentrated on E and has marginals P_1 and P_2?

2.5.1. EXERCISE. Prove: This problem includes the supply-demand problem. (See Corollary 2.1.5.)

A rather trivial necessary condition for the existence of a measure Q with the above properties is the following condition $(*)$. For every $X \subseteq S_1$, let EX denote the set of those $y \in S_2$ for which there is an $x \in X$ with $(x,y) \in E$.

$(*)$ For every $X \in \mathcal{A}_1$, and $Y \in \mathcal{A}_2$ such that $EX \subseteq Y$, we have $P_1(X) \le P_2(Y)$.

In fact, assume that a probability measure Q exists which is concentrated on E and has marginals P_1 and P_2. Then

$$P_1(X) = Q(X \times S_2) = Q(X \times S_2 \cap E) = Q(X \times Y \cap E) = Q(X \times Y),$$

and so

$$P_2(Y) = Q(S_1 \times Y) \ge Q(X \times Y) \ge P_1(X).$$

To see that $(*)$ is not always sufficient for the existence of the measure Q, let, for example, $(S_1, \mathcal{A}_1, P_1) = (S_2, \mathcal{A}_2, P_2)$ be the σ-algebra of Borel subsets of [0,1], with the usual Lebesgue measure. Let

$$E = \{(x,y) \mid 0 \le x < y \le 1\}.$$

Then $(*)$ is fulfilled; in fact, EX contains all but at most one point of X, for every $X \subseteq (0,1)$. Suppose that there exists a measure Q concentrated on E with marginals P_1 and P_2. Consider the rectangle

$$R_t = [0,t] \times (t,1].$$

Then

$$\begin{aligned} Q(R_t) &= Q\big([0,t] \times [0,1]\big) - Q\big([0,t] \times [0,t]\big) \\ &= Q\big([0,t] \times [0,1]\big) - Q\big([0,1] \times [0,t]\big) \\ &= P_1[0,t] - P_2[0,t] = t - t = 0. \end{aligned}$$

So $Q(R_t) = 0$. But since E is the union of countably many rectangles R_t, it follows that $Q(E) = 0$. This is a contradiction, since

$$Q(E) = Q\big([0,1] \times [0,1]\big) = 1.$$

So no such measure Q can exist.

On the other hand, if some additional information about E is given, the condition $(*)$ may turn out to be sufficient. We prove the following result (P. Major, unpublished; for a similar application of matchings to measures, see Dudley (1968) and Lovász and Major (1973)). (Cf. also Strassen (1965).)

2.5.2. THEOREM. *Let E be a closed subset of $[0,1] \times [0,1]$ and let λ denote the Lebesgue measure. Suppose that for any two Borel sets $X, Y \subseteq [0,1]$ such that $EX \subseteq Y$, we have $\lambda(X) \le \lambda(Y)$. Then there exists a measure μ on the Borel sets of $[0,1] \times [0,1]$ which is concentrated on E, both marginals of which are λ.*

Let M denote the set of all probability measures on the Borel subsets of $[0,1] \times [0,1]$, and let A be the set of those measures in M with both marginals λ, that is,

$$A = \{\, \mu \in M \mid \mu\big(X \times [0,1]\big) = \big(\mu[0,1] \times X\big) = \lambda(X) \quad \text{for all Borel sets } X \subseteq [0,1] \,\}.$$

The combinatorial "backbone" of the proof of this theorem is the following lemma.

2.5.3. LEMMA. *Let $[0,1] = \bigcup_{i=1}^{N} A_i = \bigcup_{j=1}^{M} B_j$ be two partitions of the interval $[0,1]$ into disjoint Borel sets. Then there exists a measure $\nu \in A$ such that $\nu(A_i \times B_j) = 0$ for all (i,j) such that $E \cap (A_i \times B_j) = \emptyset$.*

PROOF. Build a bipartite digraph G on the set $V(G) = \{A_1, \dots, A_N, B_1, \dots, B_M\}$ by connecting A_i to B_j if and only if $(A_i \times B_j) \cap E \ne \emptyset$. Let A_i have supply $\lambda(A_i)$ and B_j have demand $\lambda(B_j)$. Then the conditions of the Supply-Demand Theorem (Corollary 2.1.5) are fulfilled:

(1) $\sum_{i=1}^{N} \lambda(A_i) = \sum_{j=1}^{M} \lambda(B_j) = 1$;

(2) For any subset $I \subseteq \{1, 2, \dots, N\}$,

$$\begin{aligned}&\sum_{A_iB_j\in E(G)\text{ for some } i\in I}\lambda(B_j)\\&=\sum_{(A_i\times B_j)\cap E\neq\emptyset\text{ for some } i\in I}\lambda(B_j)\\&\geq\lambda\Big(E\big(\bigcup_{i\in I}A_i\big)\Big)\geq\lambda\Big(\bigcup_{i\in I}A_i\Big)\\&=\sum_{i\in I}\lambda(A_i).\end{aligned}$$

Thus there exist flow values f_{ij} $(1\leq i\leq N,\ 1\leq j\leq M)$ such that $\sum_{j=1}^{M}f_{ij}=\lambda(A_i)$, $\sum_{i=1}^{N}f_{ij}=\lambda(B_j)$, $f_{ij}\geq 0$ and $f_{ij}=0$ if $(A_i\times B_j)\cap E=\emptyset$.

Define, for each Borel subset $X\subseteq[0,1]\times[0,1]$,

$$\nu(X)=\sum_{i=1}^{N}\sum_{j=1}^{M}f_{ij}\lambda^2\big(X\cap(A_i\times B_j)\big)$$

(here λ^2 is the 2-dimensional Lebesgue measure). Then the properties of ν required in the lemma are easily verified. ∎

As an immediate corollary, we have

2.5.4. COROLLARY. *Let $G_1,\dots,G_n,H_1,\dots,H_n\subseteq[0,1]$ such that $(G_i\times H_i)\cap E=\emptyset$. Then there exists a measure $\nu\in A$ such that $\nu(G_i\times H_i)=0$ for $i=1,\dots,n$.*

PROOF. Let $A_1,\dots,A_N$ be the atoms of the Boolean algebra generated by $G_1,\dots,G_n$ and $B_1,\dots,B_M$, the atoms of the Boolean algebra generated by $H_1,\dots,H_n$. Then the measure defined in the preceding lemma clearly satisfies the conditions. ∎

PROOF (of Theorem 2.5.2). We can introduce a topology on M by letting the neighborhoods of $\mu\in M$ be all sets

$$\{\nu\in M\mid |\int f_i d\mu-\int f_i d\nu|<\epsilon\text{ for } i=1,\dots,k\}$$

where $\epsilon>0$ and $f_1,\dots,f_k$ are continuous functions defined on $[0,1]\times[0,1]$. It is well known (see Billingsley (1968)) that M is compact in this topology. Further, A is a closed subset of M.

Let G,H be open subsets of $[0,1]$ such that $(G\times H)\cap E=\emptyset$. Let

$$A(G\times H)=\{\nu\in A\mid \nu(G\times H)=0\}.$$

Then $A(G\times H)$ is a closed subset of M. By Corollary 2.5.4,

$$A(G_1\times H_1)\cap\cdots\cap A(G_n\times H_n)\neq\emptyset$$

for any finitely many open subsets $G_1,\ldots,G_n,H_1,\ldots,H_n$ such that $(G_i\times H_i)\cap E=\emptyset$. So by the compactness of M,

$$\bigcap_{G,H} A(G\times H)\neq\emptyset,$$

where G and H range over all pairs of open subsets of $[0,1]$ such that $(G\times H)\cap E=\emptyset$. Let $\mu\in\bigcap_{G,H}A(G\times H)$. Then $\mu\in A$ and $\mu(G\times H)=0$ for all pairs G,H of open subsets of $[0,1]$ with $(G\times H)\cap E=\emptyset$. But $[0,1]\times[0,1]-E$ is the union of countably many such sets of the form $G\times H$. Thus $\mu\big([0,1]\times[0,1]-E\big)=0$ and the proof is complete. ∎

It may be worth while to mention the following corollary of Theorem 2.5.2 (see Kamae, Krengel and O'Brien (1977)).

2.5.5. COROLLARY. *Let F_1 and F_2 be distribution functions such that $F_1(x)\geq F_2(x)$ for all $x\in\Re$. Then there exists a random vector variable (ξ_1,ξ_2) such that ξ_i has distribution function F_i and $\xi_1\leq\xi_2$.*

PROOF. We offer a sketch only. Let

$$E=\{\,\big(F_1(x),F_2(y)\big)\mid x\leq y\,\}.$$

Then E satisfies the conditions in Theorem 2.5.2, and hence there exists a measure μ on the Borel subsets of $[0,1]\times[0,1]$, concentrated on E, with marginals λ. Define

$$F(x,y)=\mu\Big(\big(0,F_1(x)\big)\times\big(0,F_2(y)\big)\Big).$$

Then any vector variable with distribution function $F(x,y)$ satisfies the requirements. ∎

REMARK. The same proof shows the sufficiency of $(*)$ if S_1 and S_2 are compact metric spaces.

Another — probably even more striking — application of matching to measure theory is the construction of the Haar measure on locally compact topological groups. This application was mentioned by Rota and Harper (1971), who elaborated upon the idea in the example of the construction of a translation invariant integral for almost periodic functions on an arbitrary group. Here we shall describe how to construct a translation invariant integral for continuous functions on a compact topological group. The existence of such an integral, essentially equivalent to the existence of Haar measure on such groups, was proved by von Neumann (1934).

Recall that a **topological group** is a group G endowed with a topology in which the group operations xy and x^{-1} are continuous. (For a general reference on topological groups see Pontrjagin (1939).) An **invariant integration** for continuous functions is a linear functional L defined on the continuous real-valued functions on G, having the following properties:

(1) $L(\alpha f+\beta g)=\alpha L(f)+\beta L(g)$. (linearity)
(2) If $f\geq 0$ then $L(f)\geq 0$. (monotonicity)
(3) If l denotes the identity function, then $L(l)=1$. (normalization)
(4) If s and t are in G and f and g are two continuous functions such that $g(x)=f(sxt)$ for every x, then $L(g)=L(f)$. (double translation invariance)

2.5.6. THEOREM. *For every compact topological group there exists an invariant integration for continuous functions.*

Before proving this theorem we need some new ideas, notation and lemmas. If A is a finite subset of group G and f any function defined on G, we set

$$\bar{f}(A)=\frac{1}{|A|}\sum_{a\in A} f(a).$$

For $U\subseteq G$, we define

$$\delta(f;U)=\sup\{|f(x)-f(y)| \mid x,y\in sUt \text{ for some } s,t\in G\}.$$

Here $sUt=\{sut \mid u\in U\}$ and is known as the **translate** of U.

The compactness of G implies that if f is continuous then it is also uniformly continuous in the sense that for every $\epsilon>0$ there exists a non-empty open set U such that $\delta(f;U)<\epsilon$.

A subset $A\subseteq G$ is called an U-**net** ($U\subseteq G$) if $A\cap sUt\neq\varnothing$ for every $s,t\in G$. It follows from the compactness of G by standard arguments that if U is open and non-empty then there exists a finite U-net.

The heart of the proof of Theorem 2.5.6 is the following lemma.

2.5.7. LEMMA. *Let A and B be minimum cardinality U-nets. Then*

$$|\bar{f}(A) - \bar{f}(B)| \le \delta(f;U).$$

PROOF. Define a bipartite graph G_0 with bipartition (A, B) by connecting $x \in A$ to $y \in B$ if and only if there exists $s, t \in G$ such that $x, y \in sUt$. We claim that this bipartite graph has a perfect matching. Trivially $|A| = |B|$. Let T be any subset of $V(G_0)$ covering all lines. We show that T is a U-net. Let $s, t \in G$. Since A is a U-net, there exists an element $x \in A \cap sUt$. Similarly, there exists an element $y \in B \cap sUt$. So xy is an line of G_0 and by the definition of T, T must contain at least one of x and y. Thus T is a U-net. Since A is a U-net with minimum cardinality, we must have $|T| \ge |A|$. Since this holds for every T, König's Minimax Theorem 1.1.1 implies that G_0 has a perfect matching.

So let $\{a_1b_1, \dots, a_nb_n\}$ be a perfect matching in G_0. Then

$$\begin{aligned}|\bar{f}(A) - \bar{f}(B)| = \frac{1}{n}\left|\sum_{i=1}^{n}\big(f(a_i) - f(b_i)\big)\right| &\le \frac{1}{n}\sum_{i=1}^{n}|f(a_i) - f(b_i)| \\ &\le \frac{1}{n}\big(n\delta(f;U)\big) = \delta(f;U).\end{aligned}$$ ■

We in fact shall use the following corollary of this lemma.

2.5.8. LEMMA. *Let A be a minimum finite U-net and B a minimum finite V-net. Then*

$$|\bar{f}(A) - \bar{f}(B)| \le \delta(f;U) + \delta(f;V).$$

PROOF. For every $b \in B$, Ab is also a U-net with minimum cardinality, so by Lemma 2.5.7,

$$|\bar{f}(A) - \bar{f}(Ab)| \le \delta(f;U).$$

Hence

$$|\bar{f}(A) - \bar{f}(AB)| = |\bar{f}(A) - \frac{1}{|B|}\sum_{b\in B}\bar{f}(Ab)| \le \frac{1}{|B|}\sum_{b\in B}|\bar{f}(A) - \bar{f}(Ab)| \le \delta(f;U).$$

Similarly,

$$|\bar{f}(B) - \bar{f}(AB)| \le \delta(f;V).$$

Hence $|\bar{f}(A) - \bar{f}(B)| \le \delta(f;U) + \delta(f;V)$ and the proof is complete. ■

PROOF (of Theorem 2.5.6). The integration can be constructed as follows. Let f be a continuous function. Let U_n be a sequence of open sets such that $\delta(U_n, f) \to 0$. Let A_n be any U_n-net with minimum cardinality. Then the sequence $\{\bar{f}(A_n)\}$ satisfies Cauchy's criterion of convergence by Lemma 2.5.8 and hence it tends to a limit $L(f)$. Lemma 2.5.8 also implies that this limit is independent of the choice of U_n and A_n, and so it is well defined.

Conditions (1)—(4) are also satisfied. To show (1), choose non-empty open sets U_n, U'_n, U''_n so that

$$\delta(\alpha f + \beta g; U_n) \to 0, \quad \delta(f; U'_n) \to 0, \quad \delta(f; U''_n) \to 0.$$

Since $\delta(f; U)$ is invariant under the translation of U, we may assume without loss of generality that $V_n = U_n \cap U'_n \cap U''_n \neq \emptyset$. Let A_n be a V_n-net with minimum cardinality. Then

$$\begin{aligned} L(\alpha f + \beta g) &= \lim\left(\alpha \bar{f}(A_n) + \beta \bar{g}(A_n)\right) \\ &= \alpha \lim \bar{f}(A_n) + \beta \lim \bar{g}(A_n) = \alpha L(f) + \beta L(g). \end{aligned}$$

Conditions (2) and (3) are trivial. Finally, to show (4) let U_n be a sequence of non-empty open sets with $\delta(f; U_n) \to 0$ and A_n a minimum U_n-net. Note that by the hypotheses, $\delta(g; U_n) = \delta(f; U_n)$ and that sA_nt is also a minimum U_n-net. Hence

$$L(g) = \lim \bar{g}(A_n) = \lim \bar{f}(sA_nt) = L(f).$$

■

This Page Intentionally Left Blank

3

Size and Structure of Maximum Matchings

3.0. Introduction

Suppose we are confronted with the following "real-life" problem. At Nashpest University we are given the task of assigning students to two-person dormitory rooms. In an attempt to accomodate the students as best we can, we ask to be provided with a list of pairs of students who would be willing to share a room. Based on this information, can we assign each person an acceptable roommate? Failing that, how can we fill a maximum number of rooms? Clearly these are matching problems in graphs where points represent students and two are joined by a line if and only if they find each other acceptable as roommates. But note immediately that the resulting graph is no longer bipartite. Indeed, if we had to consider bipartite graphs only, we could just as well have brought the book to a close right here!

But how do things really change when we attempt to do matching in a non-bipartite graph? In the bipartite case we have the Marriage Theorem 1.1.4 which serves as a good characterization of those graphs with perfect matchings. Together with these results we have a polynomial algorithm to determine a perfect (or maximum) matching in a bipartite graph. This algorithm is clearly polynomial and was realized to be so — at least as soon as mathematicians began to concern themselves with the question of polynomiality.

But a good characterization of *non*-bipartite graphs with perfect matchings had to wait until the late 1940's. And a good algorithm for finding perfect (or maximum) matchings was not found until the 1960's! In this chapter we will become acquainted with Tutte's good characterization theorem. Although much of the intervening material will be relevant to algorithmic work on matching, we will not present a polynomial matching algorithm — due to Edmonds — until Chapter 9.

The most important result in non-bipartite matching theory is due to Tutte (1947). This powerful result implies almost all of the results on matchings known previously — results like Petersen's Theorem 3.4.1 and the Marriage Theorem. Its power comes from the fact that it sets forth conditions which are both necessary and sufficient for the existence of a perfect matching in a general graph. We treat this theorem, as well as its "defect" version which has come to be called the Berge Formula, in Section 3.1.

In Section 3.2 we derive an important canonical structure theorem due independently to T. Gallai and J. Edmonds. Many interesting results are corollaries of this structure theorem, including Tutte's Theorem itself.

As we shall soon see, Tutte's Theorem shows that if a graph has no perfect matching then a certain type of cutset of points must exist. We call these, naturally enough, "Tutte sets". In Section 3 of this chapter we take a closer look at these Tutte sets and other related sets in graphs.

In the last section of this chapter we offer a survey of some of the various known sufficient conditions for the existence of perfect matchings and "large" matchings in general.

3.1. Tutte's Theorem, Gallai's Lemma and Berge's Formula

Suppose we have a graph G which contains a set of points S with the property that the number of components of odd (point) cardinality exceeds $|S|$. It is then clear that G cannot possibly have a perfect matching. The importance of the following theorem of Tutte (1947) lies in the demonstration that, conversely, if such a set S does not exist in G then G has a perfect matching. Let $c_o(G)$ denote the number of odd components of a graph G.

3.1.1. THEOREM. (*Tutte's Theorem*). *A graph G has a perfect matching if and only if $c_o(G-S) \leq |S|$, for all $S \subseteq V(G)$.*

Before presenting a proof of this theorem, a few remarks are in order.

This theorem gives a good characterization of the existence of a perfect matching (cf. Box 1A). If we want to exhibit the fact that a graph has a perfect matching, it suffices to list one. If we want to exhibit the fact that a graph G does not have a perfect matching (or say, the counselor of the dormitory at N.U. wants to convince a group of irate students of this), we can do so by specifying a Tutte set.

We have noted previously that the "only if" part of the proof is clear. At the risk of repeating ourselves, suppose G has a perfect matching F and that $S \subseteq V(G)$. Every odd component of $G-S$ must "send" at least one line of F to S, but every point in S "receives" at most one of these lines and so $G-S$ has at most $|S|$ odd components. Variations of this simple — but very important — counting argument will come up several times in the material to follow.

Tutte's original proof of the "if" part of this theorem used the idea of the Pfaffian of a matrix. Indeed, we shall have occasion to use the Pfaffian in Chapter 8 when counting perfect matchings in planar graphs. But it was not long before simpler purely combinatorial proofs of Tutte's Theorem (i.e., proofs avoiding the use of Pfaffians) were discovered. See Maunsell (1952), Tutte (1952), Gallai (1963a), Halton (1966) and Balinski (1970). In fact at least three proofs now exist which derive Tutte's result from the bipartite theorem of P. Hall (cf. Gallai (1963a), I. Anderson (1971), and Mader (1973)).

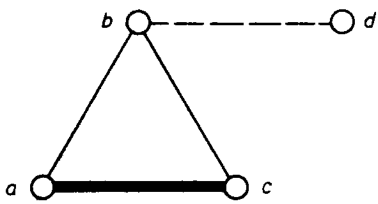


FIGURE 3.1.1.

We shall take a different approach from any of the above. (Cf. Hetyei (1972), Lovász (1975b)). First we characterize the structure of **saturated non-factorizable** graphs, that is, those graphs which have no perfect matching, but if any line is added, the resulting graph does.

3.1.2. LEMMA. *If G is saturated non-factorizable and if S is the set of points of G joined to every other point of G, then the components of $G-S$ are complete graphs.*

PROOF. Let ab and bc be adjacent lines of $G-S$. We claim that a and c are adjacent. Suppose, to the contrary, they are not. Then by the definition of S, there is a point d of G not adjacent to b (Figure 3.1.1).

Now $G+ac$ has a perfect matching F_1 by the maximality of G and similarly, let F_2 be a perfect matching of $G+bd$ containing bd. Now the symmetric difference of F_1 and F_2 consists of alternating (and therefore even) cycles and lines ac and bd each lie on such a cycle, say C_1 and C_2 respectively. We must treat two cases.

First suppose $C_1 \neq C_2$. In this case form the symmetric difference $F_3 = F_1 \oplus E(C_1)$. Then F_3 is a perfect matching of G, a contradiction.

Thus we may suppose $C_1 = C_2$. Traverse C_1 from b through d and continue until one of a and c is reached. Say, without loss of generality, a is reached first (cf. Figure 3.1.2). Let the $b-a$ path just traversed be P. Then $P + ab$ is an alternating cycle with respect to F_2. Form the

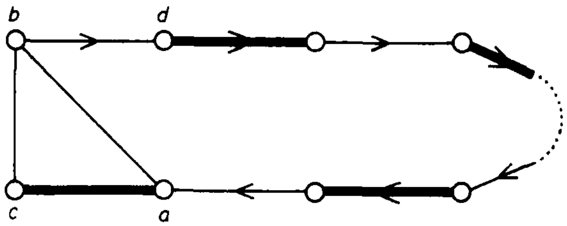


FIGURE 3.1.2.

symmetric difference $F_4 = F_2 \oplus E(P+ab)$. Then F_4 is a perfect matching of G, and again we have a contradiction.

Thus the components of $G - S$ are complete graphs as claimed. ■

3.1.3. THEOREM. *A graph G is saturated non-factorizable if and only if it has the following structure: either $|V(G)|$ is odd and G is complete, or $|V(G)|$ is even and G consists of point-disjoint complete subgraphs $S_0, G_1, \ldots, G_k$ such that $k = |S_0| + 2$, $G_1, \ldots, G_k$ are odd and every point of every G_i is connected to every point of S_0. (See Figure* 3.1.3.)

Before proceeding to the proof, we point out that this result has been derived as an easy consequence of Tutte's Theorem by a barely countable number of graph theorists; for published versions see Skupień (1973) and Homenko-Vivrot (1971, 1973). Here we do the opposite, namely we derive Tutte's Theorem from Theorem 3.1.3.

PROOF (of Theorem 3.1.3). This is an easy consequence of Lemma 3.1.2. If $|V(G)|$ is odd then trivially G is complete. If $|V(G)|$ is even then let S be the set defined in Lemma 3.1.2 and let $G_1, \ldots, G_k$ be the connected components of $G - S$. By the preceding lemma, these are complete graphs and by the definition of S, S spans a complete graph and every point of S is adjacent to every point in every G_i.

If at most $|S|$ components of $G - S$ are odd then a perfect matching of G is easily found. But this is a contradiction. So at least $|S| + 1$, and by parity, at least $|S| + 2$ components of $G - S$ are odd. If $G - S$

has more than $|S|+2$ odd components then by connecting two of these with a new line we still get a graph G_1 such that $c_o(G_1 - S) > |S|$. But then G_1 has no perfect matching, which contradicts the assumption that

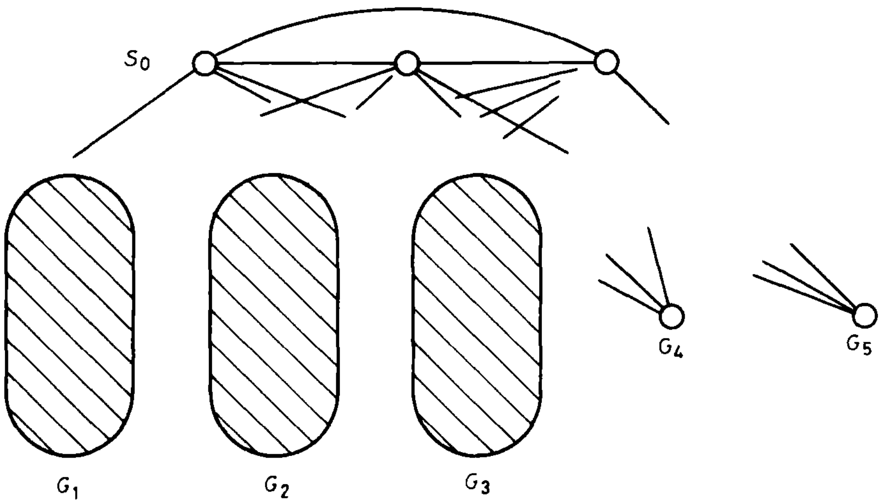


FIGURE 3.1.3.

G is saturated and non-factorizable. Thus $G - S$ has exactly $k+2$ odd components. It follows similarly that $G - S$ has no even components. ∎

PROOF (of Theorem 3.1.1). As noted earlier, necessity is trivial. To prove sufficiency suppose, to the contrary, that G has no perfect matching. Add lines to G as long as possible so that the resulting graphs have no perfect matchings. Let G' be an extremal line-saturated graph resulting from this procedure.

If $|V(G)|$ is odd, let $S = \varnothing$ and we are done. Hence suppose $|V(G)|$ is even. By Theorem 3.1.3, if S' is the set of all points adjacent to every point of G' and $H' = V(G) - S'$, then $H' \neq \varnothing$ since G' is even and has no perfect matching. Moreover, $H' = G_1' \cup \cdots \cup G_k'$ where the G_i''s are point-disjoint complete subgraphs and $k = |S'|+2$.

So $G' - S'$ has more than $|S'|$ odd components. If we now remove the lines of $E(G') - E(G)$ we inserted, each of these odd components of $G' - S'$ gives rise to at least one odd component of $G - S'$. Thus S' violates the hypothesis and this contradiction proves Tutte's Theorem.∎

3.1.4. EXERCISE. Derive the Marriage Theorem (Corollary 1.1.4) from Tutte's Theorem.

Just as in the bipartite case, if a graph has no perfect matching, it is natural to ask about the size of the maximum matchings. Before stating and proving a minimax formula for this number, we prove a few facts about matchings, some of which can be used in the proof of the minimax result.

The following results were proved by Edmonds and Fulkerson (1965).

3.1.5. LEMMA. *Let G be a graph, A and $B \subseteq V(G)$ and suppose $|A| < |B|$.*

(a) *Assume that there exists a matching which covers A and one which covers B. Then there exists a matching which covers all of A and at least one point of $B - A$.*

(b) *Assume that there exists a maximum matching which avoids A and one which avoids B. Then there exists a maximum matching which avoids A and at least one point of $B - A$.*

PROOF. We prove part (a); the proof of (b) is essentially the same. (Cf. also box 3A).

Let M_1 and M_2 be matchings which cover A and B respectively. Consider $M_1 \cup M_2$. The components of this subgraph are common lines of M_1 and M_2, alternating cycles and alternating paths. From every point in $B-A$ there starts an alternating M_2-M_1 path. Since $|B-A| > |A-B|$ by hypothesis, at least one of these paths, say P, does not end in $A-B$. But then the matching $M_1 \oplus E(P)$ still covers A, and, in addition, at least one point of $B-A$. ■

The following immediate corollaries will be of considerable use.

3.1.6. COROLLARY. *If a set of points is covered by some matching, then it is also covered by a maximum matching.* ■

3.1.7. COROLLARY. *Let G be a graph, $D \subseteq V(G)$ the set of those points of G which are missed by some maximum matching and for all $u, v \in D$, define $u \sim v$ if and only if $u = v$ or no maximum matching misses both u and v. Then $\sim$ is an equivalence relation.* ■

3.1.8. EXERCISE. Let G be any graph and suppose $X \subseteq V(G)$. Prove that G contains a matching covering all points in X if and only if for each $S \subseteq V(G)$, the graph $G - S$ has at most $|S|$ odd components which lie entirely in X.

A graph G is said to be **factor-critical** (or **hypomatchable**) if $G-v$ has a perfect matching for every $v \in V(G)$ (Gallai (1963a)). This property is a strong one indeed and it is perhaps surprising that (i) such graphs form a relatively rich collection for study and (ii) such graphs will play an important role in the decomposition theory we shall soon develop. This decomposition theory, in turn, will help us obtain estimates on the size of a maximum matching in an arbitrary graph.

3.1.9. EXERCISE. No bipartite graph is factor-critical.

3.1.10. EXERCISE. Show that a graph G is factor-critical if and only if G has an odd number of points and $c_o(G-S) \leq |S|$ for all $\emptyset \neq S \subseteq V(G)$ (that is, $\emptyset$ is the only Tutte set in G). (Note that replacing $|S|$ by $|S|-1$ on the right hand side results in an equivalent condition by parity.)

3.1.11. EXERCISE. Let G be factor-critical and suppose $v \in V(G)$. Split v into two new points v', v'' and partition the lines incident with v between v' and v'' subject only to the demand that the resulting graph G' remains connected. Prove that G' has a perfect matching.

3.1.12. EXERCISE. Let G be a connected graph such that every block of G is a triangle and every point has degree 2 or 4. Let S be the set of points having degree 2. (a) Prove that G is factor-critical. (b) Prove that for every subset $X \subseteq S$ with $|X|$ odd, $G-X$ has a unique perfect matching.

It is clear that if G is a factor-critical graph then $\nu(G-u)=\nu(G)$ for each $u \in V(G)$. Gallai (1963a) proved that for connected graphs the converse holds as well.

3.1.13. THEOREM. (*Gallai's Lemma*). *If graph G is connected and $\nu(G-u)=\nu(G)$ for each $u \in V(G)$, then G is factor-critical.*

We remark that an easy proof would follow from Tutte's Theorem, but here we choose a more direct proof based on Corollary 3.1.7.

PROOF. Let D be the point set defined in the statement of Corollary 3.1.7. Then by the hypothesis of the present theorem, $D=V(G)$. Now consider the equivalence relation $\sim$ defined in Corollary 3.1.7. Obviously, any two adjacent points are in relation $\sim$, since a matching missing both of them can be augmented by the line connecting them and so cannot be maximum. By the connectivity of G, any two points of G must be equivalent. But this means that no maximum matching misses more than one point; in other words,

$$\nu(G) \geq \tfrac{1}{2}\big(|V(G)| - 1\big) \qquad (3.1.1).$$

On the other hand, by hypothesis, for any $v \in V(G)$,

$$\nu(G) = \nu(G - v) \leq \tfrac{1}{2}|V(G - v)| = \tfrac{1}{2}\big(|V(G)| - 1\big) \qquad (3.1.2)$$

and upon comparing inequalities (3.1.1) and (3.1.2) the proof is complete. ■

A lot more can be said about the family of factor-critical graphs, but these remarks are better suited elsewhere (cf. Section 3.2 and also Chapter 5).

In a manner quite analogous to our approach for bipartite graphs in Chapter 1, we now want to determine the size of a maximum matching for graphs in general, in the case when no perfect matching exists. Fortunately, Ore's concept of deficiency can be suitably extended to general graphs. Recall from Theorem 1.3.1 that in a bipartite graph $G = (A, B)$ the (A)-deficiency of G is the number of points of A not matched by a maximum matching in G. For *general* (i.e., *non-bipartite*) graphs G, and for any matching M in G, let us define the **defect** of the matching M to be the number of points of G not matched by M. Now define the **deficiency** of G, def(G), by the equation def$(G) = |V(G)| - 2\nu(G)$. Hence def(G) is the number of points left uncovered by any maximum matching, that is, it is the minimum defect over the set of all matchings.

An analogue of Ore's Deficiency Theorem, Theorem 1.3.1, has been proved by Berge (1958a). The equation in this theorem is often called the "Berge Formula". The proof given here does not use Tutte's Theorem.

3.1.14. THEOREM. (*The Berge Formula*). *If* def(G) *denotes the deficiency of a graph* G, *then* def$(G) = \max\{\, c_o(G - X) - |X| \mid X \subseteq V(G)\,\}$.

PROOF. Choose any $X \subseteq V(G)$ and let M be any maximum matching in G. Denote the odd components of $G - X$ by $G_1, \ldots, G_k$ where $k = c_o(G - X)$. Among these components, renumbering if necessary, let $G_1, \ldots, G_i$ be those containing a point not covered by M. (Note that if def$(G) = 0$, that is, if G has a perfect matching, then the theorem is trivially true.) Thus for each of the $k - i$ components, $G_{i+1}, \ldots, G_k$, there is at least one line to M from the component of X in G. Thus $k - i \leq |X|$. On the other hand, def$(G) \geq i$ since each of $G_1, \ldots, G_i$ contains an uncovered point. Hence def$(G) \geq i \geq k - |X| = c_o(G - X) - |X|$.

We must now show that there is a set X for which equality holds, that is, one for which $\operatorname{def}(G) = c_o(G-X)-|X|$. Our proof is by induction on $|V(G)|$.

Verification for small values of $|V(G)|$ is trivial. Suppose now that G is any graph and that the theorem holds for all graphs with fewer points than G. There are two cases to consider:

Case 1. Suppose there is a point $v \in V(G)$ such that $\nu(G-v) < \nu(G)$. Then $\nu(G-v) = \nu(G)-1$ and we have $\operatorname{def}(G-v) = |V(G)-\{v\}| - 2\nu(G-v) = |V(G)| - 1 - 2((\nu(G)-1) = |V(G)| - 2\nu(G) + 1 = \operatorname{def}(G)+1$. Now by induction hypothesis, there exists a set $X' \subseteq V(G-v)$ such that $\operatorname{def}(G)+1 = c_o(G-v-X') - |X'|$. Setting $X = X' \cup \{x\}$, we obtain $\operatorname{def}(G)+1 = c_o(G-X) - |X| + 1$, that is, $\operatorname{def}(G) = c_o(G-X) - |X|$.

Case 2. Suppose $\nu(G-v) = \nu(G)$ for every $v \in V(G)$. Denote the components of G by $H_1, \ldots, H_r$. For a fixed H_i choose a point $v \in V(H_i)$. Then $\nu(G) = \nu(H_1) + \cdots + \nu(H_r) = \nu(G-v) = \nu(H_1) + \cdots + \nu(H_i - v) + \cdots + \nu(H_r)$ and hence $\nu(H_i) = \nu(H_i - v)$.

Since v is an arbitrary point in H_i it follows from the preceding lemma of Gallai (Theorem 3.1.13) that each H_i is factor-critical — and hence odd — and thus $\operatorname{def}(H_i) = 1$. But then $\operatorname{def}(G) = \sum_{i=1}^{r} \operatorname{def}(H_i) = r = c_o(G-\emptyset)$ and thus the choice of $X = \emptyset$ yields equality and the theorem is proved. ■

3.1.15. EXERCISE. Deduce Tutte's Theorem from Berge's Formula.

There are other ways to obtain the Berge Formula. Our approach is motivated by the fact that it distinguishes between points covered by every maximum matching and those which are not. (That is, point v is covered by every maximum matching if and only if $\nu(G-v) < \nu(G)$.) This distinction will reappear in the next section at the very beginning of our development of the so-called Gallai-Edmonds canonical decomposition.

One especially important alternate path (no pun intended!) to a proof of the Berge Formula is via Tutte's Theorem itself. Thus Tutte's Theorem too belongs to that class of results called "self-refining" just as do the theorems of König, P. Hall and Frobenius (cf. Chapter 1).

3.1.16. EXERCISE. Derive Berge's Formula using Tutte's Theorem. (Outline: Letting

$$\delta'(G) = \max\{\, c_o(G-X) - |X| \mid X \subseteq V(G)\},$$

the inequality $\delta'(G) \le \operatorname{def}(G)$ is obtained exactly as before. To obtain the reverse inequality construct a new graph G' from G by adjoining a set H of $\delta'(G)$ new points to G, joining each point of H to each point of G and to all other points of H. Now use Tutte's Theorem to show G' has a perfect matching.)

The following minimax formula for $\nu(G)$ for general graphs (cf. Edmonds (1965a)) is an analogue of König's Minimax Theorem for bipartite graphs. In fact, in Chapter 7 we will see that this is the "natural" analogue in the context of linear programming duality.

An **odd set cover** of a graph G is a set $\mathcal{H} = \{S_1, \ldots, S_k, v_1, \ldots, v_r\}$ of odd subsets $S_i \subseteq V(G)$ and points $v_i \in V(G)$ such that each line of G is either covered by one of the points v_i or spanned by one of the S_i's. The **weight** of an odd set cover $\mathcal{H}$ is $w(\mathcal{H}) = r + \sum_{i=1}^{k}(|S_i| - 1)/2$.

3.1.17. EXERCISE. For any graph G, $\nu(G) = \min w(\mathcal{H})$ where $\mathcal{H}$ ranges over all odd set covers.

Before ending our discussion of deficiency, we would like to make one more point. The equality of the two quantities in the Berge Formula is of more than passing interest. What this equation really says is that the quantity $\operatorname{def}(G)$ (and thereby $\nu(G)$) is *well-characterized* (see Box 1A). Suppose we are given a graph G and a non-negative integer d. First suppose $d \geq \operatorname{def}(G)$. How can we convince King Arthur of this? Let $r = (|V(G)| - d)/2$. Now if we can find a matching of size $\geq r$, we know — as does King Arthur — that $\operatorname{def}(G) = |V(G)| - 2\nu(G) \leq |V(G)| - 2r = d$, so he is convinced.

On the other hand, suppose $d \leq \operatorname{def}(G)$. How can we convince him in this case? If we can produce a set of points $X_0 \subseteq V(G)$ such that $c_o(G - X_0) - |X_0| \geq d$, then we both know that $\operatorname{def}(G) = \delta'(G) = \max\{\, c_o(G - X) - |X| \mid X \subseteq V(G)\,\} \geq c_o(G - X_0) - |X_0| \geq d$ and once more he is convinced. Thus we survive long enough to proceed to Section 3.2!

BOX 3A. Matching Matroids and Matroid Duality

The assertion of Lemma 3.1.5 enables us to define two matroids on $S = V(G)$. Let

$$\mu_1 = \{\, A \subseteq S \mid A \text{ is covered by some matching }\}$$
$$\mu_1^* = \{\, A \subseteq S \mid A \text{ is missed by some maximum matching}\}.$$

Then Lemma 3.1.5 asserts that both $(V(G), \mu_1)$ and $(V(G), \mu_1^*)$ satisfy the axiom of independent sets for matroids. Since the other two axioms are trivially valid, it follows that $(V(G), \mu_1)$ and $(V(G), \mu_1^*)$ are matroids. $(V(G), \mu_1)$ is called the **matching matroid** of G.

It is interesting to note that the matroid μ_1^* has the following relation to μ_1 : $A \in \mu_1^*$ if and only if μ_1 has a maximum independent set (i.e., a basis) disjoint from A. This property could be used to define a **dual matroid** (S, μ^*) for every matroid (S, μ). Of course, it takes some proof that this really defines a matroid, but this is not difficult and once it is done, parts (a) and (b) of Lemma 3.1.5 become equivalent. This fact explains why the proofs of statements (a) and (b) are essentially identical.

It is easy to see that the dual of the dual is the original matroid. Duality is a most important operation on matroids which generalizes, in an appropriate sense, duality of planar graphs, the current-voltage duality of electrical networks, transposition of matrices, orthogonality of linear subspaces, and more. We must refer the reader to the literature on matroids such as Tutte (1971), von Randow (1975) and Welsh (1976).

Returning to matching matroids, it can be shown that every transversal matroid is the restriction of a matching matroid. However, not every transversal matroid is a matching matroid; for example, every matching matroid must have even rank.

A considerably deeper result of Edmonds and Fulkerson (1965) asserts that every matching matroid is transversal. Unfortunately, no direct proof of this fact is known. The known proof makes use of the Gallai-Edmonds Structure Theorem 3.2.1 which will be the main topic of the next section.

3.2. The Gallai-Edmonds Structure Theorem

We now present a result which has consequences in matching theory of the order of magnitude of Tutte's Theorem, but is by no means as well known. This is the so-called Gallai-Edmonds Structure Theorem. As we have remarked earlier, Tutte's Theorem provides us with a good characterization of graphs with no perfect matching. Namely, to exhibit that a graph G has no perfect matching we can present a set S of points and at least $|S|+1$ odd components in $G-S$. (Such a separating set is often called a **Tutte set**.) Of course there may be many such Tutte sets. Is any one of these to be preferred over all the others?

Gallai (1963b, 1964a) and Edmonds (1965a) have proved independently (and in quite different ways!) that there is one of these Tutte decompositions which is "canonical" in a sense. Moreover, in general much more information is available about the various pieces of this particular decomposition than is known for the others. Besides just the size of a maximum matching, we may well want to know other things. For example, we may want to know which lines lie in *some* maximum matching, which points are covered by *every* maximum matching, etc.

It is comforting to know that we can obtain this canonical decomposition efficiently (that is, in polynomial time) via the Edmonds Matching Algorithm. But we shall postpone the algorithmic aspects of this question until Chapter 9 when we study matching algorithms.

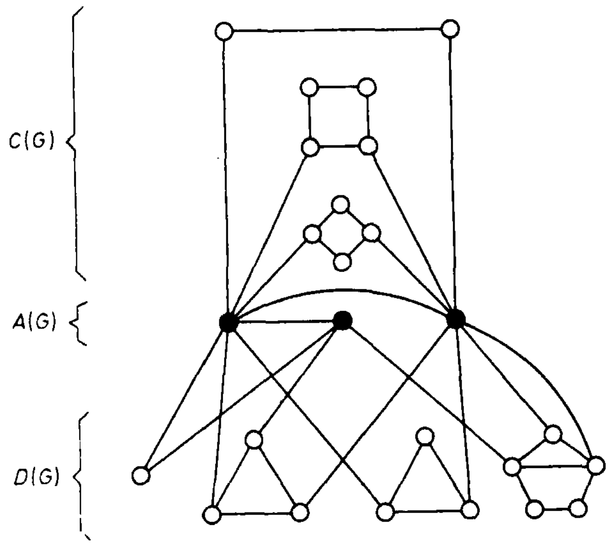


FIGURE 3.2.1. The Gallai-Edmonds decomposition of a graph G

First, let us determine the important properties of this canonical decomposition. Let G be any graph. Denote by $D(G)$ the set of all points in G which are not covered by at least one maximum matching of G. Let $A(G)$ be the set of points in $V(G) - D(G)$ adjacent to at least one point in $D(G)$. Finally let $C(G) = V(G) - A(G) - D(G)$.

The reader is invited to check the decomposition of the graph G in Figure 3.2.1. Note that $\nu(G) = 12$.

A **near-perfect** matching in a graph G is one covering all but exactly one point of G.

3.2.1. THEOREM. (*The Gallai-Edmonds Structure Theorem*).

If G is a graph and $D(G)$, $A(G)$, and $C(G)$ are defined as above, then:

(a) *the components of the subgraph induced by $D(G)$ are factor-critical,*

(b) *the subgraph induced by $C(G)$ has a perfect matching,*

(c) *the bipartite graph obtained from G by deleting the points of $C(G)$ and the lines spanned by $A(G)$ and by contracting each component of $D(G)$ to a single point has positive surplus (as viewed from $A(G)$),*

(d) *if M is any maximum matching of G, it contains a near-perfect matching of each component of $D(G)$, a perfect matching of each component of $C(G)$ and matches all points of $A(G)$ with points in distinct components of $D(G)$,*

(e) $\nu(G) = \frac{1}{2}\Big(|V(G)| - c\big(D(G)\big) + |A(G)|\Big)$, *where $c\big(D(G)\big)$ denotes the number of components of the graph spanned by $D(G)$.*

The statements in this theorem will follow relatively easily from part (a) of the following result which will be referred to later as the "Stability Lemma". Parts (b) and (c) will not be used until later, but it is natural to include them here.

3.2.2. LEMMA. (*The Stability Lemma*). *Let G be any graph and let $A(G)$, $C(G)$ and $D(G)$ be as defined above.*

(a) *Let $u \in A(G)$. Then $A(G-u) = A(G) - u$, $C(G-u) = C(G)$ and $D(G-u) = D(G)$.*

(b) *Let $u \in C(G)$. Then $A(G-u) \supseteq A(G)$, $C(G-u) \subseteq C(G) - u$ and $D(G-u) \supseteq D(G)$.*

(c) *Let $u \in D(G)$. Then $A(G-u) \subseteq A(G)$, $C(G-u) \supseteq C(G)$ and $D(G-u) \subseteq D(G) - u$.*

PROOF. (a) It clearly suffices to show $D(G-u) = D(G)$. Let M be a maximum matching of G. Then M covers u, since $u \notin D(G)$ (in fact, M covers $A(G)$). Hence $\nu(G-u) = \nu(G) - 1$. Moreover, if M is a maximum matching of G, $M - u$ is a maximum matching of $G - u$.

First we show $D(G) \subseteq D(G-u)$. Choose any $v \in D(G)$. Let M_v be a maximum matching of G which misses v. Then $M_v - u$ is a maximum matching of $G-u$ and moreover, $M_v - u$ misses v too, so $v \in D(G-u)$. Thus $D(G) \subseteq D(G-u)$.

To show $D(G-u) \subseteq D(G)$, choose a point $v \in D(G-u)$. Then there is a maximum matching M' of $G-u$ which misses v. Let w be any point in $D(G)$ adjacent to u in G and let M be a maximum matching of G which misses w. If M misses v, then $v \in D(G)$ follows, so suppose that M covers v. Consider $M \cup M'$. By definition, M' avoids v. Thus the component of $M \cup M'$ covering v must be a path P starting at v with a line of $M - M'$.

Suppose P ends with a line of M'. Then the symmetric difference of M and $E(P)$ is a new matching M'' in G which misses v. Moreover, $|M''| = |M|$ so M'' is a maximum matching and $v \in D(G)$.

So P ends with a line of M. Form yet another matching M_3, the symmetric difference of M' and $E(P)$. Then $|M_3| > |M'|$ and so $M_3 \not\subseteq E(G-u)$ by the maximality of M'. Thus P must end at u. But now the symmetric difference of M and $E(P+uw)$ is a maximum matching of G avoiding v and thus $v \in D(G)$.

(b) First note that $D(G) \subseteq D(G-u)$ follows by the same argument as that in assertion (a). To show $A(G) \subseteq A(G-u)$, let $x \in A(G)$. By (a), $C(G-x) = C(G)$ and hence $u \in C(G-x)$. Thus $\nu(G-x-u) = \nu(G-x)-1 = \nu(G)-2 = \nu(G-u)-1$, which implies that no maximum matching of $G-u$ can miss x; that is, $x \notin D(G-u)$. But since $x \in A(G-u)$ it has a neighbor in $D(G) \subseteq D(G-u)$ and so by definition of $A(G-u)$, $x \in A(G-u)$. Thus $A(G) \subseteq A(G-u)$. The fact that $C(G)-u \supseteq C(G-u)$ is immediate from the other two relations.

(c) To begin with, note that $\nu(G-u) = \nu(G)$ and thus every maximum matching of $G-u$ is a maximum matching of G as well. Hence $C(G) \subseteq C(G-u)$ and $D(G) \supseteq D(G-u)$ follow immediately. To show that $A(G) \supseteq A(G-u)$, let $x \in A(G-u)$. By definition, x is adjacent to some $y \in D(G-u)$. By definition there is a maximum matching of G missing u and y, but no maximum matching of G missing u and x. Also, trivially, no maximum matching of G can miss the adjacent points x and y. By Corollary 3.1.7 this may only happen if no maximum matching of G misses x, that is, if $x \in A(G) \cup C(G)$. Since x is adjacent to $y \in D(G-u) \subseteq D(G)$, it follows that $x \in A(G)$. This proves that $A(G-u) \subseteq A(G)$. ■

Now we turn to the proof of the Gallai-Edmonds Structure Theorem.

PROOF (**of Theorem 3.2.1**). Let us delete the points of $A = A(G)$ one by one. Lemma 3.2.2 implies that in every step the point deleted belongs to the set A of the graph present at that step and so

$$\nu(G-A) = \nu(G) - |A|. \tag{3.2.1}$$

Moreover,

$$C(G-A) = C(G) \text{ and } D(G-A) = D(G), \tag{3.2.2}$$

and if M is any maximum matching of G, then $M \cap E(G-A)$ is a maximum matching of $G-A$.

Denote by $G_1, \ldots, G_t$ the components of $G-A$ lying in D. Since no line joins D and C, these components partition D. Let H be the subgraph of G induced by C. By equations (3.2.2) every maximum matching of $G-A$ covers H, but each point of D will be missed by some such matching. Hence the maximum matchings of $G-A$ each consist of a perfect matching of H and near-perfect matchings M_i of each G_i. So H has a perfect matching which proves (b). Moreover, since each point of G_i is missed by some maximum matching of G_i, it is immediate by Gallai's Lemma that each G_i is factor-critical, proving (a).

Next we show (d) to be true. Since $M \cap E(G-A)$ is a maximum matching of $G-A$, the facts that $M \cap E(H)$ is a perfect matching of H and $M \cap E(G_i)$ is a near-perfect matching of G_i are immediate. It follows from equation (3.2.2) that $|M \cap E(G-A)| = |M| - |A|$ and hence M contains no line spanned by A. Since M covers all points of A, it matches A with $G-A$. But clearly it matches no point of A with a point in $C(G)$ and no two points of A to the same component G_i. This proves (d).

To show (c) true we need to prove that every nonempty set $X \subseteq A$ is adjacent to at least $|X|+1$ components G_i. Let G_h be any component adjacent to some point of X (since $X \subseteq A$, every point in X is adjacent to at least one point in $D(G)$, so such a G_h exists). Let u be an element of $V(G_i)$ and let M be a maximum matching missing u. By (d) already proven, M matches the points of X with different components of G_i, and in fact since $M \cap E(G_h)$ is a near-perfect matching of G_h, it follows that the points of X are matched with components G_i different from G_h. So together with G_h there are at least $|X|+1$ components G_i adjacent to X and (c) is proved.

Finally, (e) follows by equation (3.2.1) and parts (a) and (b) of the present theorem:

$$\begin{aligned}
\nu(G) &= \nu(G-A) + |A| \\
&= \sum_{i=1}^{t} \frac{|V(G_i)|-1}{2} + \frac{|C(G)|}{2} + |A| \\
&= \tfrac{1}{2}\big(|V(G)| - t + |A|\big). \qquad \blacksquare
\end{aligned}$$

Some important consequences of this theorem are immediate.

(1) From part (e) it is immediate that the deficiency of G, $\operatorname{def}(G) = c\big(D(G)\big) - |A(G)|$; that is, it is the number of components of $D(G)$ left unmatched into $A(G)$ by any arbitrary maximum matching of G.

(2) The non-trivial part of the proof of Tutte's Theorem is an easy corollary of the Gallai-Edmonds result, for suppose $c_o(G-X) \le |X|$ holds for all $X \subseteq V(G)$. Then in particular, $c_o\big(G - A(G)\big) \le |A(G)|$. Then substituting into the equation of part (e),

$$\begin{aligned}\nu(G) &= \tfrac{1}{2}\Big(|V(G)| - c\big(D(G)\big) + |A(G)| \Big) \\ &= \tfrac{1}{2}|V(G)|\end{aligned}$$

and hence G has a perfect matching.

Similarly, the non-trivial part of the proof of validity of Berge's Formula is a consequence of this theorem, since $X = A(G)$ gives equality in that formula.

(3) If G has no perfect matching, then every line incident with a point of $D(G)$ lies in some maximum matching of G. We need only the definition of $D(G)$ to see this. For let $u \in D(G)$ and uv any line meeting u. By definition there is a maximum matching M missing u. Hence by the maximality of M, there is a line vw in M which covers v. But then $M - vw + uv$ is the desired matching.

It follows from (d) that no line induced by $A(G)$ or connecting $A(G)$ to $C(G)$ belongs to any maximum matching. The lines induced by $C(G)$ show a much more complicated behavior in this respect and this problem will be one of our main concerns in Chapter 5.

(4) If G is itself a factor-critical graph, then we have a degenerate case of the Gallai-Edmonds decomposition; namely, $D(G) = V(G)$, so $A(G) = C(G) = \emptyset$.

(5) If G has a perfect matching we have a different degenerate case; namely, $D(G) = \emptyset$, so $A(G) = \emptyset$ and hence $C(G) = V(G)$.

(6) If G is a bipartite graph with bipartition (A, B) and if it has positive surplus (from A), then $A(G) = A$ and $D(G) = B$. In fact, trivially $\nu(G) = |A|$ and the P. Hall Theorem implies that $G - v$ has a complete matching of A into $B - v$ for every $v \in B$. Hence $B = D(G)$ and thus $A = A(G)$.

Thus the Gallai-Edmonds canonical decomposition is trivial in three cases: for factor-critical graphs, graphs which have perfect matchings and positive surplus bipartite graphs. On the other hand, it tells us that every graph can be built up from graphs of these three types. More precisely, we have the following result.

Let G_0 be a bipartite graph with bipartition (U, W) with positive surplus from U. For every $w \in W$, let G_w be a factor-critical graph. Let H be a factorizable graph. Define a graph G on the point set

$$V(G) = U \cup V(H) \cup \bigcup_{w \in W} V(G_w)$$

as follows. Keep all lines of the graphs H and G_w. For every line $uw \in E(G_0)$, $u \in U$, $w \in W$, add a line connecting u to an arbitrary point of G_w. Add lines connecting U to $V(H)$ and points of U to each other arbitrarily.

3.2.3. THEOREM. *The graph G constructed above has $A(G) = U$, $C(G) = V(H)$ and $D(G) = \bigcup_{w \in W} V(G_w)$. Every graph arises by the above construction uniquely.*

PROOF. Clearly, $\mathrm{def}(G) \geq c_o(G-U) - |U| = |W| - |U|$. On the other hand, select any matching M_0 of G_0 covering U and let M_0' be the set of corresponding lines of G. Let $x_w \in V(G_w)$ be the endpoint of the line of M_0' incident with G_w if such a line exists and an arbitrary point of G_w otherwise. Let M_w be a near-perfect matching of $G_w - x_w$. Let M'' be a near-perfect matching of H. Then $M = M_0' \cup M'' \cup \bigcup_{w \in W} M_w$ is a matching of G with defect $|W| - |U|$. Hence $\mathrm{def}(G) = |W| - |U|$ and M is a maximum matching. A similar argument shows that a point $u \in V(G)$ is missed by some maximum matching if and only if $u \in \bigcup_{w \in W} V(M_w)$. From this the first assertion follows.

The second assertion is an immediate consequence of the Gallai-Edmonds Structure Theorem. ∎

For a more in-depth study of the structure of maximum matchings we must thus treat three mutually disjoint classes of graphs separately: graphs with perfect matchings, factor-critical graphs and positive surplus bipartite graphs. Of these, graphs with perfect matchings have by far the most interesting structure and we shall study them in the next two chapters. Factor-critical graphs will be treated in Section 5.5. Positive surplus bipartite graphs were already discussed in Section 1.3 and we shall not have much more to say about them.

In the case of bipartite graphs, the canonical decomposition is notably simpler. This structure was worked out by Dulmage and Mendelsohn (1958, 1959, 1967) before the general result of Gallai and Edmonds. Their work relates the canonical structure to minimum point covers. Some of their results are summarized in the next theorem. Their decomposition theory for bipartite graphs with perfect matchings will be discussed in Section 4.3.

Let G be a bipartite graph with bipartition (U_1, U_2). We can think of $V(G)$ as being "doubly partitioned" as shown in Figure 3.2.2. Here the "vertical" partition is the canonical decomposition of Gallai-Edmonds and the "horizontal" one is just the given bipartition (U_1, U_2).

3.2.4. THEOREM. *Let bigraph $G = (U_1, U_2)$ and for $i = 1, 2$ let $A_i = A \cap U_i$, $C_i = C \cap U_i$ and $D_i = D \cap U_i$ where A, C and D are the three sets of the Gallai-Edmonds decomposition for G. Then*

(1) *$D = D_1 \cup D_2$ is an independent set of points,*
(2) *the subgraph $G[C_1 \cup C_2]$ has a perfect matching and hence $|C_1| = |C_2|$,*
(3) *$\Gamma(D_1) = A_2$ and $\Gamma(D_2) = A_1$,*
(4) *every maximum matching of G consists of a perfect matching of $G[C_1 \cup C_2]$, a matching of A_1 into D_2 and a matching of A_2 into D_1,*
(5) *if T is any minimum point cover for G,*

$$A_1 \cup A_2 \subseteq T \subseteq A_1 \cup A_2 \cup C_1 \cup C_2,$$

(6) *$C_1 \cup A_1 \cup A_2$ and $C_2 \cup A_1 \cup A_2$ are minimum point covers. Consequently, $A_1 \cup A_2$ is the intersection of all minimum point covers, and*
(7) *The subgraphs induced by $A_1 \cup D_2$ and $A_2 \cup D_1$ have positive surplus when viewed from A_1 and A_2 respectively.*

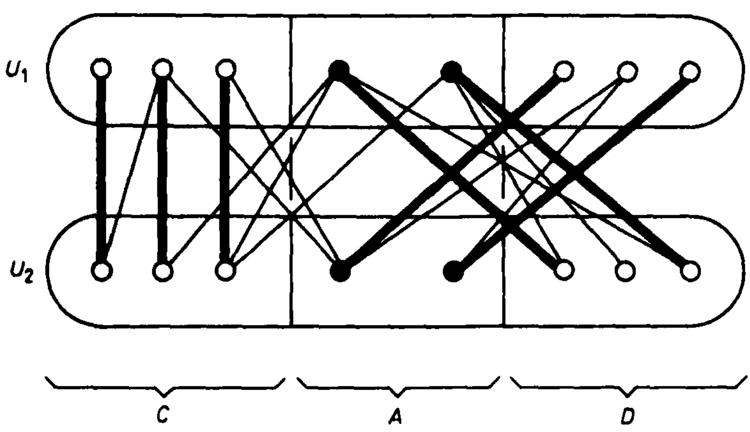


FIGURE 3.2.2.

PROOF. To get (1), observe that if any component of $G[D]$ were not a single point, it would be factor-critical and hence could not be bipartite, a contradiction.

Statements (2), (3) and (4) are immediate by the properties of the Gallai-Edmonds decomposition and the fact that G is bipartite.

To prove (5) suppose $a \in A_1 - T$. Then there is a point $d \in D$ such that line ad belongs to $E(G)$ and hence $d \in T$. Moreover, there is exactly one d in D adjacent to a. If there were more, say $\{d_1, \ldots, d_r\} \subseteq \Gamma_D(a)$,

$(r \geq 2)$, then $T - \{d_1, \ldots, d_r\} \cup \{a\}$ is a cover of G of smaller cardinality than T, a contradiction.

Now let M be a maximum matching of G which does not contain d. Then M covers a with a line in $G[A_1 \cup A_2]$ contradicting part (d) of the Gallai-Edmonds Theorem. Thus $A_1 \subseteq T$ and by a symmetric argument, $A_2 \subseteq T$.

Next suppose $T \cap D \neq \emptyset$, say $d \in T \cap D$. Suppose, without loss of generality, that $d \in D_2$. Then $\Gamma_G(d) \subseteq A_1 \subseteq T$. Hence $T - \{d\}$ is a cover for G, contradicting the minimality of T. Hence $T \cap D_2 = \emptyset$. Similarly, $T \cap D_1 = \emptyset$ and (5) is proved.

Assertion (6) follows from $\nu(G) = |C_1| + |A| = |C_1| + |A_1| + |A_2| = |C_2| + |A_1| + |A_2| = \tau(G)$ by König's Minimax Theorem. But then since $C_1 \cup A_1 \cup A_2$ and $C_2 \cup A_1 \cup A_2$ are point covers for G, they must be minimum point covers.

Finally, to prove (7), assume indirectly that for some non-empty $X \subseteq A_1$, say, $|X| \geq |\Gamma(X) \cap D_2|$. Then $T = (A_1 - X) \cup (\Gamma(X) \cap D_2) \cup A_2 \cup C_1$ is a cover for G, and since $|T| \leq |A_1| + |A_2| + |C_1| = \tau(G)$, it is a minimum cover. But this contradicts (5), since $A_1 \not\subseteq T$. ■

By Theorem 3.2.4 every bipartite graph has a unique decomposition into three graphs: $G_0 = G[C_1 \cup C_2]$, $G_1 = G[A_1 \cup D_2]$ and $G_2 = G[A_2 \cup D_1]$ together with additional lines between A_1 and A_2, A_1 and C_2, and between A_2 and C_1. These additional lines do not occur in any maximum matching. By (7) bigraphs G_1 and G_2 have positive surplus. Such graphs have been described in Section 1.3. Graph G_0, on the other hand, is a bigraph with a perfect matching. Such bigraphs do have an interesting further decomposition which will be discussed in the next chapter.

The following exercises illustrate several other applications of the Gallai-Edmonds Structure Theorem.

3.2.5. EXERCISE. Prove that every connected graph with a point-transitive automorphism group is either factor-critical or has a perfect matching. (Hint: all points must belong to the same class of the Gallai-Edmonds decomposition.)

3.2.6. EXERCISE. (Erdős and Gallai (1961).) Suppose G is a k-connected graph with $\nu(G) \leq p/2 - 1$. Then

(a) $\nu(G) \geq k$, and

(b) $\tau(G) \leq 2\nu(G) - k$.

Let us close this section by mentioning another problem — solved only recently — for which the Gallai-Edmonds Structure Theorem provides the natural setting. The problem is as follows. Give a *good* characterization (see Box 1A) of those graphs in which every matching extends to a maximum matching; that is, graphs in which every maxim*al* matching is maxim*um*. We shall call such graphs **equimatchable**. Since a graph is equimatchable if and only if all its connected components are equimatchable, we will restrict ourselves below to connected graphs.

Although the good characterization is quite tedious even to state in detail, the bottom line is that it gives rise to a *polynomial* algorithm for deciding membership in the class of equimatchable graphs. We give only a brief outline here; the reader is referred to Lesk, Plummer and Pulleyblank (1984) for details.

The class of equimatchable graphs having perfect matchings is quite restricted. Such graphs have been named **randomly matchable** by Sumner (1979) who proved the next result.

3.2.7. EXERCISE. A connected graph G is randomly matchable if and only if $G = K_{2n}$ or $G = K_{n,n}$ for some $n \geq 1$.

If equimatchable graph G does not have a perfect matching the problem becomes more interesting. Let $V(G) = D(G) \cup A(G) \cup C(G) = (D, A, C)$ be the Gallai-Edmonds decomposition for G. The following result can then be proved.

3.2.8. THEOREM. *Let G be a connected graph containing no perfect matching. Then G is equimatchable if and only if*

(1) *$C = \emptyset$ and A is independent,*
(2) *for every $v \in A$, graph $G - v$ has no matching covering $\Gamma(v)$,*
(3) *for every pair of points u, v in the same component of D with $uv \notin E(G)$, graph $G - u - v$ contains no matching covering $\Gamma(u) \cup \Gamma(v)$, and*
(4) *there are no two independent lines joining any component of D to A.* ∎

Using the result of Exercise 3.1.8, it is easy to see that this theorem gives a good characterization of equimatchable graphs. Using a polynomial matching algorithm (see Chapter 9) we can decide in polynomial time whether a graph is equimatchable.

3.3. Toward a Calculus of Barriers

Motivated by the Berge Formula 3.1.14, we define a **barrier** as a set $X \subseteq V(G)$ such that $c_o(G - X) = |X| + \text{def}(G)$. (Note that $\text{def}(G) \geq 0$.) A very large part of this book is concerned with maximum matchings, that is, those matchings which minimize the defect. (See the left hand side of the Berge Formula, Theorem 3.1.14.) Much less is known about the barriers, that is, those subsets of points which produce the maximum value on the right hand side of Berge's Formula, although even the superficial study of them described in this section yields some interesting results and some tools to be used in later chapters.

The empty set may or may not be a barrier in general. However, if G has a perfect matching, $\emptyset$ is a barrier. By the Gallai-Edmonds Structure Theorem 3.2.1, the set $A(G)$ is a barrier. In general $A(G)$ is not a unique barrier and one of our concerns will be to characterize $A(G)$ among all barriers.

The definition of barriers suggests the following question.

What happens to the deficiency function when we remove points from G?

3.3.1. LEMMA. *If G is any graph and $X \subseteq V(G)$, then* $\operatorname{def}(G-X) \leq \operatorname{def}(G)+|X|$.

PROOF. Let M be a maximum matching of G. Then $\nu(G)=|M|= |E(M)\cap E(G-X)|+|E(G)\cap E(X)|+|E(M)\cap \nabla(X)| \leq \nu(G-X)+|X|$ which, by the definition of deficiency, is equivalent to the inequality sought. ■

We will call a set of points in G **extreme** if equality holds in Lemma 3.3.1, that is, if $\operatorname{def}(G-X)=\operatorname{def}(G)+|X|$. Note that the empty set is always extreme as is the set $A(G)$ in the Gallai-Edmonds decomposition. We point out a few additional simple properties of extreme sets in the following series of exercises for the reader.

3.3.2. EXERCISE. Let G be any graph. If X is extreme in G and M is a maximum matching in G then M consists of a matching of all of X into $V(G)-X$ and a maximum matching of $G-X$. Thus no line in $G[X]$ lies in any maximum matching of G.

3.3.3. EXERCISE. In any graph every subset of an extreme set is extreme.

3.3.4. EXERCISE. Let G be any graph. If Y is an extreme set [barrier] in G and $X \subseteq Y$ then $Y-X$ is an extreme set [barrier] in $G-X$.

3.3.5. EXERCISE. Let G be any graph. If X is extreme in G and Z is extreme in $G-X$, then $X \cup Z$ is extreme in G.

3.3.6. EXERCISE. Let G be any graph. A singleton set $\{u\} \subseteq V(G)$ is extreme if and only if $u \in A(G) \cup C(G)$, where $A(G)$ and $C(G)$ are as in the Gallai-Edmonds decomposition of G.

3.3.7. EXERCISE. Let $e = xy$ be a line in G and suppose G has a perfect matching. Then $\{x, y\}$ is extreme in G if and only if e lies in no perfect matching of G.

Obviously every barrier is extreme. The following lemma describes a certain converse relationship.

3.3.8. LEMMA. *Let G be any graph. Then every extreme set of G lies in a barrier.*

PROOF. Let X be any extreme set in G and let $T = A(G-X) \cup X$. Then

$$\begin{aligned}
c_0(G-T) &= c_0(G-(A(G-X)\cup X)) \\
&= c_0(G-X-A(G-X)) \\
&= c_0(D(G-X)) \\
&= |A(G-X)| + \operatorname{def}(G-X), \\
&\text{(using part (e) of the Gallai-Edmonds Theorem)} \\
&= |A(G-X)| + \operatorname{def}(G) + |X|, \\
&\text{(since } X \text{ is extreme)} \\
&= |T| + \operatorname{def}(G)
\end{aligned}$$

and hence T is a barrier. ■

Let us note that the proof of the trivial half of Berge's Formula implies that every maximum matching covers every point of any barrier, and hence every barrier is contained in $A(G) \cup C(G)$. The following two results shows that every element of $A(G) \cup C(G)$ is contained in some barrier.

3.3.9. EXERCISE. Let G be any graph. Then if X is a barrier in G and $x \in X$, then $X - x$ is a barrier in $G - x$.

3.3.10. LEMMA. *Let G be any graph, $x \in A(G) \cup C(G)$, and let $G' = G - x$. Then $A(G') \cup \{x\}$ is a barrier in G. Moreover, if $x \in C(G)$ and B' is any barrier in G', then $B' \cup \{x\}$ is a barrier in G.*

PROOF. (i) Suppose $x \in A(G)$. Then $A(G') = A(G) - x$ by part (a) of the Stability Lemma 3.2.2 and hence $A(G') \cup \{x\} = A(G)$ is a barrier in G (since by the Gallai-Edmonds Theorem, $A(G)$ is *always* a barrier of G).

(ii) Suppose $x \in C(G)$. Then obviously $\mathrm{def}(G') = \mathrm{def}(G) + 1$. So $c_o(G - (B' \cup \{x\})) = c_o(G' - B') = |B'| + \mathrm{def}(G') = |B'| + \mathrm{def}(G) + 1 = |B' \cup \{x\}| + \mathrm{def}(G)$. ∎

In a sense barriers are analogues of the tight sets studied in Section 1.3. Lemma 1.3.3 provides us with a very handy "calculus" of tight sets. Is there any analogue of this for barriers? It seems that the situation here is considerably more complicated than it is for tight sets, and we are far from a description of the structure of all barriers.

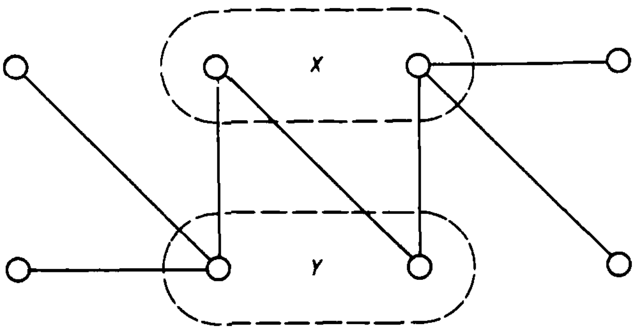


FIGURE 3.3.1. Two barriers whose union and intersection are not

The difficulty in handling barriers seems due to the fact that the intersection and union of barriers is not a barrier in general. In Figure 3.3.1, the sets X and Y are barriers, but neither their union nor their intersection is such. But let us see how much can be salvaged from the wreckage!

3.3.11. THEOREM. *The intersection of two (inclusionwise) maximal barriers is a barrier.*

PROOF. Let X and Y be two maximal barriers. Let $G_1, \ldots, G_k$ be the connected components of $G - X$ and $H_1, \ldots, H_m$ the connected components of $G - Y$. Note that by the maximality of X and Y, these components are odd, and since X and Y are barriers, $k = |X| + \mathrm{def}(G)$ and $m = |Y| + \mathrm{def}(G)$. Let $X_i = X \cap V(H_i)$, $Y_i = Y \cap V(G_i)$ and $A = X \cap Y$. Without loss of generality we may assume that $X_1, \ldots, X_{m_1} \neq \emptyset$, $Y_1, \ldots, Y_{k_1} \neq \emptyset$, but $X_{m_1+1} = \cdots = X_m = Y_{k_1+1} = \cdots = Y_k = \emptyset$, and also that $k_1 \leq m_1$.

Observe that $G_{k_1+1}, \ldots, G_k$ are components of $G-X-Y$ and hence each of them is contained in a connected component of $G-Y$. If $G_i \subseteq H_j$ where $k_1+1 \le i \le k$, $m_1+1 \le j \le m$, then $G_i = H_j$ and G_i is a connected component of $G-A$. The number of such G_i's is at most $|A|+\mathrm{def}(G)$, by the definition of $\mathrm{def}(G)$. Consider now those components G_i which are contained in an H_j, $1 \le j \le m_1$. Such a G_i is a component of $H_j - X_j$. Note that by the maximality of Y,

$$c_o(H_j - X_j) \le |X_j|$$

(since otherwise $X_j \cup Y$ would be a larger barrier). Since H_j has an odd number of points, the usual parity argument yields

$$c_o(H_j - X_j) \le |X_j| - 1.$$

So H_j contains at most $|X_j| - 1$ components G_i $(k_1 + 1 \le i \le k)$. So the total number of components G_i disjoint from Y is at most

$$\begin{aligned} |A| + \mathrm{def}(G) + \sum_{j=1}^{m_1} \big(|X_j| - 1\big) &= |A| + \mathrm{def}(G) + |X - A| - m_1 \\ &= |X| + \mathrm{def}(G) - m_1 \\ &= k - m_1 \\ &\le k - k_1. \end{aligned}$$

Since this number is, obviously, exactly $k-k_1$, we infer that equality must hold throughout and in particular $G-A$ must have exactly $|A| + \mathrm{def}(G)$ odd components. Thus A is a barrier. ∎

Note that we have not used any part of the preceding theory in the proof of Theorem 3.3.11. In fact, a proof of Tutte's Theorem can be based on this result.

3.3.12. EXERCISE. Let G be a graph such that $c_o(G-X) \le |X|$ holds true for every $X \subseteq V(G)$, but deleting any line of G destroys this property. Prove that G consists of disjoint lines. Give a new proof of Tutte's Theorem based on this result. (Difficult!)

For which graphs can a result similar to that of Theorem 3.3.11 be obtained for *arbitrary* barriers? The following theorem and corollary, which will have an important application later in Chapter 5, provides such a class of graphs.

3.3.13. THEOREM. *Let G be a graph such that $C(G) = \emptyset$ and let X and Y be barriers such that no line connects $X - Y$ to $Y - X$. Then $X \cap Y$ and $X \cup Y$ are barriers.*

PROOF. The condition on X and Y implies, by Lemma 1.3.2, that

$$c(G - (X \cap Y)) + c(G - (X \cup Y)) \geq c(G - X) + c(G - Y).$$

Let $G_1, \ldots, G_k$ be the connected components of $G - (X \cap Y)$. Since $X \cap Y \subseteq A(G)$ it follows by the Stability Lemma 3.2.2 that $C(G - (X \cap Y)) = C(G) = \emptyset$, and hence none of $G_1, \ldots, G_k$ can have a perfect matching. Thus

$$\operatorname{def}(G - (X \cap Y)) = \sum_{i=1}^{k} \operatorname{def}(G_i) \geq k = c(G - (X \cap Y)).$$

Similarly,

$$\operatorname{def}(G - (X \cup Y)) \geq c(G - (X \cup Y)).$$

Hence,

$$\begin{aligned} |X \cap Y| + |X \cup Y| + 2\operatorname{def}(G) &\geq \operatorname{def}(G - (X \cap Y)) + \operatorname{def}(G - (X \cup Y)) \\ &\geq c(G - (X \cap Y)) + c(G - (X \cup Y)) \\ &\geq c(G - X) + c(G - Y) \\ &\geq c_o(G - X) + c_o(G - Y) \\ &= |X| + |Y| + 2\operatorname{def}(G). \end{aligned}$$

Since the first and last expressions are equal, we must have equality throughout. In particular, $c(G - (X \cap Y)) = \operatorname{def}(G - (X \cap Y)) = |X \cap Y| + \operatorname{def}(G)$. But the first equality implies that $\operatorname{def}(G_i) = 1$ for every component G_i of $G - (X \cap Y)$, and so all components of $G - (X \cap Y)$ are odd. So $c_o(G - (X \cap Y)) = c(G - (X \cap Y)) = |X \cap Y| + \operatorname{def}(G)$; that is, $X \cap Y$ is a barrier. Similarly, $X \cup Y$ is a barrier. ∎

3.3.14. COROLLARY. *Let G be a graph in which $C(G) = \emptyset$ and $A(G)$ is an independent set of points. Then the union and intersection of any two barriers is a barrier.*

PROOF. If X and Y are any two barriers, each lies in $A(G) \cup C(G) = A(G)$ and hence $X \cup Y$ is independent in G. The result follows immediately from Theorem 3.3.13. ∎

Maximal barriers yield a nice characterization of $A(G)$.

3.3.15. THEOREM. *$A(G)$ is the intersection of all maximal barriers in graph G.*

PROOF. Let X be any maximal barrier. The maximality of X implies that $A(G-X)=\emptyset$ (otherwise $X\cup A(G-X)$ would be a barrier properly containing X). By the Stability Lemma 3.2.2, we have $A(G)\subseteq A(G-X)\cup X$, and so $A(G-X)=\emptyset$ implies $A(G)\subseteq X$. Since this is true for every maximal barrier X, it follows that the intersection of all maximal barriers contains A.

To complete the proof we have to show that if $x\in V(G)-A(G)$ then there exists a maximal barrier which does not contain x. If $x\in D(G)$ then we already know that no maximal barrier contains x. So suppose $x\in C(G)$. Let M be any maximum matching and let xy be the line of M incident with x. Then $A(G)\cup\{y\}$ is, obviously, a barrier. Let X be a maximal barrier containing $A(G)\cup\{y\}$. Again, as in the proof of Berge's Formula, X cannot span a line of M, and so $x\notin X$. ∎

3.3.16. EXERCISE. Show by an example that $A(G)$ is not necessarily contained in every barrier of graph G.

The set $A(G)$ is itself a barrier. How may we distinguish it among all barriers? The previous proof shows that if $C(G)\neq\emptyset$, then $A(G)$ is not a maximal barrier. Figure 3.3.2 shows that, in general, $A(G)$ is not a minimal barrier. So to characterize $A(G)$ as a barrier with certain extremal properties, we need to consider a slightly more complicated notion. Let $D_1(X)$ denote the union of odd components of $G-X$ (for $X\subseteq V(G)$). The next theorem says that $A(G)$ is the barrier X for which the set $D_1(X)$ is (inclusionwise) minimal.

3.3.17. THEOREM. *For every barrier X in a graph G, $D_1(X)\supseteq D_1(A(G))=D(G)$.*

PROOF. It is evident that $D_1(A(G))=D(G)$. Let $x\in D(G)$ and let M be a maximum matching missing x. By the proof of Berge's Formula, the connected component G_0 of $G-X$ containing x must have exactly one point not covered by M, and no line of M can connect G_0 to X. Hence G_0 must have an odd number of points and so $x\in D_1(X)$. ∎

3.3.18. EXERCISE. Let G be any graph and X a maximal barrier in G.

(a) Prove, using P. Hall's Theorem, that there exists a matching M in G which matches each point of X with points in different components of $G-X$.

(b) Prove, using Exercise 3.1.10, that every component of $G-X$ is factor-critical.

(c) Conversely, show that any set X satisfying (a) and (b) is a maximal barrier.

(d) Give a new proof of Berge's Formula from the above.

(e) Try to eliminate the use of Exercise 3.1.10 from the proof in (d) by using induction instead.

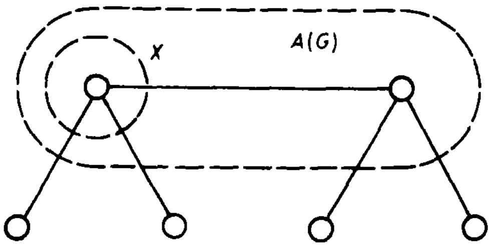


FIGURE 3.3.2. A barrier X properly contained in $A(G)$

We conclude with a property of *minimal* barriers. A similar property of minimal Tutte sets was observed by Sumner (1974, 1976). In fact the proofs and implications are identical. A **claw** is an induced subgraph isomorphic to the complete bipartite graph $K_{1,3}$. Claw-free graphs will also be studied in connection with vertex packing in Section 12.4.

3.3.19. LEMMA. *Let X be an (inclusionwise) minimal barrier in a graph G. Then every point in X is the center of a claw in G.*

PROOF. Assume that $x \in X$ is not the center of a claw. Then x is connected to at most two components of $G-X$. But then $c_o(G-(X-x)) = c_o(G-X)-1$ and therefore $X-x$ is a barrier, contradicting the minimality of X. ∎

Sumner used the previous observation to prove the following fact, observed independently by Las Vergnas (1975):

3.3.20. THEOREM. *Every connected claw-free graph with an even number of points contains a perfect matching.*

PROOF. By Lemma 3.3.19 every minimal barrier of G is empty. But by the assumptions on G, $c_o(G) = 0$ and so $\text{def}(G) = 0$. ∎

Jünger, Pulleyblank and Reinelt (1983) have recently extended the above result by proving that if G has an *odd* number of points and is claw-free, then G contains a near-perfect matching.

3.3.21. EXERCISE. Let G be a connected graph which does not contain a "cherry", that is, two points of degree one with a common neighbor. Prove that G has two adjacent points whose removal does not disconnect G. Give a new proof of Theorem 3.3.20 based on this fact. (See Lovász (1979c).)

3.4. Sufficient Conditions for Matchings of a Given Size

Let us reconsider the question 'When does a graph possess a perfect matching?'. Of course the theorem of Tutte gives necessary and sufficient conditions for a perfect matching to exist. However, a number of interesting results exist in which sufficient (but not necessary) conditions are provided. Not surprisingly, most of the proofs involved are indirect and hence start with a Tutte decomposition of a graph assumed to have no perfect matching. Then a counting argument leads to a contradiction. Modulo this idea, these arguments usually become quite straightforward, though sometimes tedious.

In his pioneering 1891 paper (already mentioned in the historical remarks in the Preface to this book), Petersen proved the following classical result.

3.4.1. THEOREM. *Every connected cubic graph with no more than two cutlines has a perfect matching.*

Let us note at the onset that Petersen's original proof was somewhat tedious and shorter proofs were subsequently given by Brahana (1917-18), by Frink (1925-26, 1926) and by König (1936). However, with Tutte's Theorem in hand the following even shorter proof is possible.

PROOF. Let G be a connected cubic graph with no more than two cutlines. (We remind the reader that graph G need not be simple.) Of course $|V(G)|$ is even. Suppose indirectly that G has no perfect matching. Then G contains a cutset S with $c_o(G - S)$ odd components and $|S| \leq c_o(G - S) - 2$. Since G is cubic each odd component is joined to S by some odd number of lines. If this odd number is one, the corresponding line must be a cutline. By hypothesis there are at most two such in G. Hence the odd components of $G-S$ send at least $3(c_o(G-S)-2)+2 = 3c_o(G-S)-4 \geq 3(|S|+2)-4 = 3|S|+2$ lines to S. But of course S sends at most $3|S|$ lines to the odd components and a contradiction results. ∎

Note that two cutlines is "best possible" in the sense that there are cubic multigraphs with three cutlines, but no perfect matching (cf. the graph in Figure 3.4.1).

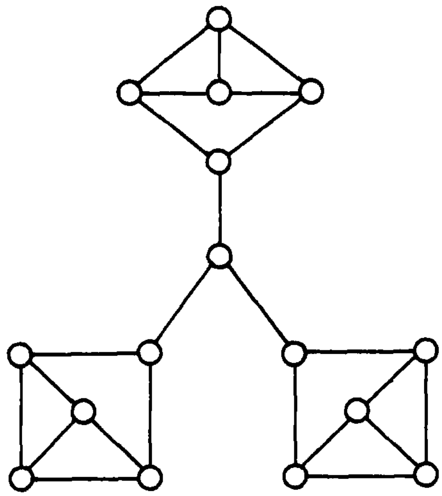


FIGURE 3.4.1.

This kind of argument can be — and has been — generalized to more complicated situations. Here is a representative result of this kind, still dealing with regular graphs.

3.4.2. THEOREM. (*Plesník* (1972)). *Let G be an r-regular graph with an even number of points which is $(r-1)$-line-connected. Then, if any $r-1$ lines are deleted from G, the resulting graph has a perfect matching.*

PROOF. Let G' denote the graph resulting from G upon deletion of some $r-1$ lines and suppose, indirectly, that G' has no perfect matching. Then by Tutte's Theorem there is a set S with $c_o(G'-S) > |S|$ and hence by parity — since $|V(G')|$ is even — $c_o(G-S) \geq |S|+2$. Let $G_1,\ldots,G_k$

denote the odd components and $G_{k+1},\ldots,G_t$, the even components of $G'-S$.

Now let us count lines. To this end let α_i be the number of lines in $E(G)-E(G')$ joining G_i to S, and β_i, those in $E(G)-E(G')$ joining G_i to other G_j's, and γ_i those lines of G' between G_i and S. So $\alpha_i+\beta_i+\gamma_i$ is the total number of lines joining G_i to $G-G_i$ and hence by hypothesis is at least $r-1$.

We now claim if G_i is an odd component that, in fact, $\alpha_i+\beta_i+\gamma_i \geq r$. Suppose not; that is, suppose $\alpha_i+\beta_i+\gamma_i = r-1$. Then the sum of the degrees of subgraph G_i is $r|V(G_i)|-(r-1) = r\big(|V(G_i)|-1\big)+1$ which is an odd number contradicting the fact that G_i is odd. So the claim is proved and we have the inequality:

$$\sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \beta_i + \sum_{i=1}^{t} \gamma_i \geq rt. \tag{3.4.1}$$

Counting all lines removed from G to get G' we have $\sum_{i=1}^{t} \alpha_i + \frac{1}{2}\sum_{i=1}^{t} \beta_i$ and so:

$$2\sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \beta_i \leq 2(r-1). \tag{3.4.2}$$

On the other hand, there are altogether $\sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \gamma_i$ lines entering S and hence:

$$\sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \gamma_i \leq r|S|. \tag{3.4.3}$$

Adding inequalities (3.4.2) and (3.4.3), we obtain

$$3\sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \beta_i + \sum_{i=1}^{t} \gamma_i \leq r\big(|S|+2\big)-2. \tag{3.4.4}$$

Now comparing inequalities (3.4.1) and (3.4.4) we have:

$$\begin{aligned} rt &\leq \sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \beta_i + \sum_{i=1}^{t} \gamma_i \\ &< 3\sum_{i=1}^{t} \alpha_i + \sum_{i=1}^{t} \beta_i + \sum_{i=1}^{t} \gamma_i \\ &\leq r\big(|S|+2\big)-2 < r\big(|S|+2\big), \end{aligned}$$

so $t < |S|+2$, a contradiction. ∎

3.4.3. COROLLARY. *If G is an r-regular, $(r-1)$-line-connected graph on an even number of points then every line of G is contained in a perfect matching.*

PROOF. If $e = uv$ and if the other lines at u are $e_1, \ldots, e_{r-1}$, consider a perfect matching of $G - e_1 - \cdots - e_{r-1}$. ∎

Let us remark here that for the first time we have met a class of graphs which have the property that each line lies in a perfect matching. We shall call such a line **allowed** and any line which is not allowed will be called **forbidden**. Graphs in which every line is allowed are called **1-extendable** (or sometimes **matching covered**). They will be studied extensively in the next three chapters.

Let us also remark in passing that we have really proved a bit more than claimed in Theorem 3.4.2. In particular, we do not need the full strength of the assumption that G is $(r-1)$-line-connected.

3.4.4. EXERCISE. (Cruse (1977)). Show that the conclusion of Theorem 3.4.2 still holds for r-regular graphs if we omit the assumption that G is $(r-1)$-line-connected and replace it by the weaker assumption that any odd set X, $|X| \geq 3$, is joined to $G - X$ by at least $r-1$ lines.

In view of the fact that the Berge Formula is a positive deficiency version of Tutte's Theorem, it is not surprising that there are positive deficiency analogues of some of the above results. (See, for example, the results on graphs with transitive automorphism groups to be found in Section 5.5.) See also Little, Grant and Holton (1975, Theorem 3.1, 1976) and Plesník (1979, Theorem 1).

Generalizations in other directions also exist. For example, Plesník (1979) addresses the problem of finding matchings with defect d in r-regular, $(r-1)$-line-connected graphs missing sets of arbitrarily prescribed points *and* lines. In a somewhat different direction, Chartrand and Nebeský (1979) and Chartrand, Goldsmith and Schuster (1979) study sufficient conditions for perfect matchings in r-regular, but only $(r-2)$-line-connected, graphs. The interested reader is referred to these papers as the hypotheses of these theorems are too complicated to be stated here.

Now let us further explore the question of how large a matching can be "in general". But of course the question itself as posed is "too general". How large can a matching be in terms of what? Among the more reasonable graph parameters which come to mind immediately are number of points, number of lines, minimum degree, maximum degree, point connectivity and line connectivity.

3.4.5. EXERCISE. Prove that if a simple graph G on an even number of points p has more than $\binom{p-1}{2}$ lines, then it has a perfect matching.

One approach, apparently first posed in the language of set families by Erdős and Rado (1960) and continued in papers of Abbott (1966), Chvátal (1970), Abbott, Hanson and Sauer (1972) and finally Chvátal and Hanson (1976), proceeds as follows. Let $f(p,\nu,\Delta)$ denote the maximum number of lines which may be contained in any simple graph on p points having matching number ν and maximum degree Δ. Chvátal and Hanson obtain exact values for $f(p,\nu,\Delta)$, although the results are quite tedious to state. If the reader will bear with us, however, we feel at least once in this section we should be explicit!

3.4.6. THEOREM. *Let p, ν, Δ be positive integers with $p \geq 2\nu+1$ and let $f(p,\nu,\Delta)=f$. Then:*
(a) *If $\Delta \leq 2\nu$ and $p \leq 2\nu + \lfloor \nu / \lfloor (\Delta+1/2) \rfloor \rfloor$, then*

$$f = \begin{cases} \min\{\lfloor p\Delta/2 \rfloor,\ \nu\Delta + \lfloor 2(p-\nu)/(\Delta+3) \rfloor ((\Delta-1)/2)\}, & \textit{if } \Delta \textit{ is odd,} \\ p\Delta/2, & \textit{if } \Delta \textit{ is even.} \end{cases}$$

(b) *If $\Delta \leq 2\nu$ and $p \geq 2\nu + \lfloor \nu / \lfloor (\Delta+1)/2 \rfloor \rfloor$, then*

$$f = \nu\Delta + \lfloor \nu / \lfloor (\Delta+1)/2 \rfloor \rfloor \lfloor \Delta/2 \rfloor, \textit{and}$$

(c) *if $\Delta \geq 2\nu+1$, then*

$$f = \begin{cases} \max\left\{ \binom{2\nu+1}{2}, \left\lfloor \dfrac{\nu(p+\Delta-\nu)}{2} \right\rfloor \right\}, & \textit{if } p \leq \nu+\Delta, \\ \nu\Delta, & \textit{if } p \geq \nu+\Delta. \end{cases}$$

PROOF. The proof uses the Berge Formula and careful counting to show that f does not exceed the right hand side and then extremal graphs are constructed to show that equality holds. ∎

To get a bit more of the flavor here, however, let us consider a simple proof due to Bollobás (1977) of part of Theorem 3.4.6(c). In particular, let us prove that if $\Delta \geq 2\nu+1$, then $f(p,\nu,\Delta) \leq \nu\Delta$. So consider a graph G on p points with $\Delta(G)=\Delta$. By Berge's Formula there is a cutset S in G such that $c_o(G-S) = p+s-2\nu$ where $s=|S|$. Then clearly $s \leq \nu$. It is an easy exercise to see that the number of lines in $G-S$ cannot exceed the number of lines of the graph on $V(G)-S$ which consists of one "large" complete component and $c_o(G-S)-1$ isolated points.

Now there are $p-2\nu+s-1$ singletons and hence the one "large" odd component must have $p-s+p-2\nu+s-1=2\nu-2s+1$ points. But then $G-S$ contains at most $\binom{2\nu-2s+1}{2}=(\nu-s)(2\nu-2s+1)$ lines and hence G itself has at most $s\Delta+(\nu-s)(2\nu-2s+1)=s(2s+\Delta-4\nu-1)+2\nu^2+\nu$ lines. Since $1\le s\le\nu$, it is easy to verify that the maximum of this expression occurs when $s=\nu$ and is $\nu\Delta$ as claimed.

Generalizing some earlier work of Weinstein (1963, 1974) Bollobás and Eldridge (1976) have obtained more "finely tuned" results, but the results are even more cumbersome to state than those of Chvátal and Hanson so we mercifully omit a complete treatment! Suffice it to say that these two authors define and determine functions $m(p,\delta,\Delta)$ and $m(p,\delta,\Delta,\kappa)$ and $m(p,\delta,\Delta,\lambda)$ where κ and λ are the point and line connectivity and where, for example, $m(p,\delta,\Delta,\lambda)$ is the minimum size of a maximum matching in any graph with p points, minimum and maximum degrees δ and Δ respectively and line connectivity λ.

Now what other parameters might be useful in estimating the size of a maximum matching? I. Anderson (1973) introduced the new idea of the *binding number* of a graph. In particular, the **binding number** of G, $\mathrm{bind}(G)=\min\{|\Gamma(X)|/|X| \mid \emptyset\ne X\subseteq V(G) \text{ and } \Gamma(X)\ne V(G)\}$. Anderson proved that any graph with binding number at least $\frac{4}{3}$ has a perfect matching. Not long thereafter, Woodall (1973) generalized Anderson's result to obtain a lower bound in terms of the binding number for the size of a maximum matching in graphs which do not necessarily have perfect matchings. (In the same paper Woodall proved the beautiful theorem which says that any graph G with $\mathrm{bind}(G)\ge\frac{3}{2}$ has a Hamilton cycle.)

We present a result which generalizes Anderson's theorem slightly. In point of fact, however, the proof of this result is essentially identical with that of Anderson.

3.4.7. THEOREM. *Let G be a graph on an even number of points. If*

$$|\Gamma(X)|\ge\min\left\{|V(G)|,\tfrac{4}{3}|X|-\tfrac{2}{3}\right\} \text{ for all } X\subseteq V(G),$$

then G has a perfect matching.

PROOF. Suppose, to the contrary, that G has no perfect matching. Then there exists a set $S\subseteq V(G)$ with $|S|<c_o(G-S)$. Let $p=|V(G)|$ and $k=|S|$. Since p is even, by parity we may assume $c_o(G-S)\ge k+2$. Let m denote the number of singleton components of $G-S$. (Note that $\Gamma(V(G)-S)=V(G)$ if and only if $m=0$.)

Case 1. Suppose $m > 0$. Then $|\Gamma(V(G) - S)| < p$ and hence by hypothesis $|\Gamma(V(G) - S)| \geq \frac{4}{3}|V(G) - S| - \frac{2}{3}$ so $\frac{4}{3}(p - k) - \frac{2}{3} \leq |\Gamma(V(G) - S)| \leq p - m$. From this we have:

$$p + 3m \leq 4k + 2. \tag{3.4.5}$$

On the other hand, counting points in the odd components of $G - S$, we have $m + 3(k + 2 - m) \leq p - k$ and hence

$$4k - 2m + 6 \leq p. \tag{3.4.6}$$

From inequality (3.4.5) we have

$$4k - 3m + 2 \geq p. \tag{3.4.7}$$

Combining inequalities (3.4.6) and (3.4.7), we obtain $4k - 2m + 6 \leq 4k - 3m + 2$ or $m \leq 2 - 6 < 0$, a contradiction.

Case 2. Suppose $m = 0$ (and hence all odd components have at least three points). Let X be the union of any $k + 1$ of the odd components of $G - S$. Since X fails to contain at least one odd component of $G - S$, we have $|\Gamma(X)| < p$ and hence $|\Gamma(X)| \geq \frac{4}{3}|X| - \frac{2}{3}$. On the other hand, $|\Gamma(X)| \leq |X| + k$, so $|X| + k \geq \frac{4}{3}|X| - \frac{2}{3}$ and hence $3k \geq |X| - 3(\frac{2}{3})$. But $|X| \geq 3(k + 1)$ as well, so $3k \geq 3(k + 1) - 3(\frac{2}{3}) = 3k + 1 > 3k$, a contradiction. ∎

We now state Woodall's generalization of Anderson's theorem. As Woodall himself points out, his proof is completely analogous to that of Anderson, except once again Berge's Formula is used instead of Tutte's Theorem. The proof is left as an exercise.

3.4.8. THEOREM. *Let G be any graph with p points with* $\operatorname{bind}(G) = c$. *Then:*

$$\nu(G) \geq \begin{cases} pc/(c+1), & \text{if } 0 \leq c \leq \frac{1}{2}, \\ p/3, & \text{if } \frac{1}{2} \leq c \leq 1, \text{and} \\ p(3c - 2)/3c - 2(c - 1)/c, & \text{if } 1 \leq c \leq \frac{4}{3}. \end{cases}$$ ∎

Note that if G has an even number of points, Anderson's theorem follows from the last inequality above, for then $p(3c - 2)/3c = p/2$ and $2(c - 1)/c < 1$.

Let us turn our attention to yet another recently formulated parameter. A graph is "tough" if relatively few components result from the deletion of any cutset of points. More precisely, the **toughness** of G, $t(G)$ is defined to be $+\infty$, if $G = K_n$ and to be $\min(|S|/c(G-S))$, if $G \neq K_n$. Here the minimum is taken over all point cutsets S of G and $c(G-S)$ denotes the number of components of $G-S$.

The formulation of this concept is due to Chvátal (1973c) who was motivated by certain studies on Hamilton cycles in graphs. It is clear that if G is Hamiltonian then $c(G-S) \leq |S|$ for all point cutsets S in G. Hence every Hamiltonian graph G has toughness $t(G) \geq 1$.

The idea of toughness sounds temptingly similar to that of binding number and at first thought we might expect a close connection between them. Indeed Woodall (1973) proves the following elementary result which the reader may verify as an exercise.

3.4.9. THEOREM. *For any graph G,* $\operatorname{bind}(G) \leq t(G) + 1$. ∎

But binding number and toughness are not so similar as we might think! We know (Woodall (1973, Theorem 12)) that every graph G with $\operatorname{bind}(G) \geq \frac{3}{2}$ has a Hamilton cycle. Compare this result with the following conjecture by Chvátal (1973b) which remains unsettled at the time of this writing.

3.4.10. CONJECTURE. *There exists a positive real number t_0 such that for every graph G, $t(G) \geq t_0$ implies G is Hamiltonian.*

It is known that $t_o = \frac{3}{2}$ is not large enough here, for if we construct a thirty point cubic graph H from the Petersen graph by replacing each point by a triangle, H then has no Hamilton cycle, but $t(G) = \frac{3}{2}$. Chvátal conjectures that $t_o = 2$ will, in fact, be sufficient.

The analogous connection between toughness and perfect matchings is, however, much easier to settle.

3.4.11. EXERCISE. (a) Prove that if $t(G) \geq 1$ then G has a perfect matching.

(b) Show that for every $\epsilon > 0$, there are graphs $G(\epsilon)$ such that $t(G(\epsilon)) > 1 - \epsilon$, but $G(\epsilon)$ has no perfect matching.

Before bringing these sufficiency proceedings to a close, let us exhume one last parameter, the **genus** of G, $\gamma(G)$ (i.e., the minimum genus of an orientable surface into which G can be embedded; see A. White (1973)). Together with (point) connectivity, this parameter can yield yet another sufficient condition for a perfect matching. First we need the following well-known result. The proof is once more remanded to the custody of the reader. (See Harary (1969)).

3.4.12. LEMMA. *Let G be a simple bipartite graph with $p \geq 3$ points, q lines and genus $\gamma(G)$. Then*

$$q \leq 2p + 4\gamma(G) - 4.$$

■

Nishizeki (1978, 1979) has shown the following to be true.

3.4.13. THEOREM. *If $k = \kappa(G) \geq 4$, if $p = |V(G)|$ is even and if $\gamma(G) < k(k-2)/4$, then G has a perfect matching.*

PROOF. We use Tutte's criterion. Let $X \subseteq V(G)$, we show that $c_o(G - X) = t \leq |X|$ $(= s)$. This is trivial if X does not separate the graph, so assume that X is a separating set. Then $s \geq k$ by hypothesis.

Form a simple bigraph G' with $p' = s + t$ points and some q' lines by contracting every odd component of $G - X$ to a single point, deleting all even components of $G - X$ and all lines spanned by X and identifying parallel lines. Clearly $\gamma(G') \leq \gamma(G)$.

If x is a point arising from the contraction of an odd component of $G - X$, then by the k-connectivity of G, $\deg_{G'}(x) \geq k$. Hence

$$q' \geq kt. \tag{3.4.8}$$

On the other hand, Lemma 3.4.12 implies

$$\begin{aligned} q' &\leq 2p' + 4\gamma(G') - 4 \\ &= 2(s+t) + 4\gamma(G') - 4 \\ &< 2(s+t) + k(k-2) - 4. \end{aligned} \tag{3.4.9}$$

By inequalities (3.4.8) and (3.4.9) we have

$$kt < 2(s+t) + k(k-2) - 4,$$

and hence

$$t < \frac{2}{k-2}s + k - \frac{4}{k-2}.$$

Thus

$$\begin{aligned} t - s &< \left(\frac{2}{k-2} - 1\right)s + k - \frac{4}{k-2} \\ &\leq \left(\frac{2}{k-2} - 1\right)k + k - \frac{4}{k-2} \\ &= 2. \end{aligned}$$

Since $t - s$ is an even integer, this implies that $t - s \leq 0$, that is, $t \leq s$ as claimed. ■

3.4.14. COROLLARY. *Every 4-connected graph which has an even number of points and is embeddable on the torus has a perfect matching.* ∎

It is interesting to compare this corollary with the celebrated result of Tutte which states that every 4-connected planar graph has a Hamilton cycle. The analogous result is not known to hold for the torus (Grünbaum (1970b)).

3.4.15. EXERCISE. If G is planar with $\min \deg(G) \geq 3$ and $p \geq 10$, then

$$\nu(G) \geq \frac{1}{3}(p+2).$$

For reference see Nishizeki and Baybars (1979).

3.4.16. EXERCISE. (a) (Berge (1973)). Let G be any simple graph with an even number of points and having the property that any two point-disjoint odd cycles in G are joined by a line. Prove: G has a perfect matching if and only if $|X| \leq |\Gamma(X)|$ for every $X \subseteq V(G)$.

(b) (Fulkerson, Hoffman and McAndrew (1965)). Prove: if G is r-regular, has an even number of points and if any two point-disjoint cycles are joined by a line, then G has a perfect matching.

This Page Intentionally Left Blank

4

Bipartite Graphs with Perfect Matchings

4.0. Introduction

In Chapter 3 we presented the Gallai-Edmonds Structure Theorem in order to learn more about the maximum matchings of a graph. But we then saw that if we are interested, say, in the number of such matchings, we must first settle the same question for the more special class consisting of graphs with perfect matchings. The subgraph $C(G)$ of the Gallai-Edmonds decomposition is always such a graph, and for graphs with perfect matchings $C(G) = G$. So in this case, the Gallai-Edmonds result is of no help to us. We must then take up the study of graphs with perfect matchings separately and we shall do so in this chapter and the next. The present chapter is devoted to bipartite graphs.

So let us consider a bipartite graph G with a perfect matching. In the investigation of many questions of interest to us (for example, the number of perfect matchings or the number of linearly independent perfect matchings), those lines of G which do *not* occur in any perfect matching play no role. If we delete these lines we are left with a graph in which every line occurs in a perfect matching. We may restrict our attention to the connected components of this graph. Such components are called "elementary" graphs.

In the present chapter we shall begin by investigating the structure and properties of elementary bipartite graphs. We shall introduce a new kind of decomposition, called **ear decomposition**, which will turn out to be a powerful tool in the study of the structure of matchings. For example, ear decompositions help us to show that every elementary bipartite graph has a spanning elementary subgraph with no more than $3(p-6)/2$ lines. Finally, we shall show how a general bipartite graph with a perfect matching can be assembled from elementary bipartite graphs. These results will also serve as motivation for our approach to the study of elementary graphs and decomposition techniques in general.

4.1. Elementary Bipartite Graphs and their Ear Structure

Let us recall from Chapter 3 that a line of graph G (bipartite or not) is **allowed** if it lies in some perfect matching of G and **forbidden** otherwise. A graph G is said to be **elementary** if its allowed lines form a connected subgraph of G. The term "elementary" for this concept was coined by Hetyei in the 1960's, although the idea itself is much older. Indeed we can find this idea employed by König in his 1915 paper! In the case of bipartite graphs there are several useful alternate definitions of this notion. In the following theorem, due mostly to Hetyei (1964), we present four such. (See also Lovász and Plummer (1977).)

4.1.1. THEOREM. *Suppose G is a bipartite graph with bipartition (U, W). Then the following are equivalent:*
(i) *G is elementary;*
(ii) *G has exactly two minimum point covers, namely U and W;*
(iii) *$|U| = |W|$ and for every non-empty proper subset X of U, $|\Gamma(X)| \geq |X| + 1$;*
(iv) *$G = K_2$, or $|V(G)| \geq 4$ and for any $u \in U$, $w \in W$, $G - u - w$ has a perfect matching;*
(v) *G is connected and every line of G is allowed.*

PROOF. (i)$\Rightarrow$(ii). Suppose G has a minimum point cover K such that $K_U = K \cap U \neq \emptyset \neq K \cap W = K_W$. Suppose $G[K]$ contains an allowed line uw. Let M be a perfect matching for G containing uw. Then M matches $U - K_U$ into K_W and $W - K_W$ into K_U. Hence $|U - K_U| < |K_W|$. But U is a cover and $|U| = |K_U| + |U - K_U| < |K_U| + |K_W| = |K|$, contradicting the minimality of K. Thus all lines in $G[K]$ are forbidden. But removal of all lines in $G[K]$ necessarily disconnects G and thus G is not elementary.

(ii)$\Rightarrow$(iii). That $|U| = |W|$ is immediate. Suppose there is a set X, $X \subseteq U$, $\emptyset \neq X \neq U$, such that $|\Gamma(X)| \leq |X|$. Now $(U - X) \cup \Gamma(X)$ covers G and $|(U - X) \cup \Gamma(X)| = |U - X| + |\Gamma(X)| \leq |U - X| + |X| = |U|$ and hence $(U - X) \cup \Gamma(X)$ is a minimum point cover for G. But $U - X \neq \emptyset$ and hence by assumption $\Gamma(X) = \emptyset$, that is, X is a set of isolates. But then $U - X$ covers G and $|U - X| < |U|$, a contradiction.

(iii)$\Rightarrow$(iv). Suppose $G \neq K_2$. Then since $|U| = |W|$, $|V(G)| \geq 4$. Let $u \in U$, $w \in W$ and let $H = G - u - w$. We now prove that H has a perfect matching using the Marriage Theorem. Choose $X \subseteq U - u$, $X \neq \emptyset$. Suppose $|\Gamma_H(X)| < |X|$. Then $|\Gamma_G(X)| \leq |\Gamma_H(X)| + 1 \leq |X|$. Moreover, $X \neq U$ and we have a contradiction of (iii).

(iv)$\Rightarrow$(v). If $G = K_2$ we are done, so suppose $|V(G)| \geq 4$. Suppose G is not connected. Let G_1 be a component of G such that $|V(G_1) \cap U| \leq |V(G_1) \cap W|$. Let $u \in V(G_1) \cap U$ and $w \in W - V(G_1)$. Then $G-u-w$ has no perfect matching, a contradiction. Thus G is connected. That each line of G is allowed is immediate.

(v)$\Rightarrow$(i). Trivial. ∎

4.1.2. EXERCISE. Show that it can be checked in polynomial time whether or not a given bigraph is elementary.

4.1.3. EXERCISE. If G is elementary bipartite and $G \neq K_2$, then G is 2-connected.

4.1.4. EXERCISE. If e is any allowed line of an elementary (not necessarily bipartite) graph G, then any graph G' obtained from G by inserting an even number of new points in line e is again elementary.

4.1.5. EXERCISE. Let G be a bipartite graph and M a perfect matching of G. Orient all lines of G toward the same color class and contract the lines of M. Then the resulting digraph is strongly connected if and only if G is elementary.

We next present a further equivalent property. We have seen by Exercise 4.1.2 that the property of being elementary is in **P** and therefore in NP. In fact we can use property (v) to exhibit that a graph is elementary by verifying connectivity and by providing for each line a perfect matching containing it. The characterization to follow yields a more geometric procedure for showing that a graph is elementary. It can be used to easily find many (in fact, all!) elementary bipartite graphs. We shall refer to this property as the "ear structure" of an elementary bipartite graph. The idea of an ear structure for elementary graphs occurs first in Hetyei (1964).

Let x be a line. Join its endpoints by a path P_1 of odd length (the so-called "first ear"). We may now proceed inductively to build a sequence of bipartite graphs as follows: If $G_{r-1} = x + P_1 + \cdots + P_{r-1}$ has already been constructed, add an r^{th} ear P_r by joining any two points in different color classes of G_{r-1} by an odd path (i.e., P_r) having no other point in common with G_{r-1}. The decomposition $G_r = x + P_1 + \cdots + P_r$ will be called an **ear decomposition** of G_r. (We shall see in the next chapter that in the study of non-bipartite elementary graphs more complicated "ear decompositions" will be used). More specifically, the representation presented in this paragraph will henceforth be called a **Bipartite Ear Decomposition**.

The reader already familiar with graph theory may have recognized that a similar decomposition occurs in the following result due to Whitney (1932): A graph is 2-connected if and only if it can be represented as $G = P_1 + P_2 + \cdots + P_r$, where P_1 is a cycle and each P_i is a path joining two different points of $P_1 + \cdots + P_{i-1}$ having no other point in common with $P_1 + \cdots + P_{i-1}$.

A similar characterization of strongly connected digraphs in terms of ears could also be formulated. In view of Exercise 4.1.5 this is equivalent to the following theorem for which we prefer to give a direct proof.

4.1.6. THEOREM. *A graph is elementary bipartite if and only if G has a Bipartite Ear Decomposition.*

PROOF. Suppose first that $G = x + P_1 + \cdots + P_r$ is a bipartite ear decomposition. We shall proceed by induction on the number of ears. If $r = 1$, G is an even cycle and hence elementary. Now suppose $G_{j-1} = x + P_1 + \cdots + P_{j-1}$ $(r \geq j \geq 2)$ is elementary. We will show $G_j = x + P_1 + \cdots + P_j$ is also elementary. Denote the points of attachment for P_j by a and b.

Choose any line x in G_{j-1}. Since G_{j-1} is elementary by the induction hypothesis, there is a perfect matching F'_x of G_{j-1} containing x by Theorem 4.1.1. Clearly F'_x extends to a perfect matching F_x of G_j such that $x \in F_x$.

Since G_{j-1} is elementary, by Theorem 4.1.1 we know that $G_{j-1} - a - b$ has a perfect matching F' which clearly extends to a perfect matching F_j of G_j which uses all lines of P_j which are not in F_x defined above. Thus all lines of G_j lie in perfect matchings and so G_j is elementary.

Conversely, suppose G is an elementary bipartite graph. We seek an ear decomposition of G. Choose any line x of G and a perfect matching F_x of G containing x. Then if $x = ab$, suppose (without loss of generality) that $y = bc$ is a line adjacent to x. Further, let F_y be a perfect matching of G containing y. (Such must exist by Theorem 4.1.1). Then the component of $F_x \cup F_y$ containing x and y is an even cycle $C = x + P_1$ and we have our first ear.

If $G = x + P_1$ we are done, so suppose there is a line $z = ef$ in $G - (x + P_1)$ with at least one endpoint, say e, on $x + P_1$. (Such must exist, of course, since G is connected.) Let F_z be a perfect matching of G containing z. Define P to be the alternating path in $F_x \cup F_z$ obtained by starting at e, traversing $F_x \cup F_z$ through z to f and ending upon the first return to $x + P_1$. The (necessarily odd) path so traversed is our second ear P_2. Clearly we may continue to find new ears until every line of G lies in some ear. ■

Sometimes it is convenient to describe a bipartite ear decomposition by the sequence of subgraphs $(G_0, G_1, \ldots, G_r)$, where $G_0 = x$ and for $0 \le i \le r$, $G_i = x + P_1 + \cdots + P_i$. By the preceding theorem every G_i is an elementary bipartite graph and it is easy to see that G_i is a "nice" subgraph. (If G is a graph and H is a subgraph of G then H is said to be nice if $G - V(H)$ has a perfect matching.)

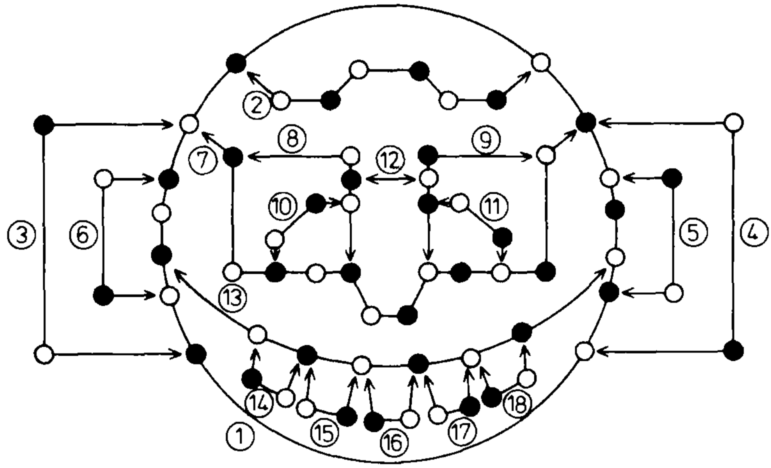


FIGURE 4.1.1. An ear decomposition

Let us make several remarks here concerning such ear decompositions. Let $(G_0, \ldots, G_{r-1}, G_r = G)$ be the sequence of graphs in any ear decomposition.

(1) Such an ear decomposition may be started with any line x of G.
(2) Such an ear decomposition is not in general unique.
(3) Single lines may serve as ears.
(4) Note that for all i, $1 \le i \le r$, the graph $G - V(G_i)$ has a perfect matching. Given any nice subgraph H of G, we can find a decomposition $G = H + P_1 + \cdots + P_r$ just as in Theorem 4.1.6, where for all

i, $1 \le i \le r$, P_i is an ear (i.e., a path of odd length) joining different color classes of $H + P_1 + \cdots + P_{i-1}$.

(5) If G is an elementary bipartite graph and G' is a nice elementary subgraph of G, then there is a bipartite ear decomposition of G in which G' occurs as one of the G_i's. For we can first find a bipartite ear decomposition of G' and then, by the preceding remark, complete this to a decomposition of G.

(6) If G has a ear decomposition $x+P_1+\cdots+P_r$, then it is easy to see that $r = q - p + 1$ where $p = |V(G)|$ and $q = |E(G)|$ and hence, although an ear decomposition may not be unique, the number of ears in any such decomposition is always the same. We shall sometimes find it convenient to denote an ear decomposition simply by $P_1 + \cdots + P_r$ where P_1 is understood to denote the beginning even cycle.

(7) A bipartite ear decomposition of a bipartite graph automatically proves that it is elementary. So the property "elementary" is in NP. In fact, this property is in P, since condition (ii) in Theorem 4.1.1 can be checked in polynomial time. But quite often an ear decomposition provides a more convenient way to exhibit the fact that G is elementary. If we draw a picture of G so that the ear decomposition is recognizable, then the reader does not have to verify the existence of a perfect matching in all the graphs obtained from G by deleting one point from each color class. (See Figure 4.1.1 for an example having 18 ears!)

It is natural to ask at this point if there is any analogue to this ear decomposition which holds for non-bipartite elementary graphs. Unfortunately it is not always possible to obtain a decomposition of such a graph into a chain of elementary subgraphs each obtained from the preceding one by adding a *single* ear. On the other hand, however, we can show, among other things, that if G is any elementary graph, there is a chain $G_0 \subset G_1 \subset \cdots \subset G_k = G$ where G_0 is a line and for $0 \le i \le k$, each G_i is elementary and for $0 \le i \le k-1$ each G_{i+1} is obtained from G_i by attaching *one or possibly two* ears. This will be discussed in more detail in the next chapter.

There are several other nice consequences of the ear structure for elementary bipartite graphs, but these are more easily treated after discussing minimal elementary bipartite graphs in the next section.

4.2. Minimal Elementary Bipartite Graphs

It is often useful in studying graphs with a given property to focus our attention on "line-minimal" graphs having this property. We have already used this approach successfully in our treatment of positive surplus bipartite graphs (see Theorem 1.3.8) and also implicitly in the first proof of König's Minimax Theorem.

An elementary graph G is said to be **minimal elementary** if $G - e$ is not elementary for all lines $e \in E(G)$. We proceed to investigate the properties and structure of such graphs which are, in addition, bipartite. (See Lovász and Plummer (1977).)

In a somewhat different direction, these graphs, reformulated in terms of non-negative matrices, prove useful in the study of doubly stochastic matrices and the permanent function. (Cf. Brualdi and Gibson (1977), Brualdi, Harary and Miller (1980), and Chapter 8 of the present book. For more details see Section 3 at the end of this chapter.)

4.2.1. THEOREM. *Any nice elementary subgraph of a minimal elementary bipartite graph is minimal elementary.*

PROOF. Every nice elementary subgraph of an elementary bipartite graph occurs in some ear decomposition, so this theorem is equivalent to saying that if $G = P_1 + \cdots + P_r$ is a minimal elementary bipartite graph where P_1 is an even cycle, then for each i, $i = 1, \ldots, r$, $G_i = P_1 + \cdots + P_i$ is also minimal elementary. Clearly it is enough to show this for $G_{r-1} = G'$. Since G' is elementary by Theorem 4.1.6, it remains only to show minimality. But if y is any line of G' and if $G' - y$ is elementary, then so is $G' - y + P_r = G - y$ again by Theorem 4.1.6, a contradiction. ■

We point out in particular that in any ear decomposition of a minimal elementary bipartite graph no ear can consist of a single line.

If we want to exhibit that a bigraph is elementary, then clearly it suffices to exhibit this for a spanning minimal elementary bipartite subgraph. The smaller this subgraph, the better! The next few results show that minimal elementary bipartite graphs must be "small" in various senses. One measure of smallness is that the graph should not contain certain subgraphs. We present a result in this direction.

It follows from Theorem 4.2.1 that if G is a minimal elementary bipartite graph with an ear decomposition $P_1 + \cdots + P_r$ $(r \geq 1)$, the starting cycle P_1 can never be a 4-cycle (unless G is itself a 4-cycle). But we can show more, namely that a 4-cycle can never be a proper subgraph of any minimal elementary bipartite graph; in other words, such graphs must have girth ≥ 6.

4.2.2. THEOREM. *If G is a minimal elementary bipartite graph and G is not a 4-cycle, then it contains no 4-cycle.*

PROOF. Suppose, on the contrary, that G contains a 4-cycle as a proper subgraph. Choose any line in G and add ears until the first 4-cycle is found. This 4-cycle had to arise by adding an ear of length one or three. But since G is minimal it must have been an ear of length three. Call this ear P, and denote by H the subgraph of G obtained in the ear construction just prior to adding P. Furthermore, let $x = ab$ be the line of H such that $x \cup P$ forms the 4-cycle. Since, as remarked prior to this theorem, no ear decomposition can start with a 4-cycle, $H \neq x$. By Theorem 4.2.1 it follows that $H + P$ is a minimal elementary bipartite graph. On the other hand, however, $H' = H + P - x$ arises from H by subdividing a line by two points and is thus elementary, a contradiction to the minimality of $H + P$. ■

The following upper bound on the number of lines in a minimal elementary bipartite graph is easy to obtain. The proof is left to the reader as is the exhibition of an extremal family of graphs. (See Lovász-Plummer (1977)).

4.2.3. THEOREM. *Let G be a minimal elementary bipartite graph with p points and q lines. Then $q \leq (3p-6)/2$ and the bound is sharp for all (even) $p \geq 6$.* ■

We certainly know too that every minimal elementary bipartite graph contains points of degree two; for example, two such points must occur in the final ear of any ear decomposition. But the next result implies the existence of rather a lot more such points in general. To this end it is convenient to introduce the following ideas. A line is said to be a **2-line** (a **3-line**) if both of its endpoints have degree $= 2$ (≥ 3) in G.

4.2.4. LEMMA. *Let $G = (U, W)$ be a minimal elementary bipartite graph and suppose $G = x + \cdots + P_r$ is an ear decomposition of G. Then every P_i contains at least one 2-line.*

PROOF. In the given ear decomposition fix any ear P_i. Among all ear decompositions of G which use a piece of P_i as an ear, choose one which uses a smallest such piece. Denote this piece by P_i' and let $x' + P_1' + \cdots + P_r'$ be the ear decomposition in which P_i' occurs. Let α and β denote the endpoints of P_i', $\alpha \in U$, $\beta \in W$.

We claim all inner points of P_i' have degree $= 2$ in G. (Note that by selection and by the lack of single-line ears, P_i' will have at least two inner points.)

So suppose this claim is false. Let $x = uv$ be any line not in P'_i, but having one endpoint — say v — an interior point of P'_i. Let F_x be a perfect matching of G containing x and let F be a perfect matching of $G - V(P'_1 + \cdots + P'_i)$. Starting with x, take an alternating $F_x - F$ path

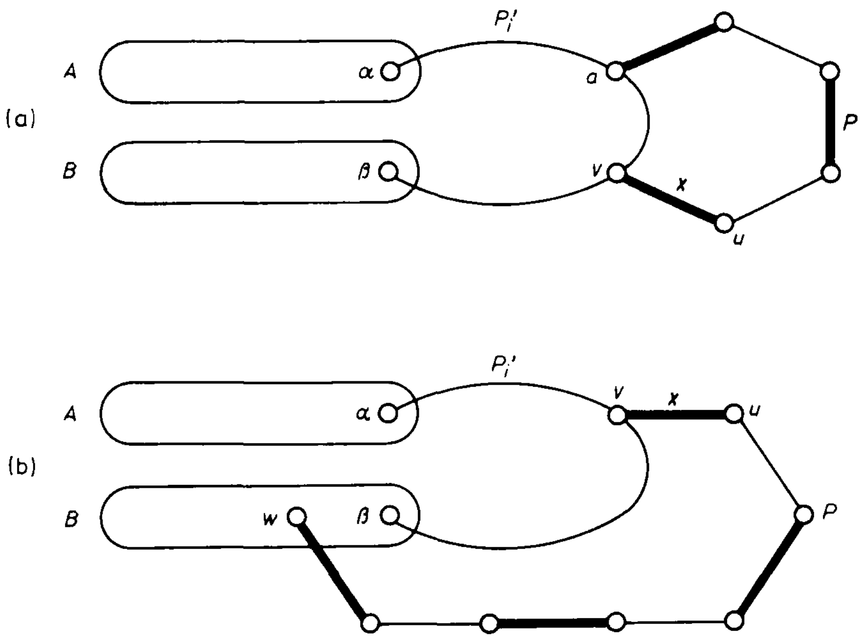


FIGURE 4.2.1.

P leaving v. Let $A = U \cap V(G'_{i-1})$ and $B = W \cap V(G'_{i-1})$. There are two cases to treat.

(1) Suppose P returns to P'_i at some point a, say, before reaching $A - V(P'_i)$ or $B - V(P'_i)$ (Cf. Figure 4.2.1(a)). Then take the ear decomposition $x' + P'_1 + \cdots + P'_{i-1} + [P'_i - P'_i[a, v] + P] + P'_i[a, v]$ which is a nice elementary subgraph and therefore can be extended to an ear decomposition of G. Since $P'_i[a, v]$ is a proper section of P'_i, we contradict the selection of P'_i.

(2) Suppose P reaches $A \cup B$ at a point $w \in B$ say, before P'_i (Cf. Figure 4.2.1(b)). Then take $x' + P'_1 + \cdots + P'_{i-1} + [P'_i - P'_i[v, \beta] + P] + P'_i[v, \beta]$ (which is again a nice subgraph) and proceed as in case (1). ■

4.2.5. COROLLARY. *If G is a minimal elementary bipartite graph which is not an even cycle and if C is any cycle in G, then C contains two 2-lines which are separated on C by points of degree ≥ 3 in G.*

PROOF. Let C be any cycle in G and let x be any non-2-line in G. Choose any ear decomposition $G = x + P_1 + \cdots + P_r$ for G and suppose C is first complete at stage $x + P_1 + \cdots + P_k$. Then by the preceding theorem, P_k, and therefore C, will contain a 2-line y.

On the other hand, we are also free to find an ear decomposition for G beginning with line y; that is, $G = y + P_1' + \cdots + P_r'$. Suppose in this ear decomposition C is first complete at $y + P_1' + \cdots + P_j'$. Then P_j' forms a piece of C and by Theorem 4.2.3 contains a 2-line z ($\neq y$). If $j > 1$, then the two ends of ear P_j are points of degree ≥ 3 on C separating y and z, while if $j = 1$, then since G is not a cycle and is non-separable, there must be two points of degree ≥ 3 separating y and z as desired. ■

4.2.6. COROLLARY. *Any minimal elementary bipartite graph G is separated by its 2-lines.*

PROOF. Suppose not. Then form G' from G by deleting all 2-lines of G. G' must then be a spanning tree for G. But then return one 2-line to G' and the resulting graph will contain a cycle C containing a single 2-line and moreover, C retains this property in G, contradicting Corollary 4.2.4. ■

4.2.7. THEOREM. *Let G be a minimal elementary bipartite graph with p points, q lines and q_2 2-lines. Then we have the following lower bounds for q_2:*
(a) $q_2 \geq q - p + 2$;
(b) $q_2 \geq \lfloor (p + 15)/6 \rfloor$.
Moreover, bound (b) *is sharp for all $p > 6$.*

PROOF. Note first that if $p = 4$ or 6 then the cycle C_p is the only minimal elementary bipartite graph on p points and each has $q_2 = p$.

Hence suppose $p > 6$. Bound (a) is immediate from Theorem 4.2.3 by letting the starting line x of the ear decomposition be a 2-line.

To obtain bound (b) let F be the forest resulting from G after removing all 2-lines. Let k denote the number of points of degree 2 in F, s, the number of isolates in F, m, the number of endpoints of F, and let c be the number of components of F.

Then $3(p - m - k - s) + m + 2k \leq 2(p - c)$ and thus

$$3s + 2m + k \geq p + 2c. \tag{4.2.1}$$

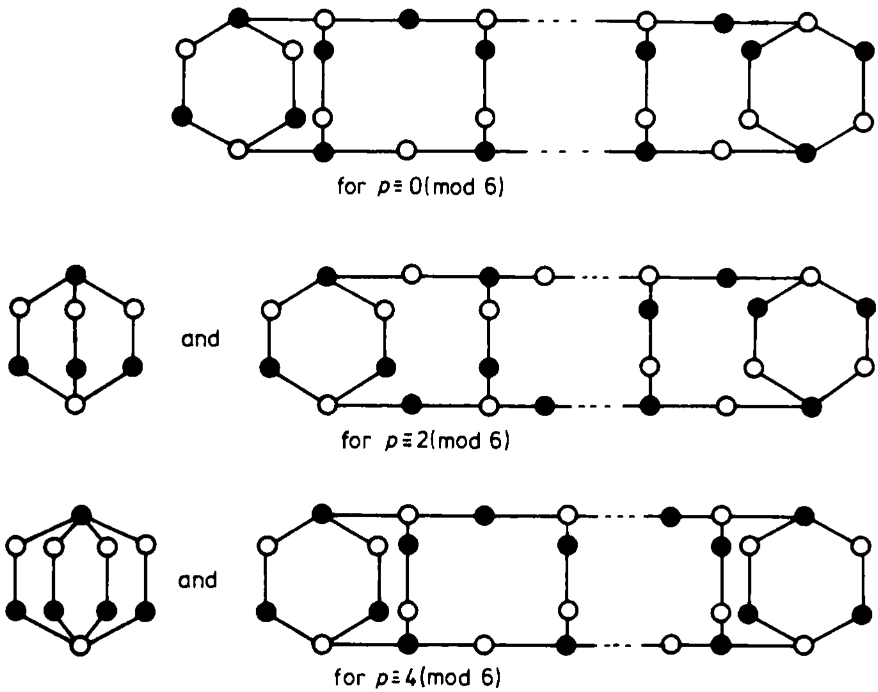


FIGURE 4.2.2. Extremal graphs for Theorem 4.2.7

On the other hand, the number of lines in F is

$$p - c \geq 2k + m \tag{4.2.2}$$

since no point counted by k is adjacent to one counted by m.

From inequality (4.2.1), $k \geq p + 2c - 2m - 3s$, so by inequality (4.2.2) we have $p - c \geq 2p + 4c - 4m + m - 6s$ and hence $3m \geq p + 5c - 6s$. But then $2s + m \geq (p + 5c)/3 \geq (p + 10)/3$ and therefore $q_2 = (2s + m)/2 \geq (p + 10)/6$ and, since q_2 is integral, $q_2 \geq \lfloor (p + 15)/6 \rfloor$. The extremal graphs realizing this bound are presented in Figure 4.2.2. ∎

4.2.8. EXERCISE. If G is a minimal elementary bipartite graph on $p \geq 8$ points and if $q_3 =$ the number of 3-lines in G, then $q_3 \leq 2\lfloor (p - 8)/4 \rfloor$ where the bound is sharp for all (even) $p \geq 8$.

The next lemma not only enables us to obtain yet another characterization of minimal elementary bipartite graphs, this time in terms of cycles and chords, but has several other interesting corollaries as well.

Here as usual, a **chord** of a cycle is any line not contained in the cycle, but having both ends on the cycle. To avoid confusion in the following, we shall call a path of length $n \geq 1$ having all interior points of degree 2 in the graph G an **n-path** in G. Moreover, if P is an n-path in G for any $n \geq 1$, call P a **pending** path.

4.2.9. LEMMA. *If G is an elementary bipartite graph and G_0 is any subgraph of G without isolated points, then there exists a nice elementary subgraph G' such that $G_0 \subseteq G' \subseteq G$ and G' has cyclomatic number no greater than $|E(G_0)| - 1$.*

PROOF. Let G' be a nice elementary subgraph of G containing G_0 and of minimum size with respect to these properties. (Such a G' exists since G is elementary.)

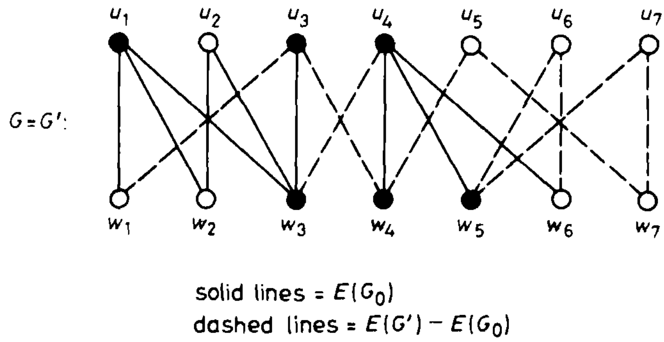
Now define a new graph G_1 as follows. Let W be the set of points of G' having degree at least 3. Then we may assume $W \neq \emptyset$, for otherwise G' is a single line or a cycle and we are done. Given any two points of W, join them by a 2-path for each even pending path joining them in G', by a single line for each odd pending path joining them in G' which has no line in common with G_0 and finally, by a 3-path for each pending path joining them and containing a line of G_0. (See, for example, Figure 4.2.3 where $G = G'$ and $W = \{u_1, u_3, u_4, w_3, w_4, w_5\}$. Note that $|E(G_0)| = 9$, while $|E(G_1)| - |V(G_1)| + 2 = |E(G')| - |V(G')| + 2 = 6$.)

We claim that G_1 is a minimal elementary bigraph. It is clearly bipartite and elementary. Suppose it is not minimal elementary and hence that $G_1 - e$ is elementary, for some line e in $E(G_1)$. Then e is not on any 2-path or 3-path in G_1 and hence the corresponding (odd) pending path in G' must have no line belonging to G_0. But then deleting this pending path from G' results in a proper subgraph G'' of G' which is nice and $G_0 \subseteq G''$. Moreover, G'' is elementary since it arises from $G_1 - e$ by replacing each pending path with one of the same parity. But this contradicts the minimality of G' and the claim is proved.

Now by Theorem 4.2.7(a), G_1 has at least $|E(G_1)| - |V(G_1)| + 2$ different 2-lines. On the other hand, G_1 has at most $|E(G_0)|$ 2-lines, by definition of G_1. Hence $|E(G_0)| \geq |E(G_1)| - |V(G_1)| + 2 = |E(G')| - |V(G')| + 2$ and the proof is complete. ∎

The first corollary is a special case of a theorem of Little (1974b) and the proof is immediate. (For the general case of Little's theorem, see Theorem 5.4.4.)

4.2.10. COROLLARY. *Any two lines of an elementary bipartite graph are contained in a nice cycle.* ∎



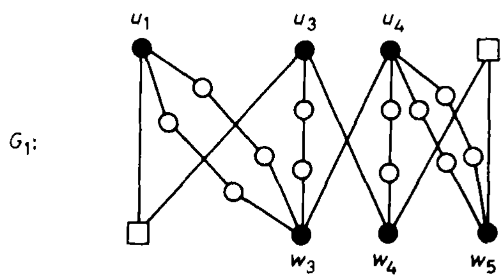


FIGURE 4.2.3.

We have seen in Theorem 4.2.2 that a minimal elementary bipartite graph cannot contain a 4-cycle. A complete list of such excluded subgraphs is not known, but the following result shows at least that the following question is algorithmically decidable: Given a graph G_0, does there exist a minimal elementary bipartite graph G containing G_0 as a subgraph? (The algorithm, however, is exponential.)

4.2.11. THEOREM. *If G_0 is a subgraph of a minimal elementary bipartite graph, then it is a subgraph of one with cyclomatic number no greater than $|E(G_0)| - 1$ and with at most $12|V(G_0)| - 10$ points.*

PROOF. Let G be a minimum minimal elementary bipartite graph containing G_0. By Lemma 4.2.9 there is a nice elementary subgraph $G' \subseteq G$ such that $G_0 \subseteq G'$ and G' has cyclomatic number at most $|E(G_0)| - 1$. From Theorem 4.2.1 we know that G' is also minimal elementary and hence $G = G'$ by the choice of G. So the cyclomatic number of G is at most $|E(G_0)| - 1$.

Let q_2 be the number of 2-lines in G. It is clear that G has no 2-lines point disjoint from G_0 since G is minimum. Furthermore any point of G_0 is incident with at most two 2-lines of G. So we have $q_2 \leq 2|V(G_0)|$.

But on the other hand, Theorem 4.2.7(b) implies that $q_2 \geq (p+10)/6$, where $p = |V(G)|$. It follows immediately that $p \leq 12|V(G_0)| - 10$. ∎

It is interesting to consider Lemma 4.2.9 from another point of view. Let G_0 be any bipartite graph and consider all elementary bipartite graphs containing G_0 as a subgraph. From any member of the class we can construct further members by (1) joining two points in different color classes by a line or (2) subdividing a line by two points. Lemma 4.2.9 implies that we can obtain *all* members of the class from a *finite* number of initial graphs — a so-called **finite basis** — by repeated applications of (1) and (2).

Next we state and prove another finite basis theorem of similar flavor, but applicable to *all* bipartite graphs. Yet another finite basis theorem will be given in Chapter 12 and will be proved using the result we present here.

Let X be a finite set and suppose $X_1, \ldots, X_t \subseteq X$. Consider the class $\mathcal{K}$ of bipartite graphs G such that $X \subseteq V(G)$ and $G - X_i$ has a perfect matching for $i = 1, \ldots, t$. From any member of $\mathcal{K}$ we can construct further members by (1) joining two points in different color classes by a line, (2) subdividing a line by two points or (3) adding two new points and a line connecting them. Then Theorem 4.2.12 below implies that we can generate all members of $\mathcal{K}$ from a finite basis by repeated applications of (1), (2) and (3). (See Lovász (1978).)

4.2.12. THEOREM. *Let X be a set, $X_1, \ldots, X_t \subseteq X$ and suppose that $|X_i| \leq r$ for $i = 1, \ldots, t$. Let G be a bipartite graph such that*
(a) *$X \subseteq V(G)$,*
(b) *$G - X_i$ has a perfect matching, and*
(c) *if any line of G is deleted, property (b) fails to hold in the resulting graph.*

Then the number of points in G with degree ≥ 3 is at most $r^3\binom{t}{3}$.

PROOF. By property (c) we can associate with each line e of G an integer index $i(e)$, $(1 \leq i(e) \leq t)$, such that $G - e - X_{i(e)}$ has no perfect matching. In other words, each perfect matching of $G - X_{i(e)}$ must contain e. Hence lines with the same index are independent.

For each point of G of degree ≥ 3, consider the indices of some three lines adjacent to it. Suppose that the conclusion of the theorem is false; that is, suppose that there are more than $r^3\binom{t}{3}$ points of degree ≥ 3. It then follows that there are more than r^3 points with the same triple of indices. Without loss of generality we may assume that in particular the indices 1, 2 and 3 occur at more than r^3 points.

Letting i and j range over $\{1,2,3\}$, choose a perfect matching F_i of $G - X_i$ and consider the graphs G_{ij} with $V(G_{ij}) = V(G)$ and $E(G_{ij}) = F_i \cup F_j$. The components of G_{ij} are
(1) the points of $X_i \cap X_j$ as isolates,
(2) the common lines of F_i and F_j as 2-point components,
(3) cycles alternating with respect to both F_i and F_j, and
(4) alternating paths the endpoints of which lie in $(X_i - X_j) \cup (X_j - X_i)$.

Note that the number of paths of type (4) is clearly at most r.

Let v be any point to which lines with all three indices 1, 2 and 3 are incident. Trivially, then, v cannot be in any component of type (1) or (2) of G_{ij}, since it is incident with lines of F_i and F_j and these are distinct. It cannot lie on an alternating cycle of G_{ij} either, for then "switching" on this cycle, we could obtain from F_i a perfect matching F'_i of $G - X_i$ which does not contain the line of index i at v, a contradiction. So v must lie in an alternating path component of G_{ij}.

Since the number of points v having the same incident triple of indices 1, 2 and 3 is greater than r^3, we can find two points u and v among them which belong to the same alternating path component of G_{ij} for all three choices of the pair $\{i,j\}$. So for each pair $\{i,j\} \subseteq \{1,2,3\}$, there exists a path P_{ij} connecting u and v and alternating with respect to F_i and F_j. If we choose the labelling appropriately we may assume that P_{12} starts with a line of F_1 and P_{23} starts with an F_2 line at point u, say. (Note that we do not make any claim of this sort about path P_{13}.)

Starting at point u, traverse P_{12} until a point of P_{23} is encountered. (At the latest this will happen when point v is reached.) Then return to point u using P_{23}. The cycle traversed has even length (since G is bipartite) and so it must alternate with respect to F_2. But then, as above, "switching" on this cycle we obtain from F_2 another perfect matching F'_2 of $G - X_2$, which does not contain the line of index 2 incident with point u, a contradiction. ∎

To illustrate these ideas the reader is invited to consider the following example.

4.2.13. EXAMPLE. Let $X = X_1 = \{1,2,3,4\}, X_2 \neq \emptyset, X_3 = \{1,2\}$ and let $X_4 = \{3,4\}$. Then by a straightforward — though tedious —

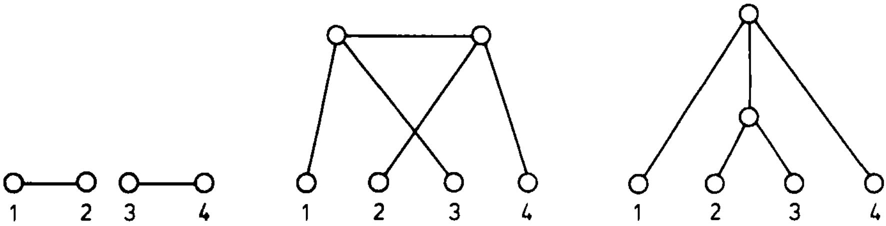


FIGURE 4.2.4. The class $\mathcal{K}$ of basis graphs

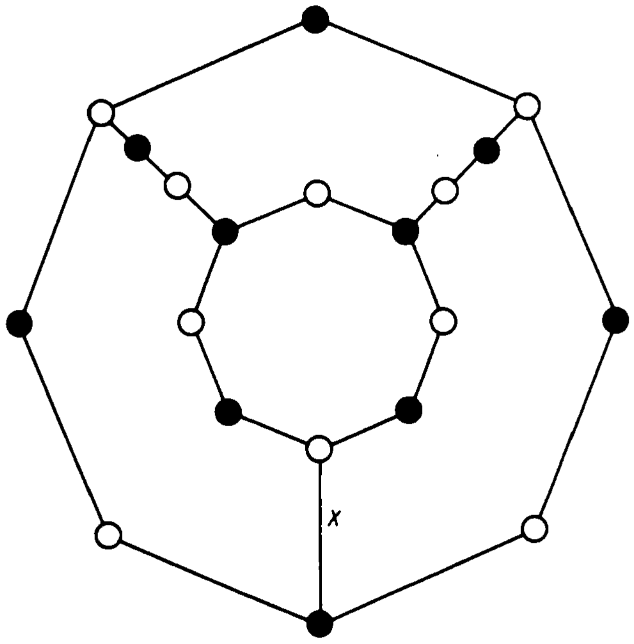


FIGURE 4.2.5. A minimal elementary bigraph with a chord x

argument, it may be shown that the basis graphs for the corresponding class $\mathcal{K}$ are precisely those shown in Figure 4.2.4.

The following remains unsettled at the time of this writing.

4.2.14. CONJECTURE. *The conclusion of Theorem* 4.2.12 *holds for non-bipartite graphs as well.*

4.2.15. EXERCISE. Use Theorem 4.2.12 to prove the finite basis consequence of Lemma 4.2.9 in the special case when G_0 is connected. (Hint: Let the X_i's be the pairs of endpoints of lines of G_0.)

Our final result of this section is the following concise characterization of minimal elementary graphs.

4.2.16. THEOREM. *Let G be an elementary bipartite graph. Then G is minimal if and only if no nice cycle has a chord.*

PROOF. If G is minimal, consider any ear decomposition starting with any given nice cycle. Then any chord of this cycle would have to be an ear which, as we have already shown, is impossible.

Conversely, suppose G is not minimal. Then it contains a line $x = ab$ such that $G - x$ is elementary. But then by Corollary 4.2.10, a and b lie on a nice cycle in $G - x$ and thus x is a chord of this cycle. ∎

We hasten to point out that Theorem 4.2.16 does not preclude the existence of chords in minimal elementary bipartite graphs. The minimal elementary bipartite graph of Figure 4.2.5 has a chord labelled x.

4.3. Decomposition into Elementary Bipartite Graphs

In Section 3.2 we began to look at the work of Dulmage and Mendelsohn (1958, 1959) on developing a canonical decomposition of arbitrary bipartite graphs and showed how their results could in fact be derived from the Gallai-Edmonds Theorem. In particular, it was demonstrated there that every bipartite graph has a unique decomposition into three point-disjoint subgraphs together, perhaps, with some additional lines joining these three graphs in certain ways. Two of these three graphs are positive surplus bipartite graphs. Such graphs were further analyzed in Section 1.3.

The remaining graph was shown to always contain a perfect matching. We now show how this type of graph can be further decomposed into elementary bipartite graphs. (In the notation of Section 3.2 we hence propose to further decompose $G[C_1 \cup C_2]$.) These results are due to König (1916a, 1916b) and Dulmage and Mendelsohn (1958, 1959). So let $G = (U, W)$ be a bipartite graph with a perfect matching. The subgraph of G consisting of all lines which are allowed in G has components, say, $L_1, L_2, \ldots, L_k$ ($k \geq 1$). Clearly, each L_i is an elementary graph. Define point sets S_i and T_i by $S_i = U \cap V(L_i)$ and $T_i = W \cap V(L_i)$ for $i = 1, \ldots, k$.

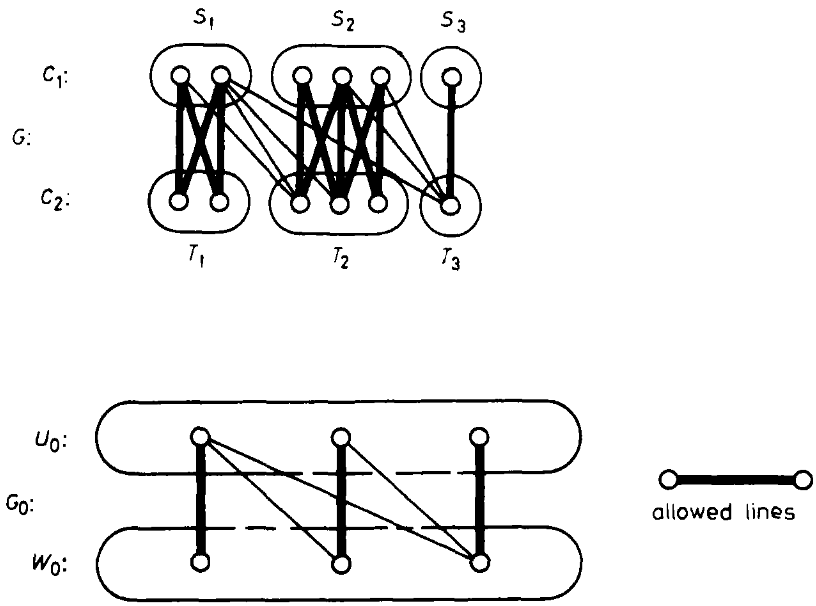


FIGURE 4.3.1.

Let G_0 be the bipartite graph obtained from G by shrinking every S_i and T_i to a point and replacing each set of parallel lines which might result by a single line. (Cf. Figure 4.3.1.)

4.3.1. LEMMA. *The bipartite graph $G_0 = (U_0, W_0)$ defined above has a unique perfect matching M consisting of the k lines $l_1, \ldots, l_k$ where l_i joins points S_i and T_i of G_0.*

PROOF. Suppose M is not unique; that is, suppose that M' is also a perfect matching for G_0 and $M' \neq M$. Say, $M' = \{l'_1, \ldots, l'_k\}$. Then back in G the lines of M' correspond to k lines at least two of which join S_i's and T_i's with different subscripts. Of course in G for each $i = 1, \ldots, k$ there is exactly one line of M' covering a point of S_i and exactly one covering a point of T_i. In deleting all points of $M' \cap G$ we have removed exactly two from each L_i, namely one from S_i and one from T_i for each $i = 1, \ldots, k$. But each L_i is elementary, so by Theorem 4.1.1, $G - V(M')$ has a perfect matching which together with the deleted lines of M' forms a perfect matching for G containing at least two lines joining different L_i's. But these two lines are thus allowed in G contradicting the definition of the L_i's they join. ∎

Next we need the following result.

4.3.2. LEMMA. *Let $G_0 = (U_0, W_0)$ be any bipartite graph with a unique perfect matching. Then the points of G_0 can be labelled $U_0 = \{u_1, \ldots, u_m\}$, $W_0 = \{w_1, \ldots, w_m\}$ such that for every line $u_i w_j$, $i \geq j$.*

PROOF. The proof is by induction on m. If $m = 1$, the result is clear. Suppose $m > 1$. Let M_0 be the unique perfect matching in G_0 and let P be an alternating path starting and ending with a line of M_0 and having

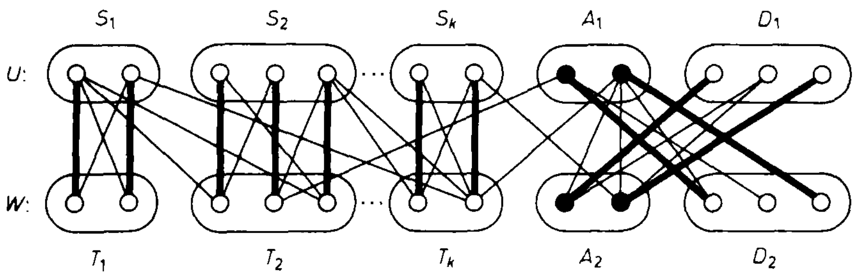


FIGURE 4.3.2. Dulmage-Mendelsohn decomposition of a bigraph

maximum length with respect to these properties. Let u be an endpoint of P and without loss of generality we may assume $u \in U_0$. We claim $\deg_{G_0}(u) = 1$. But this is clear, for u can be adjacent to only one point in W_0 or else we contradict either the uniqueness of the perfect matching in G_0 or the maximality of P.

Let uw be the line of M_0 incident with u. Then $G' = G_0 - u - w$ also satisfies the hypotheses of this lemma. Thus by the induction hypothesis we may index the points $U' = \{u_2, \ldots, u_m\}$, $W' = \{w_2, \ldots, w_m\}$ such that for each line $u_i w_j$ in G', $i \geq j$. But then letting $w = w_1$ and $u = u_1$, we have the desired labelling of G_0. ■

4.3.3. EXERCISE. Let G be a bipartite graph with a unique perfect matching. Orient all lines of G toward the same color class and contract the lines of M. Show that the resulting digraph is acyclic. Use this to produce an alternate proof of Lemma 4.3.2.

Applying Lemma 4.3.2 to our bipartite graph G having a perfect matching we see immediately that the components $L_1, \ldots, L_k$ can be ordered so that for any line $uv \in G$ with $u \in S_i$ and $v \in T_j$ we have $i \leq j$.

Now consider an arbitrary bipartite graph G. By Theorem 3.2.5, the subgraph $G[C_1 \cup C_2]$ has a perfect matching, and so we can apply the previous decomposition results. Let T_i, S_i be defined as above for $G[C_1 \cup C_2]$ instead of G. (Cf. Figure 4.3.2.)

Recalling from part (6) of Theorem 3.2.4 that $C_1 \cup A_1 \cup A_2$ and $C_2 \cup A_1 \cup A_2$ are both minimum point covers of G, we see that there is a sequence of "intermediate" minimum point covers for G:

$$\begin{gathered}
A_1 \cup A_2 \cup T_1 \cup \cdots \cup T_k, \\
A_1 \cup A_2 \cup S_1 \cup T_2 \cup \cdots \cup T_k, \\
\vdots \\
A_1 \cup A_2 \cup S_1 \cup \cdots \cup S_i \cup T_{i+1} \cup \cdots \cup T_k, \\
\vdots \\
A_1 \cup A_2 \cup S_1 \cup \cdots \cup S_k.
\end{gathered}$$

This decomposition theory can be of some assistance in sparse matrix inversion. Let A be any $n \times n$ matrix over any field. Construct a bipartite graph $G(A)$ with bipartition (U, W) where $U = \{u_1, \ldots, u_n\}$, $W = \{w_1, \ldots, w_n\}$ and $u_i w_j \in E(G(A))$ if and only if $a_{ij} \neq 0$. If $G(A)$ has no perfect matching then $\det A = 0$ and A is not invertible. So suppose $G(A)$ has a perfect matching. Let $S_1, \ldots, S_k$ and $T_1, \ldots, T_k$ be constructed as above. Then the corresponding rearrangement of the rows and columns of A result in rectangular blocks of zeros which in turn facilitate the computation of A^{-1} (cf. Figure 4.3.3). This rearrangement can of course be carried out by pre- and post-multiplication of A by appropriate permutation matrices P and Q.

This rearrangement of the rows and columns of A may be useful in particular if we have to invert several matrices with the same pattern of zeros, since then the decomposition need only be done once.

REMARK. Several of the results on minimal elementary bipartite graphs presented in this chapter were in fact stated and proved first in the language of non-negative matrices. Let A be any non-negative square matrix and $G(A) = (U, W)$ the corresponding bigraph. A matrix A is **partly decomposable** if its rows and columns may be permuted so as to yield a matrix of the form $\begin{pmatrix} X & 0 \\ Z & Y \end{pmatrix}$ where X and Y are square. Otherwise A is **fully indecomposable.** It is clear, using Theorem 4.1.1, that A is fully indecomposable if and only if $G(A)$ is elementary. If a fully indecomposable matrix A loses this property whenever any non-zero entry is changed to zero, it is said to be **nearly decomposable**. Clearly this corresponds to $G(A)$ being a minimal elementary bipartite graph. Couched in this matrix terminology we can find proofs of Theorem 4.1.6 due to Hartfiel (1970) and Theorem 4.2.6 by Minc (1972).

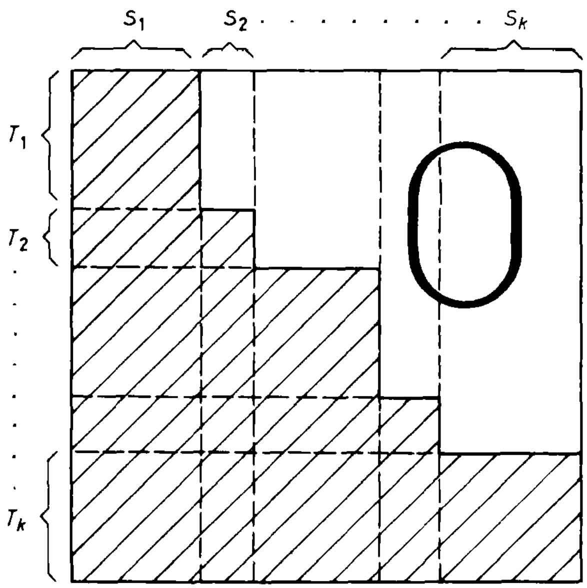


FIGURE 4.3.3.

For the general problem of finding a perfect matching in $G(A)$ — or as sparse matrix researchers prefer to call it, "transversal selection" — we have available for bigraph matching the Hopcroft–Karp Algorithm (1971, 1973) (or any other for that matter). But heuristics often perform better in practice than these algorithms do. For a comprehensive survey of the state of the art from the sparse matrix researcher's viewpoint, see Duff (1977).

This Page Intentionally Left Blank

5

General Graphs with Perfect Matchings

5.0. Introduction

In Chapter 4 we learned about bipartite elementary graphs, their so-called "ear structure" and how to express the structure of any bipartite graph having a perfect matching in terms of elementary bigraphs. In this chapter we want to pursue the same goals in the more general case when the graph is no longer necessarily bipartite.

After becoming familiar with some useful basic properties of general elementary graphs in the first section we proceed to introduce a canonical partition for any elementary graph in Section 2. Historically, this idea seems to have originated with Kotzig (1959a, 1959b, 1960) and was further developed by Lovász (1972a). Two equivalent ways of defining this partition are: (a) the blocks of the partition are exactly the maximal barriers in G, or (b) the associated equivalence relation on $V(G)$ is defined by: x and y are equivalent if and only if $G-x-y$ has no perfect matching.

A new general decomposition procedure for all elementary graphs is then developed. It involves this canonical partition and uses as building blocks the elementary bipartite graphs of Chapter 4 together with a new kind of graph called "bicritical". A graph G is **bicritical** if G contains a line and $G - x - y$ has a perfect matching for every pair of distinct points x and y in G. This decomposition can be pushed one step further by decomposing bicritical graphs into 3-connected bicritical graphs, also called **bricks**. We will refer to this entire procedure as the **Brick Decomposition Procedure**.

We shall call any simple graph with a perfect matching **saturated** if the addition of any new line increases the number of perfect matchings. In Section 3 we will return to the idea of the canonical partition and use it to help us obtain a construction (the so-called "Cathedral Construction") for saturated non-elementary graphs.

The chart shown in Figure 5.0.1 describes the canonical procedure which we obtain by combining the methods of Chapters 3, 4 and 5.

We do not know of any "canonical" decomposition procedure applicable to the four classes forming the four endpoints of the above tree.

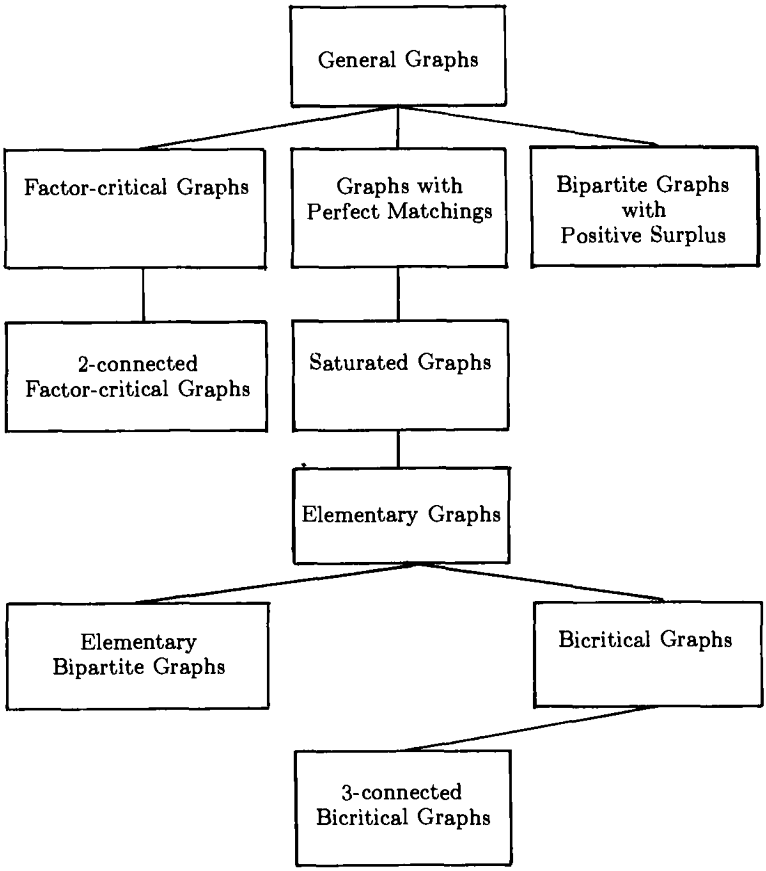


FIGURE 5.0.1. The Brick Decomposition Procedure

So we shall have to find a new approach to describe structural properties of these graphs.

One such approach is the technique of ear decomposition. We introduced and used ear decompositions of elementary bipartite graphs in Chapter 4. Similar decompositions can be defined for general 1-extendable graphs and for factor-critical graphs. These are introduced and studied in Sections 5.4 and 5.5.

A second approach to obtaining structural information about these families of "building blocks" is to study those members of the families which are minimal with respect to line deletion. Our main result on bipartite graphs with positive surplus (Theorem 1.3.8) tells us that the minimal graphs with this property are forests with all points in one of the two color classes having degree 2. In Chapter 4 we proved several properties of minimal elementary bipartite graphs. In Sections 5.4 and 5.5 we derive some similar results for 1-extendable, minimal bicritical and minimal factor-critical graphs. In this study ear decomposition techniques are the main tools.

One of the most important uses of the decomposition methods in this chapter is to help us estimate how many *different* perfect matchings a graph has. Exact determination of this number seems quite a difficult job. In fact, it is NP-hard (see L. Valiant (1979a)). However, our results will be used to determine some useful lower bounds for this number. More on this subject can be found in Chapter 8.

5.1. Elementary Graphs: Elementary Properties

Let us begin by extending to all graphs a concept introduced for bipartite graphs in Chapter 4. A graph G with a perfect matching is **elementary** if its allowed lines form a connected subgraph. If G has a perfect matching, is connected and all of its lines are allowed, G is said to be **1-extendable** (or sometimes **matching covered.**) Thus every 1-extendable graph is elementary. Note that for connected bipartite graphs these two properties are the same by Theorem 4.1.1. But this is not always so for non-bipartite graphs as one readily sees by considering $K_4 - e$ where e is any line in K_4.

The following remarks, although transparently obvious, will prove useful subsequently:

(1) If G is any graph (other than K_2) and G' results from G by subdividing a line with two points, then G' is 1-extendable if and only if G is 1-extendable.

(2) If G is any elementary graph (other than K_2) and G' results from G by subdividing a line e with two points, then G' is elementary if and only if e is allowed in G.

(3) Every elementary graph is 2-connected.

(4) The allowed lines of any elementary graph form a 1-extendable spanning subgraph.

In order to begin to ascertain the structure of general graphs with perfect matchings we will find it convenient to employ two additional ideas, one old and one new. Suppose for the moment that G is an arbitrary graph which may or may not contain a perfect matching. The old idea we need here was introduced in Section 3.3. Recall that a set $X \subseteq V(G)$ is a **barrier** if $c_0(G-X) = |X| + \text{def}(G)$. Thus set $A(G)$ of the Gallai-Edmonds decomposition of G is always a barrier.

First we give a characterization of elementary graphs in terms of the Gallai-Edmonds structure. The following two lemmas lead us toward such a characterization.

5.1.1. LEMMA. *If G is elementary, then $C(G-x) = \emptyset$, for all $x \in V(G)$.*

PROOF. Suppose that $C(G-x) \neq \emptyset$ for some $x \in V(G)$. Let f be any allowed line in G and let M_f be a perfect matching of G which contains f. Let $e = xy$ be the line of M_f covering x. Then $M_f - e$ is a maximum matching of $G-x$ which misses y and hence $y \in D(G-x)$. From the Gallai-Edmonds Theorem we know that no line of $M_f - e$ joins $C(G-x)$ and $A(G-x)$ and hence no line of M_f joins these two sets. In particular, f does not join these sets and since f was an arbitrary allowed line in G, no allowed line in G joins $A(G-x)$ and $C(G-x)$. Moreover, we saw above that all allowed lines of G covering x have their other endpoints in $D(G-x)$.

Thus the allowed lines of G form a disconnected subgraph of G contradicting remark (4) above. ∎

5.1.2. LEMMA. *Let G be any graph. If X is a maximal barrier in G, $x \in X$ and $C(G-x) = \emptyset$, then $A(G-x) = X - x$.*

PROOF. By Exercise 3.3.9, $X - x$ is a barrier in $G - x$ and hence, by the remark in Section 3.3, $X - x \subseteq A(G-x) \cup C(G-x) = A(G-x)$. On the other hand, by Lemma 3.3.10, $A(G-x) \cup \{x\}$ is a barrier in G and so by the maximality of X, $X = A(G-x) \cup \{x\}$ and the lemma follows.∎

At this point the reader may wish to remind himself of the definition of **an extreme** set of points given in Section 3.3.

5.1.3. THEOREM. *A graph G is elementary if and only if* $\text{def}(G) = 0$ *and for all $x \in V(G)$, $C(G-x) = \emptyset$.*

PROOF. The necessity of the condition is immediate from Lemma 5.1.1.

Now suppose $\text{def}(G) = 0$ and $C(G-x) = \emptyset$ for all $x \in V(G)$, but G is not elementary. That is, suppose $V(G) = V_1 \cup V_2$, $V_1 \neq \emptyset \neq V_2$, $V_1 \cap V_2 = \emptyset$ and every line of $\nabla(V_1, V_2)$ is forbidden. Thus every perfect matching of G must match V_1 onto itself and it follows that $|V_1|$ and $|V_2|$ are both even.

Claim 1. G is connected.

Suppose it is not. Since $\text{def}(G) = 0$, G is the disjoint union of graphs $G_1, G_2, \ldots, G_k$, $(k \geq 2)$, such that $\text{def}(G_i) = 0$ for all i. Thus each $|V(G_i)|$ is even.

Let $U_1 = V(G_1)$ and $U_2 = \bigcup_{i=2}^{k} V(G_i)$. Choose any $x \in U_1$ and let M be any maximum matching of $G - x$. Since $\text{def}(G) = \text{def}(G_1) = 0$, $\text{def}(G - x) = \text{def}(G_1 - x) = 1$ and so M misses exactly one point of U_1 and hence matches U_2 with itself. So $D(G - x) \subseteq U_1$ and hence $D(G - x) \cup A(G - x) \subseteq U_1$. But then $C(G - x) \supseteq U_2 \neq \emptyset$ which is a contradiction. This proves Claim 1.

So G is connected and in particular it follows that $\nabla(V_1, V_2) \neq \emptyset$; that is, there is a line $v_1 v_2$, $v_i \in V_i$, which is forbidden in G by the definition of V_1 and V_2.

Thus $\{v_1, v_2\}$ is extreme in G by Exercise 3.3.7 and hence lies in a barrier S of G by Lemma 3.3.8. Without loss of generality, assume S is an (inclusion-wise) maximal barrier.

Claim 2. Every line xy with $x \in S$ and $y \notin S$ is allowed in G.

To see this, let xy be such a line. Now $C(G - x) = \emptyset$ by hypothesis and so by Lemma 5.1.2 we have $S = A(G - x) \cup \{x\}$; that is, $A(G - x) \subseteq S$. But then $y \notin A(G - x)$ and again since $C(G - x) = \emptyset$, $y \in D(G - x)$. So there is a maximum matching M_x of $G - x$ missing y (and of course no other point of $G - x$, since $\text{def}(G - x) = 1$). But then $M_x \cup \{xy\}$ is a perfect matching of G, proving Claim 2.

Claim 3. $G - S$ has no odd component joined to both $S \cap V_1$ and $S \cap V_2$.

Suppose, on the contrary, that H is such a component. Let uv be a line with $u \in V(H)$ and $v \in S \cap V_1$. By Claim 1 there is a perfect matching F of G which contains uv. Thus $u \in V_1$ by definition of the partition (V_1, V_2) (cf. Figure 5.1.1.).

Since G has a perfect matching F, graph $G - v$ has a Gallai-Edmonds decomposition in which $D(G - v) \neq \emptyset$, but $C(G - v) = \emptyset$. (Cf. Figure 5.1.2 and recall that $A(G - v) \cup \{v\} = S$ as argued in the proof of Claim 2.)

If e_v is the line of F covering v, then $F - e_v$ is a maximum matching of $G - v$. Thus $F - e_v$ covers $A(G - v)$ and has at most one line joining each odd component of $D(G - v)$ (of which H is one) to $A(G - v)$. Upon reinserting v and e_v we have exactly one line of F joining S to each component of $G - S$ and in particular, exactly one line of F joining H

to S. So $E(H) \cap F$ is a near-perfect matching of H. Thus $E(H - V_2) \cap F$ is a near-perfect matching in $H - V_2$, and so $|V(H) \cap V_1|$ is odd.

Similarly, $|V(H) \cap V_2|$ is also odd and hence $|V(H)|$ is even, a contradiction. This proves Claim 3.

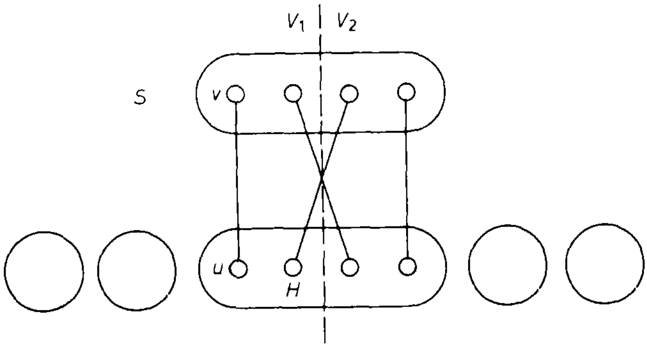


FIGURE 5.1.1.

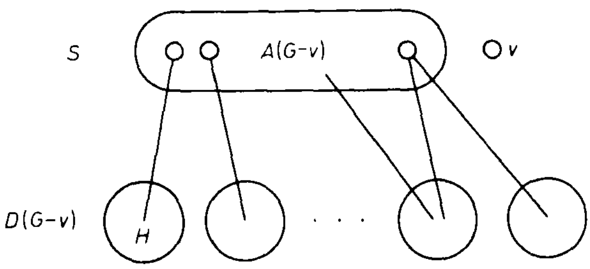


FIGURE 5.1.2.

So by Claims 1 and 3 all odd components of $G - S$ are joined to $S \cap V_1$ or $S \cap V_2$, but not to both. For $i = 1, 2$, let a_i be the number of odd components of $G - S$ joined to $S \cap V_i$. Then $a_1 + a_2 = c_0(G - S) = |S| + \text{def}(G) = |S|$. Without loss of generality we may assume $a_1 \leq |S \cap V_1|$. Let $x \in S \cap V_2$. Then recall $S - x = A(G - x)$ from Lemma 5.1.2 and hence $A(G-x)$ has a subset of $k = |S \cap V_1|$ points joined to at most k odd components of $(G-x) - A(G-x)$ contradicting part (c) of the Gallai-Edmonds Structure Theorem. This completes the proof of the theorem. ∎

The following result shows that for elementary graphs the idea of a calculus of barriers can be pushed somewhat further than in the case of general graphs. (Cf. Theorem 3.3.11.)

5.1.4. LEMMA. *Suppose G is elementary and that X and Y are barriers in G such that $X \cap Y \neq \emptyset$ and no line of G joins $X - Y$ to $Y - X$. Then $X \cap Y$ and $X \cup Y$ are barriers in G.*

PROOF. Suppose $v \in X \cap Y$ and denote $G-v$, $X-v$ and $Y-v$ by G', X' and Y' respectively. Then X' and Y' are barriers in G' by Exercise 3.3.9 and clearly no line joins $X'-Y'$ to $Y'-X'$. Also G is elementary so $C(G') = \emptyset$ by Theorem 5.1.3. But then by Theorem 3.3.13, $X' \cap Y'$ and $X' \cup Y'$ are barriers in G'. Hence $X \cap Y = (X' \cap Y') \cup \{v\}$ and $X \cup Y = (X' \cup Y') \cup \{v\}$ are barriers in G by Lemma 3.3.10. ∎

The careful reader may wonder what has become of the *even* components of $G - S$ when G is elementary and S is a barrier! But relax, dear reader, the next result assures us that there are none!

5.1.5. THEOREM. *Let G be elementary. Suppose X is a non-empty extreme set in G. Then*
(a) *$c(G-X) \leq |X|$, and*
(b) *if $c(G-X) = |X|$, then all components of $G-X$ are odd.*

PROOF. Let $G_1, \ldots, G_k$ be the components of $G-X$. Then since G has a perfect matching, $\operatorname{def}(G-X) = |X| = \sum_{i=1}^{k} \operatorname{def}(G_i)$.

We claim that $\operatorname{def}(G_i) > 0$ for each i. For suppose $\operatorname{def}(G_1) = 0$, say. Let F be any perfect matching of G. Now at least $\operatorname{def}(G_i)$ lines of F leave G_i and therefore at least $\sum_{i=1}^{k} \operatorname{def}(G_i) = |X|$ lines of F leave $G - X$. Since X has only $|X|$ points to receive these lines, we have equality throughout and in particular no lines of F leave G_1. But F was arbitrary, so all lines in $\nabla(G_1)$ are forbidden, a contradiction, and our claim is proved.

Thus $\operatorname{def}(G_i) \geq 1$ for each i and hence $|X| = \sum_{i=1}^{k} \operatorname{def}(G_i) \geq k$ proving assertion (a).

If $k = |X|$, we must have $\operatorname{def}(G_i) = 1$ for all i and hence $|V(G_i)|$ is odd, proving (b). ∎

We will now conclude this section by using the above theorem to obtain a "Tutte-like" characterization for all elementary graphs.

5.1.6. THEOREM. *A graph G is elementary if and only if $c_0(G-X) \leq |X|$ for all $X \subseteq V(G)$ and if equality holds for some $X \neq \emptyset$ (i.e., if X is a non-empty barrier) then $G-X$ has no even components.*

PROOF. ($\Leftarrow$). G must have a perfect matching by Tutte's Theorem, so $\operatorname{def}(G) = 0$. Now choose any point $x \in V(G)$ and note that $G' = G-x$ has $\operatorname{def}(G') = 1$. Let $X = A(G') \cup \{x\}$. Then $c_0(G-X) = c_0(G' - A(G')) = |A(G')| + \operatorname{def}(G') = |A(G')| + 1 = |X|$, where we use the fact that $A(G')$ is a barrier in G'.

So by assumption, $G-X$ has no even components, that is, $C(G-x) = \emptyset$. But then by Theorem 5.1.3, G is elementary.

($\Rightarrow$). Suppose now, conversely, that G is elementary. Then by Tutte's Theorem, $c_0(G-X) \leq |X|$ for all $X \subseteq V(G)$. Now assume further that $c_0(G-X) = |X|$ for some $X \neq \emptyset$, $X \subseteq V(G)$. Thus X is a barrier in G.

Suppose R is an even component of $G-X$ and let F be a perfect matching of G. Then the lines of F matching X to the odd components of $G-X$ must cover X and hence no line of F matches any point of R to X. This contradicts the fact that G is elementary and hence $R = \emptyset$ and the proof of the theorem is complete. ∎

5.2. The Canonical Partition $\mathcal{P}(G)$

We now introduce the **canonical partition** of the point set of any elementary graph. This partition will prove to be a most useful tool in obtaining the main structural results of this chapter. We define $\mathcal{P}(G)$ to be $\{S_1, \ldots, S_k\}$ where $S_1, \ldots, S_k$ are the maximal barriers in G (or equivalently, the maximal extreme sets in G). This partition is the "new" idea promised at the beginning of Section 5.1.

5.2.1. LEMMA. *If G is elementary then $\mathcal{P}(G) = \{S_1, \ldots, S_k\}$ is a partition of $V(G)$.*

PROOF. Since G is elementary, $V(G) = C(G)$ and so if $x \in V(G)$, $A(G-x) \cup \{x\}$ is a barrier in G by Lemma 3.3.10. Thus every point of G lies in a maximal barrier of G.

On the other hand, suppose X and Y are maximal barriers in G and suppose $x \in X \cap Y$. By Lemma 5.1.1, $C(G-x) = \emptyset$ for all $x \in V(G)$. Thus if $x \in X \cap Y$, by Lemma 5.1.2 we have $A(G-x) = X - x = Y - x$ and so $X = Y$. ∎

The next theorem summarizes some important properties of the partition $\mathcal{P}(G)$.

5.2.2. THEOREM. *Let $\mathcal{P}(G) = \{S_1, \ldots, S_k\}$ be the canonical partition of an elementary graph G. Then:*

(a) *$X \subseteq V(G)$ is extreme in G if and only if $X \subseteq S_i$, for some i, $1 \leq i \leq k$.*

(b) *If x and y are points of G, then $G - x - y$ has a perfect matching if and only if x and y lie in different classes of $\mathcal{P}(G)$. In particular, a line $e = xy \in E(G)$ is allowed in G if and only if x and y lie in different classes of $\mathcal{P}(G)$.*

(c) *If $\emptyset \neq X \subseteq S_i \in \mathcal{P}(G)$, then $A(G - X) = S_i - X$ and $C(G - X) = \emptyset$.*

(d) *Let $S \subseteq V(G)$. Then $S \in \mathcal{P}(G)$ if and only if $G - S$ has exactly $|S|$ components and each is factor-critical.*

PROOF. (a) is immediate by Lemma 3.3.8 and by Exercise 3.3.3.

(b). Let x and y be any two points in G. Then $G - x - y$ has no perfect matching if and only if $\operatorname{def}(G - x - y) > 0$ and by parity this inequality holds if and only if $\operatorname{def}(G - x - y) \geq 2$. But this in turn holds if and only if $\{x, y\}$ is extreme, that is, (by part (a)), $\{x, y\} \subseteq S_i$ for some $S_i \in \mathcal{P}(G)$.

(c). Suppose $\emptyset \neq X \subseteq S_i \in \mathcal{P}(G)$ and that $x \in X$. Since G is elementary, it follows by Lemma 5.1.1 that $C(G - x) = \emptyset$, for all $x \in V(G)$. But then by Lemma 5.1.2, $A(G - x) = S_i - x$ and hence $X - x \subseteq A(G - x)$.

Now since $G - X = (G - x) - (X - x)$, we have $A(G - X) = A((G - x) - (X - x)) = A(G - x) - (X - x)$, using Lemma 3.2.2(a) $|X - x|$ times. But this in turn $= (S_i - x) - (X - x) = S_i - X$.

On the other hand, if $y \in A(G - x)$, then $C(G - x - y) = C(G - x)$ by Lemma 3.2.2(a) and this set is empty by Lemma 5.1.1. Again, $G - X = (G - x) - (X - x)$ for all $x \in V(G)$, so again by repeatedly applying Lemma 3.2.2(a) and letting $x = x_1$ we have $C(G - X) = C(G - \{x_1, \ldots, x_k\}) = C((G - x_1 - \cdots - x_{k-1}) - x_k) = C(G - x_1 - \cdots - x_{k-1}) = \cdots = C(G - x_1) = C(G - x) = \emptyset$, using Lemma 5.1.1 again at the last step.

(d). First suppose $S \in \mathcal{P}(G)$. Then by part (c), $A(G - S) = \emptyset = C(G - S)$ and so by the Gallai-Edmonds Theorem, $G - S$ consists of factor-critical components. Since S is a barrier in G, there must then be precisely $|S|$ such components.

Conversely, suppose $S \subseteq V(G)$ is such that $G - S$ has exactly $|S|$ components and each is factor-critical. Then $c_0(G - S) = |S| = |S| + \operatorname{def}(G)$, so S is a barrier and hence $S \subseteq S_i$ for one of the maximal barriers

S_i in G. So by part (c), $A(G-S) = S_i - S$. But $G-S$ is given as a union of disjoint factor-critical graphs so $A(G-S) = \emptyset$ and hence $S = S_i$. ∎

Let us pause at this point to make several additional remarks about the partition $\mathcal{P}(G)$.

1. Note that if G is elementary and $e = xy$ is allowed in G and if we obtain a new graph G' by subdividing the line e by the insertion of two new points u and v, then the classes of $\mathcal{P}(G')$ are the same as those of G except u joins the class of y and v joins the class of x. (Note that the path resulting from e is $xuvy$.)

2. If G is elementary and if for some $e \in E(G)$ the graph $G-e$ is still elementary, then $\mathcal{P}(G)$ is a refinement of $\mathcal{P}(G-e)$. To see this, simply observe that if $S \in \mathcal{P}(G)$ then S is an extreme set in $G-e$, for we have $\operatorname{def}(G'-S) \geq \operatorname{def}(G-S) = |S|$ and hence equality must hold.

3. If G is elementary and $e \notin E(G)$ is a line joining two points x and y in $V(G)$ then (a) $\mathcal{P}(G+e)$ refines $\mathcal{P}(G)$ and from Theorem 5.2.2 it is immediate that (b) line $e = xy$ is allowed in $G+e$ if and only if x and y lie in different classes of $\mathcal{P}(G)$.

4. An alternative way to define the partition $\mathcal{P}(G)$ is to show that the relation $\{(x,y) \mid x,y \in V(G)$ and either $x = y$ or $\{x,y\}$ is extreme in $G\}$ is an equivalence relation. (To verify that this is an equivalence relation "from scratch" is no easy matter, but using Theorem 5.2.2(a), it is trivial.) Of course, then, we define $\mathcal{P}(G)$ to be the corresponding equivalence classes. Recall from Exercise 3.3.7 that $\{x,y\}$ is extreme if and only if $G-x-y$ has no perfect matching.

5.2.3. EXERCISE. Prove that if the set $\mathcal{P}(G)$ of maximal barriers partitions $V(G)$ then G is elementary.

In a spirit parallel to that of our inquiry into the degenerate cases of the Gallai-Edmonds decomposition we present the next two results.

5.2.4. EXERCISE. Let G be an elementary graph. Then $\mathcal{P}(G)$ consists of precisely two classes if and only if G is bipartite.

The second degenerate case arises when each class of $\mathcal{P}(G)$ is a singleton. The next theorem tells us that this situation occurs exactly when G is bicritical. (Recall from the introduction to this chapter that G is **bicritical** if and only if G contains a line and for each pair of distinct points x and y in G, $G-x-y$ has a perfect matching.)

5.2.5. THEOREM. *The following are equivalent for any graph G:*

(a) *G is bicritical.*

(b) *G is elementary and all classes of $\mathcal{P}(G)$ are singletons.*

(c) *If* $X \subsetneq V(G)$ *and* $|X| \geq 2$ *then* $c_0(G - X) \leq |X| - 2$.

PROOF. Suppose first that G is bicritical. Then it is 1-extendable and therefore elementary. Suppose $\mathcal{P}(G) = \{S_1, \ldots, S_k\}$ and suppose $|S_1| \geq 2$, say. Then if x and y are points in S_1, we know $\{x, y\}$ is extreme and hence, by Theorem 5.2.2(a), $\text{def}(G - x - y) = 2$ and hence $G - x - y$ has no perfect matching, a contradiction. Hence (b) holds.

Next suppose that G is elementary with only singleton classes in $\mathcal{P}(G)$. By Tutte's Theorem we have $c_0(G - X) \leq |X|$ for all $X \subsetneq V(G)$. Suppose then that there is an $X \subsetneq V(G)$ with $|X| \geq 2$, but $c_0(G - X) \geq |X| - 1$. Since $|V(G)|$ is even, $c_0(G - X) \neq |X| - 1$ by parity, so $c_0(G - X) = |X|$ and hence X is a barrier in G. But then X lies in a class S of $\mathcal{P}(G)$ and so $|S| \geq 2$, a contradiction.

Finally, choose any two points x and y in $V(G)$ and let $G' = G - x - y$. Suppose $X' \subsetneq V(G')$, but that $c_0(G' - X') \geq |X'| + 1$. Then if $X = X' \cup \{x, y\}$, we have $c_0(G - X) = c_0(G' - X') \geq |X'| + 1 = |X| - 1$, contradicting (c). Thus $c_0(G' - X') \leq |X'|$ for all $X' \subsetneq V(G')$ and by Tutte's Theorem, graph G' has a perfect matching. Thus G is bicritical and the proof is complete. ∎

We now proceed to use $\mathcal{P}(G)$ to develop a canonical decomposition (or construction) procedure for all elementary graphs.

5.2.6. THEOREM. *Let G be elementary, suppose $S \in \mathcal{P}(G)$, $|S| \geq 2$, and let H be any component of $G - S$. Then*

(a) *the bipartite graph G'_S obtained from G by contracting each component of $G - S$ to a single point and deleting each line spanned by S is elementary,*

(b) *the graph H' obtained from G by contracting the set $V(G) - V(H)$ to a single point u_H is elementary, and*

(c) $\mathcal{P}(H') = \{\{u_H\}\} \cup \{T \cap V(H) \mid T \in \mathcal{P}(G)\}$.

This theorem describes a way to decompose the graph G into smaller elementary graphs H' and the elementary bigraph G'_S which functions as a "frame" for reassembling the H's to recover G. We shall refer to this operation as the **decomposition of G with respect to S**. (See Figure 5.2.1.) If one of the elementary graphs H' is not bicritical, we select a class of $\mathcal{P}(H')$ with more than one point and repeat the decomposition. We continue until a list of bicritical graphs is obtained. This procedure is called the **bicritical decomposition of G**.

PROOF (of Theorem 5.2.6). Let M be a perfect matching of G. Then no line of M lies in S and by Theorem 5.2.2(d), M must match each point of S to a point of a different component of $G - S$. So each

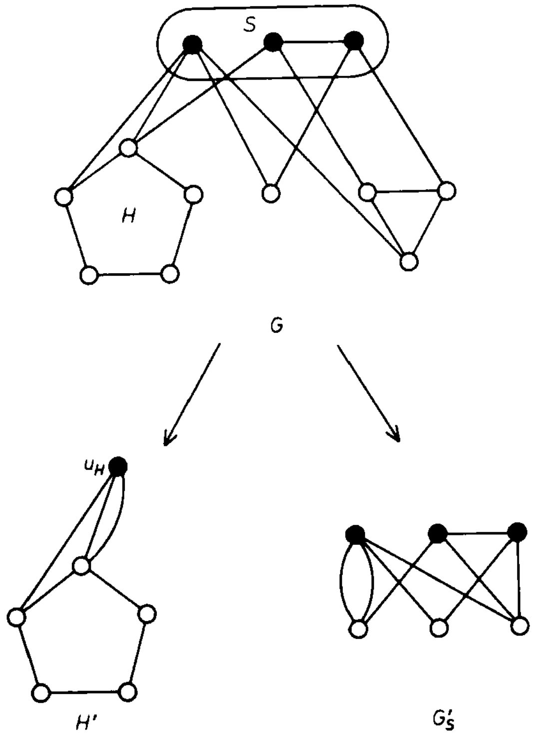


FIGURE 5.2.1.

perfect matching of G is mapped onto a perfect matching of G'_S by the contraction.

On the other hand, it is clear that each perfect matching of G is mapped onto a perfect matching of H'. So an allowed line of G, if not contracted, corresponds to allowed lines in G'_S and H'. But a contraction of a connected graph is connected and hence (a) and (b) follow.

To prove (c), it will suffice to show (i) that if $x, y \in V(H)$ then $G - x - y$ has a perfect matching if and only if $H' - x - y$ has one, and (ii) if $x \in V(H)$ then $H' - x - u_H$ has a perfect matching.

In order to show (i) let x and y be arbitrary points of $V(H)$ and suppose first that $G - x - y$ has a perfect matching M. Since $H - x - y$ is odd, M must contain exactly one line from H to S. So M is mapped onto a perfect matching of $H' - x - y$.

Conversely, suppose $H' - x - y$ has a perfect matching M'. Then M' corresponds to a matching M in G covering $H - x - y$ and having exactly one line e joining $H - x - y$ to S. But then e must join two different classes of $\mathcal{P}(G)$, so by Theorem 5.2.2(b), e is allowed in G. So let M_e be a perfect matching of G containing e and let f be the line of M' covering point u_H. But then e is the only line of M_e joining S to H by Theorem 5.2.2(d), so $(M_e - E(H)) \cup (M' - \{f\})$ is a perfect matching of $G - x - y$.

To prove (ii) simply observe that $H' - u_H = H$ is factor-critical by Theorem 5.2.2(d) and hence $H - x = H' - x - u_H$ has a perfect matching.

■

This will be enough to provide us with a procedure for constructing all elementary graphs from two families of basic "building block" graphs, namely elementary bipartite graphs and bicritical graphs.

5.2.7. THEOREM. *A graph is elementary if and only if it can be built up from elementary bipartite graphs and from bicritical graphs by iterating the following construction:*

CONSTRUCTION. *Let $G_0 = (U(G_0), W(G_0))$ be an elementary bipartite graph with more than two points. With each point $w \in W(G_0)$ associate a (previously constructed) elementary graph $G(w)$ which has the property that $\mathcal{P}(G(w))$ contains a one-element equivalence class $\{v_w\}$ such that $\deg_{G(w)}(v_w) = \deg_{G_0}(w)$. Now splice $G(w)$ to G_0 at v_w and w as follows. Let $A(w)$ and $B(w)$ denote the sets of points joined to w in G_0 and to v_w in $G(w)$ respectively. Now delete v_w and w and and construct any bipartite graph on $(A(w), B(w))$ requiring only that every point of $A(w) \cup B(w)$ is covered by as many new lines of this newly constructed bigraph as the number of deleted lines which covered it. Repeat this construction for each $w \in W(G_0)$. Finally, add lines joining pairs of points in $U(G_0)$ arbitrarily.*

We illustrate this construction in Figure 5.2.2.

PROOF (of Theorem 5.2.7). The "only if" part of the proof is obvious by Theorem 5.2.6.

Conversely, suppose G is obtained via the Construction from smaller elementary graphs. We show that G is elementary.

Suppose a_1b_1 is a line of G with $a_1 \in A(w_1)$ and $b_1 \in B(w_1)$. Then a_1w_1 is a line of G_0 and hence is allowed in G_0. Let $F_0 = \{a_1w_1, \ldots, a_kw_k\}$ be a perfect matching of G_0 containing a_1w_1. We construct a matching F_1 covering $U(G_0)$ in G by including a line joining each a_j to some b_j where $b_j \in B(w_j)$. Note that these lines a_jb_j all exist because, by the rules of the Construction, every point of $A(w_j)$ is covered by a line to $B(w_j)$. Now extend F_1 to a perfect matching of G by recalling that in each $G(w)$, every line incident with point v_w is allowed in $G(w)$ (since $\{v_w\}$ is a singleton class in $\mathcal{P}(G(w))$). Thus each $G(w_j) - v_j - b_j$ has a perfect matching. Thus all lines in G with exactly one endpoint in G_0 are allowed in G.

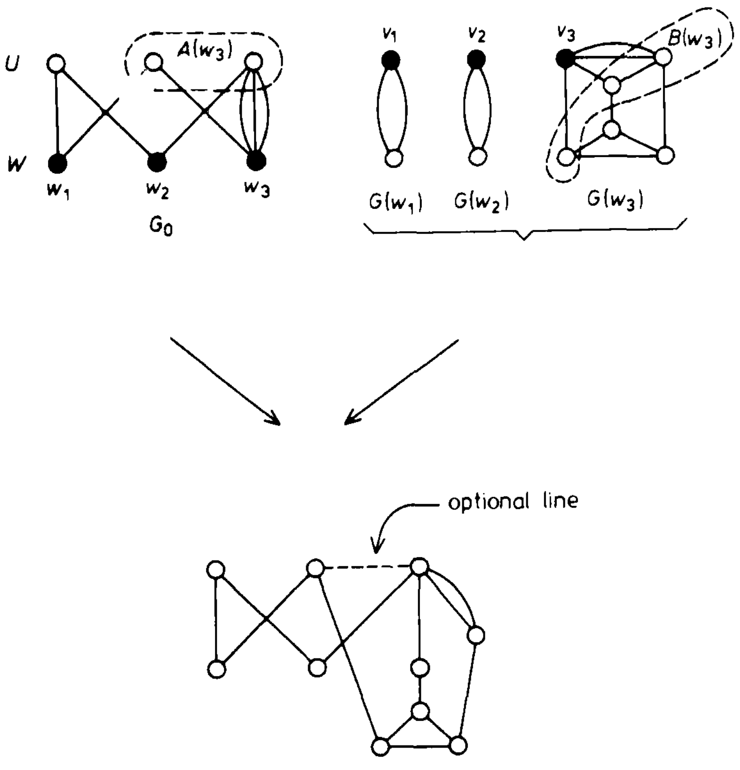


FIGURE 5.2.2.

By a similar argument, the allowed lines of each $G(w)$ are allowed in G and hence it follows that G is elementary. (Actually we have proved the slightly stronger result that if $\hat{G}$ results from G by deleting all lines with both endpoints in $U(G_0)$, then $\hat{G}$ is already elementary.) Finally we claim that $U(G_0) \in \mathcal{P}(G)$. To see this, recall that each $\{v_w\} \subsetneq V(G(w))$ is a singleton class in $\mathcal{P}(G(w))$. Thus by Theorem 5.2.2(d), graph $G(w) - v_w$ consists of $|\{v_w\}| = 1$ factor-critical component (i.e., $G(w) - v_w$ is factor-critical) and hence $G - U_0$ has $|U_0| = k$ factor-critical components. ∎

It is natural to ask in what sense is the above Construction of a given elementary graph unique. If elementary graph G is not bicritical, then $\mathcal{P}(G)$ always contains a class S with $|S| \geq 2$, but of course there may be

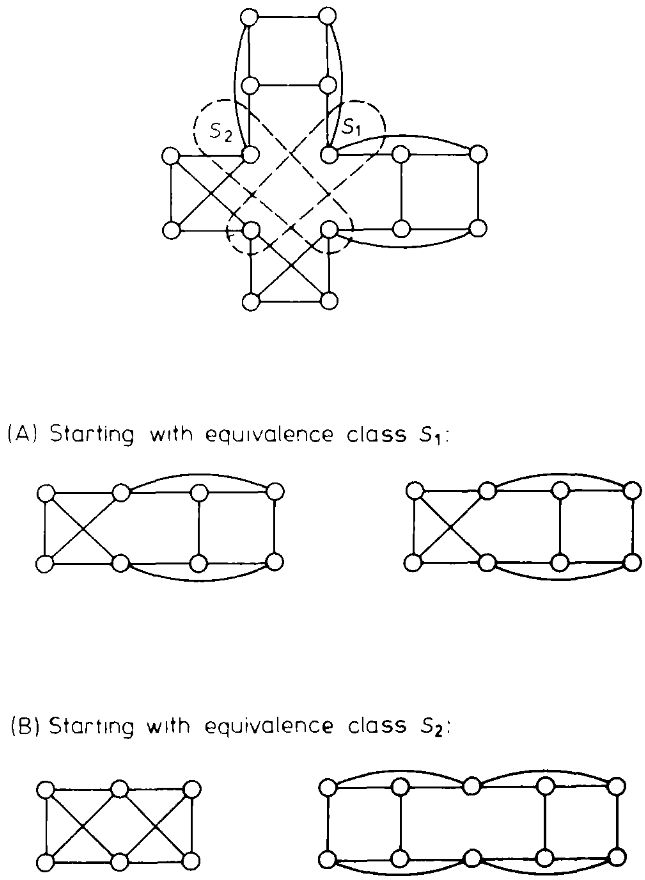


FIGURE 5.2.3.

many such classes. Is the collection of bicritical graphs obtained by the bicritical decomposition procedure of Theorem 5.2.6 independent of the class S chosen? In Figure 5.2.3 we may see that this is not the case.

We can decompose bicritical graphs one step further into 3-connected bicritical graphs called **bricks**. First, however, we need some definitions.

Let G be a 2-connected graph and let $\{x,y\}$ be a cutset of G. Then G can be written as $G_1' \cup G_2'$, where G_1' and G_2' are connected graphs different from K_2 and $V(G_1') \cap V(G_2') = \{x,y\}$. Let $G_1 = G_1' + xy$ and $G_2 = G_2' + xy$. If either one of G_1 and G_2 is not 3-connected, then we take one of its 2-element cutsets and repeat this procedure.

Thus G can be decomposed (in the sense described above) into 3-connected graphs and triangles, possibly with multiple lines. We shall call these the **3-blocks** of G. We remark that the 3-blocks of a graph are uniquely determined up to multiplicity of lines; that is, the final list of 3-blocks does not depend upon which 2-element cutset is chosen at each step of the decomposition procedure. (See Hopcroft and Tarjan (1973a).) This fact, however, will not be needed here and if we talk about the 3-blocks of a graph, we shall have in mind any list obtained by a decomposition procedure.

5.2.8. LEMMA. *A graph is bicritical if and only if it is K_2, or it is 2-connected and each of its 3-blocks is bicritical.*

PROOF. It suffices to prove that if $G = G_1' \cup G_2'$, where G_1' and G_2' are connected graphs different from K_2, $V(G_1') \cap V(G_2') = \{x,y\}$ and $G_i = G_i' + xy$, then G is bicritical if and only if both G_1 and G_2 are bicritical.

First suppose that G is bicritical. Then $G - x - y$ has a perfect matching, so G_1 and G_2 must both be even and each $G_i - x - y$ has a perfect matching.

Suppose $u, v \in V(G_i)$. Then $G - u - v$ has a perfect matching M. Since $|V(G_i')|$ is even, $M \cap E(G_i')$ covers either all points of $G' - u - v$ or misses precisely two, namely x and y. In this second case, $M + xy$ is a perfect matching of $G_i - u - v$. So G is bicritical.

Second, assume that both G_1 and G_2 are bicritical and suppose $u, v \in V(G)$. If u and v both lie in $V(G_1)$, say, then consider a perfect matching M_1 of $G - u - v$. If the line xy of $E(G_1) - E(G_1')$ belongs to M_1, then $M_1 - xy$, together with a perfect matching of G_2, forms a perfect matching of $G - u - v$. If M_1 does not contain xy, then M_1 and a perfect matching of $G_2 - x - y$ form a perfect matching of $G - u - v$.

If $u \in V(G_1) - \{x,y\}$ and $v \in V(G_2) - \{x,y\}$, say, then let M_1 be a perfect matching of $G_1 - u - x$ and M_2, a perfect matching of $G - v - y$. Then $M_1 \cup M_2$ is a perfect matching of G. Thus G is bicritical. ■

Lemma 5.2.8 implies that the 3-blocks of a bicritical graph are bricks. Currently this, then, is the limit of our decomposition procedure.

The results of Section 7.6 will imply that the number of bricks obtained, when decomposing an elementary graph, is invariant. In fact, we can actually show that the *collection* of bricks itself is uniquely determined.

If we start with a *saturated* elementary graph then the collection of bicritical building blocks, 3-connected or not, is also uniquely determined. See Theorem 5.3.7 for the details.

5.2.9. EXERCISE. A graph is factor-critical if and only if it is connected and each of its blocks is factor-critical. ∎

5.3. Saturated Graphs and Cathedrals

In this section all graphs will be assumed to be *simple*. A graph G with a perfect matching is said to be **saturated** if $G+e$ has more perfect matchings than G for all lines $e \in E(\overline{G})$. Thus, in particular, a saturated graph must be connected. In Figure 5.3.1 we show two saturated graphs: G_0 which is elementary and G which is not. Note that a 1-extendable graph is saturated if and only if it is bicritical.

From any graph G with a perfect matching we can obtain a saturated graph by joining pairs of non-adjacent points as long as no new perfect matchings are produced. Many important graph-theoretic properties are preserved by this procedure. But the most important point for us is that the set of perfect matchings is unchanged. The structural results of this section could be extended to unsaturated graphs by saturating them as above.

At the end of Section 5.2 we developed a construction procedure for obtaining any elementary graph using bicritical graphs and elementary bipartite graphs as fundamental building blocks. In the present section we shall first investigate that construction procedure when the elementary graph produced is saturated and then we shall extend the construction to obtain all saturated *non*-elementary graphs. Our approach follows closely that initiated by Kotzig (1959a, 1959b, 1960) and further developed by Lovász (1972a).

We begin with three lemmas about saturated graphs in general which we shall eventually need to analyze their structure. For convenience we denote $A(G-x) \cup \{x\}$ by $A^*(G-x)$ for every $x \in V(G)$.

5.3.1. LEMMA. *Let G be any saturated graph and let $x \in V(G)$. Then:*

(a) *$D(G-x)$ consists precisely of those points of $V(G)-\{x\}$ which are not adjacent to x or which are joined to x by an allowed line of G.*

(b) *Every point of* $A^*(G-x) \cup C(G-x)$ *is adjacent to every other point of* $A^*(G-x)$ *and all such adjacencies are forbidden lines in* G.

(c) $C(G-x)$ *induces a saturated graph.*

PROOF. Since G has a perfect matching, we have $V(G) = C(G)$ and thus by Lemma 3.3.10, the set $A^*(G-x)$ is a barrier in G. But again since G has a perfect matching, if we delete $A^*(G-x)$ from G, we must have exactly $|A^*(G-x)|$ odd components remaining. But then every perfect matching of G matches all of $A^*(G-x)$ into $D(G-x)$, and hence no line joining $C(G-x)$ to $A^*(G-x)$ is allowed in G. If we add any new line f joining $A^*(G-x)$ to $A^*(G-x)$ or $A^*(G-x)$ to $C(G-x)$, the same reasoning shows that f is forbidden in $G+f$. This proves part (b).

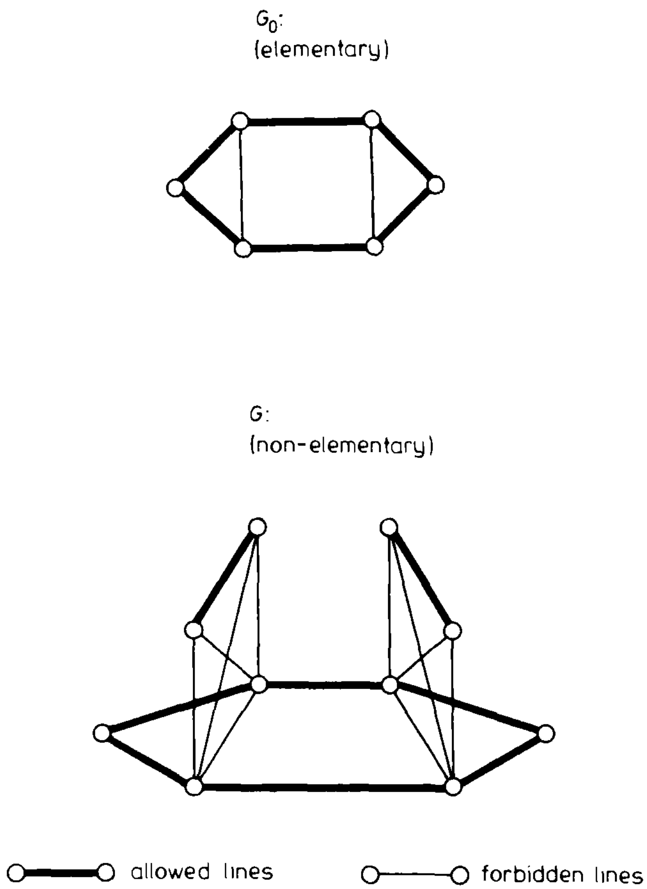


FIGURE 5.3.1.

If any new line f is added joining two points of $C(G-x)$, since G is saturated, f must lie in a new perfect matching M_f of $G+f$. But then since all lines joining $C(G)$ to $A^*(G-x)$ are forbidden, line f in fact must lie in a new perfect matching of $C(G-x)$. Hence (c) is proved.

Now suppose $y \in D(G-x)$ and suppose y is adjacent to x. Let M be a maximum matching of $G-x$ which misses y. Then $M \cup \{xy\}$ is a perfect matching of G and hence xy is allowed in G. So if D' is the set of all points in $V(G)-\{x\}$ which are not adjacent to x or are joined to x via an allowed line, we have shown $D(G-x) \subseteq D'$.

The converse inclusion is obvious by part (b). This completes the proof of (a) and hence the proof of the lemma. ∎

5.3.2. LEMMA. *Let G be saturated and $x \in V(G)$. Then:*

(a) *if $z \in A(G-x)$, then $A^*(G-z) \subseteq A^*(G-x)$ and $D(G-z) \subseteq D(G-x)$, and*

(b) *if $z \in C(G-x)$, then $A^*(G-z) \supseteq A^*(G-x)$ and $D(G-z) \supseteq D(G-x)$.*

PROOF. First suppose $z \in A(G-x)$. Then by Lemma 5.3.1(b) all points of $A^*(G-x) \cup C(G-x)$ are joined to z by forbidden lines of G. Suppose $y \in D(G-z)$. Then, by Lemma 5.3.1(a), either y is not adjacent to z or line yz is allowed in G. But if $y \in A^*(G-x) \cup C(G-x)$, then by Lemma 5.3.1(b), line yz is forbidden in G, a contradiction. Thus $y \in D(G-x)$ and hence $D(G-z) \subseteq D(G-x)$.

Now suppose $x \in A(G-z)$. Then an argument symmetric to that in the preceding paragraph yields $D(G-x) \subseteq D(G-z)$. Thus $D(G-z) = D(G-x)$ and hence $A^*(G-z) = A^*(G-x)$.

On the other hand, suppose $x \notin A^*(G-z)$. We want to show $A^*(G-z) \subseteq A^*(G-x)$. Now if $x \in D(G-z)$, then $x \in D(G-x)$ which is absurd. So $x \in C(G-z)$ and hence by Lemma 5.3.1(b) all points of $A^*(G-z)$ are joined to x by a forbidden line of G. But no point of $D(G-x)$ is joined to x by a line forbidden in G by Lemma 5.3.1(a), hence $A(G-z) \cap D(G-x) = \emptyset$. Moreover, no point of $C(G-x)$ is adjacent to any point of $D(G-x)$ and hence no point of $C(G-x)$ is adjacent to any point of $D(G-z)$. Thus $C(G-x) \cap A(G-z) = \emptyset$ and hence it follows that $A(G-z) \subseteq A(G-x)$, completing the proof of (a).

To prove (b), assume $z \in C(G-x)$. Then no point of $D(G-x)$ is adjacent to z and hence $D(G-x) \subseteq D(G-z)$ by Lemma 5.3.1(a) applied to $G-z$. Using Lemma 5.3.1(b), we know that every point of $A(G-x)$ is joined to z by a forbidden line. So $A(G-x) \cap D(G-z) = \emptyset$. Moreover, each point of $A(G-x)$ is joined to at least one point of $D(G-x)$ and consequently to at least one point of $D(G-z)$. Thus $A(G-x) \subseteq A(G-z)$ and the proof is complete. ■

5.3.3. LEMMA. *Let G be any graph having a perfect matching. Then G is saturated if and only if every extreme set in G induces a complete graph, all lines of which are forbidden in G.*

PROOF. First assume G is saturated, X is extreme in G and $\{x, y\} \subseteq X$. Thus by Exercise 3.3.3, $\{x, y\}$ is also extreme in G.

Suppose x and y are not adjacent in G. Then $\operatorname{def}(G + xy - x - y) = \operatorname{def}(G - x - y) = \operatorname{def}(G) + 2$ since $\{x, y\}$ is extreme in G. But $\operatorname{def}(G + xy) = \operatorname{def}(G) = 0$, so we have $\operatorname{def}(G + xy - x - y) = \operatorname{def}(G + xy) + 2$ and hence $\{x, y\}$ is extreme in $G + xy$. Thus xy is forbidden in $G + xy$ by Exercise 3.3.7, and hence G is not saturated, a contradiction. So we must have $xy \in E(G)$ and thus xy is forbidden in G by Exercise 3.3.7.

Conversely, suppose G is not saturated and thus there are two points $x, y \in V(G)$ with $xy \notin E(G)$, such that G and $G + xy$ have the same perfect matchings. Thus xy is forbidden in $G + xy$. But again using Exercise 3.3.7, $\{x, y\}$ is extreme in $G + xy$. Hence $\operatorname{def}(G - x - y) = \operatorname{def}(G + xy - x - y) = \operatorname{def}(G + xy) + 2 = 0 + 2 = \operatorname{def}(G) + 2$ and thus $\{x, y\}$ is extreme in G. But this contradicts the hypothesis that xy must be a forbidden line of G. ■

Now let us assume further that our saturated graphs are elementary. As a first result we have the following immediate consequence of Lemma 5.3.3 which tells us exactly how the forbidden lines of G are distributed vis-à-vis the classes of the partition $\mathcal{P}(G)$.

5.3.4. THEOREM. *If G is a saturated elementary graph, then the forbidden lines of G constitute point-disjoint complete subgraphs induced by the classes of $\mathcal{P}(G)$.* ■

5.3.5. EXERCISE. Show that if G is a saturated elementary graph and $S \in \mathcal{P}(G)$, then $\mathcal{P}(G)$ must have at least $|S|$ singleton classes different from S.

Now let us recall the decomposition procedure developed in Section 5.2. In particular, we saw there that the bicritical building blocks were not uniquely determined. But if G happens to be saturated they are unique in the sense of Theorem 5.3.7 below.

5.3.6. LEMMA. *Let G be a saturated elementary graph, $S \in \mathcal{P}(G)$ and let H be one of the components of $G-S$. Let H' be obtained by contracting $V(G)-V(H)$ to a single point. Then H' is a saturated elementary graph. Moreover, if $X \in \mathcal{P}(G)$ is such that $X \cap V(H) \neq \emptyset$, then $X \subseteq V(H)$ and, in fact, $X \in \mathcal{P}(H')$.*

PROOF. Graph H' is elementary by Theorem 5.2.6. Suppose we add a new line e' to H'. Let e be the corresponding line added to G. Since G is saturated there is a perfect matching M_e of G containing e. But from Theorem 5.2.2(d), we know M_e matches each point of S to a different component of $G-S$ and hence M_e corresponds to a perfect matching of H' containing line e'. So H' is saturated.

Now suppose $X \in \mathcal{P}(G)$ and $X \cap V(H) \neq \emptyset$. (Of course then $X \neq S$.) Suppose also that $X \cap V(H_1) \neq \emptyset$ for some $H_1 \neq H$, where H_1 is another component of $G-X$. But since no lines of G can join points lying in different components of $G-S$, it follows that $G[X]$ is a disconnected subgraph of G contradicting the fact that $G[X]$ is complete according to Theorem 5.3.4. Thus $X \subseteq V(H)$ and, in fact, by Theorem 5.2.6(c), $X \in \mathcal{P}(H')$. ∎

To underline the importance of Lemma 5.3.6 let us describe once more what it implies. It tells us that if G is saturated elementary, if $S \in \mathcal{P}(G)$, if H is a component of $G-S$ and if $S_{i_1}, \ldots, S_{i_k}$ are the classes of $\mathcal{P}(G)$ such that $S_{i_j} \cap V(H) \neq \emptyset$, then, in fact, each such S_{i_j} lies *entirely within* H and, moreover, $\mathcal{P}(H') = \{\{t_S\}, S_{i_1}, \ldots, S_{i_k}\}$, where t_S is the point obtained by contracting $G-H$ to a point. Since in addition, H' is saturated elementary, we begin to see the basis for an inductive reduction procedure here which helps us prove the next theorem. This next result tells us that the Construction of general elementary graphs discussed in Section 5.2 is, in a sense, "more unique" when the graph is saturated.

Graph G'_S is as defined in the statement of Theorem 5.2.6.

5.3.7. THEOREM. *Let G be a saturated elementary graph and let G^* be the graph obtained by contracting each class S of $\mathcal{P}(G)$ to a point t_S. Then:*

(a) *the blocks of G^* (i.e., the maximal 2-connected subgraphs of G^*) are bicritical,*
(b) *the number of blocks of G^* containing t_S is exactly $|S|$,*
(c) *If one builds G by iterating the construction of Theorem 5.2.7, the initial bicritical building block graphs must be precisely the blocks of G^*, and*

(d) *the elementary bipartite graphs used in the last step of the Construction of G are uniquely determined in the sense that they must be of the form G'_S, for some $S \in \mathcal{P}(G)$.*

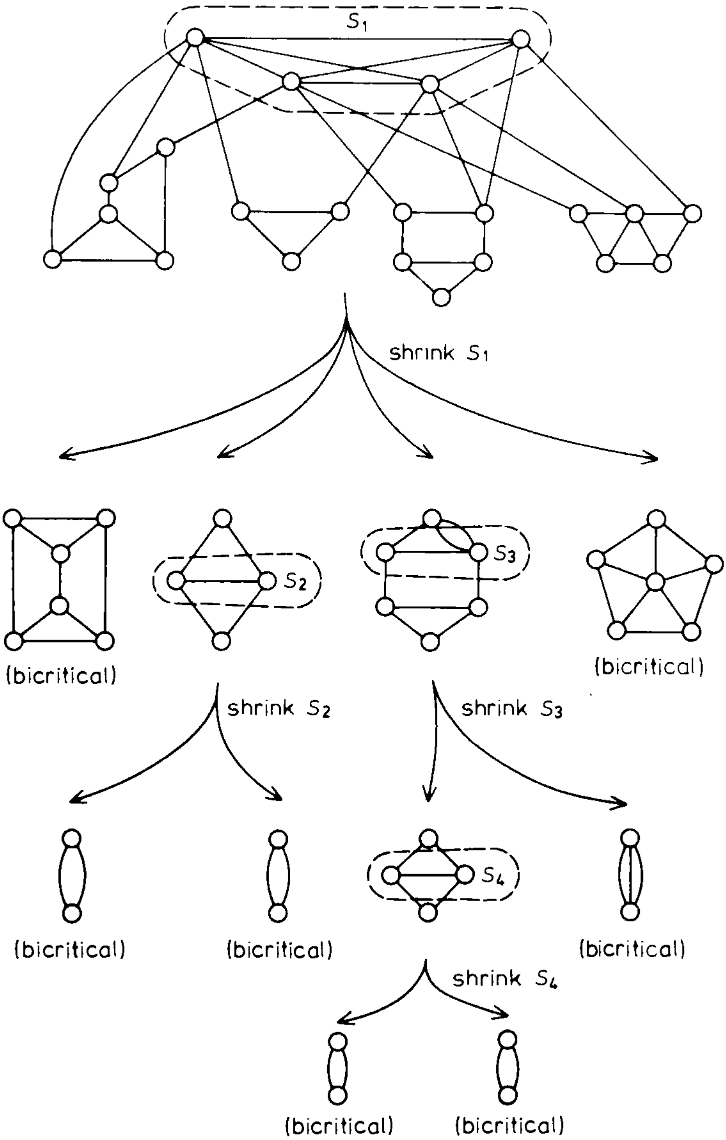
PROOF. Choose any $S \in \mathcal{P}(G)$ and contract S to a single point to obtain a graph G'. Then by Lemma 5.3.6, the blocks of G' are saturated elementary graphs and every class of $\mathcal{P}(G) - \{S\}$ is a class in $\mathcal{P}(H)$, for one of these blocks H. If S' is a non-singleton class in $\mathcal{P}(H)$, shrink it to a point to obtain an even smaller graph all blocks of which are saturated elementary. Continuing in this manner, we ultimately arrive at a graph G^* such that all blocks of G^* are bicritical. So (a) is proved.

Part (b) follows immediately from Theorem 5.2.2(b).

Now suppose we have built G via the Construction and that the last graph in the procedure is the elementary bipartite graph $G_0 = (U_0, W_0)$. Then recall from the proof of Theorem 5.2.7 that $U_0 \in \mathcal{P}(G)$. This proves part (d).

But now from Lemma 5.3.6 we know that for any $S \in \mathcal{P}(G)$ and any component H of $G - S$, graph H' (where H' is as in Lemma 5.3.6) is saturated elementary and has, as its partition classes, $\{t_S\}$ and some of the classes in $\mathcal{P}(G)$. But graph H' is then either bicritical or is itself built up using the Construction, where we must use a non-singleton class of $\mathcal{P}(H')$ as the beginning "U_0". But this class of $\mathcal{P}(H')$ is also a class in $\mathcal{P}(G)$! Thus we may view the given Construction of G in reverse, so to speak, by successively shrinking non-singleton classes of G to single points, while any remaining non-singleton classes remain non-singleton classes in the shrunken graph. Ultimately we will shrink all non-singleton classes of G to single points and obtain G^*. After each shrinking step, the block graphs, that is, those playing the role of a H', remain elementary and we stop only when all classes of each H' at hand are singletons. But this means all such blocks are bicritical by Theorem 5.2.5(b). Thus no matter how we start to dissect a Construction of graph G, we must end the dissection with the blocks of G^* as the initial bicritical building blocks. This proves part (c) and, consequently, the theorem. ∎

It is instructive, we think, to study the Construction of the saturated elementary graph G shown in Figure 5.3.2. The reader is invited to convince himself that if he dissects G by shrinking non-singleton classes of G in an order different from the one we chose, he will nevertheless end up with the bicritical blocks of G^* as his beginning bicritical building blocks just as we did.



S_1
shrink S_1
S_2
S_3
(bicritical)
(bicritical)
shrink S_2
shrink S_3
S_4
(bicritical)
(bicritical)
(bicritical)
shrink S_4
(bicritical)
(bicritical)

FIGURE 5.3.2.

We are now prepared to describe a canonical construction procedure for all saturated non-elementary graphs.

THE CATHEDRAL CONSTRUCTION. Let G_0 be any saturated elementary graph. To each class $S \in \mathcal{P}(G_0)$ assign an (already constructed) saturated graph G_S or the empty set. For each $S \in \mathcal{P}(G_0)$ join every point of S to every point of G_S. In the case when G_S is not empty we will call the subgraph G_S the **tower** over S and S, the **foundation** of that tower. (Note that a tower and its foundation are point-disjoint.)

We now present our main result on saturated non-elementary graphs.

5.3.8. THEOREM. (*The Cathedral Theorem*).

(a) *Every graph G, built up by iterating the Cathedral Construction using smaller saturated graphs, is itself saturated.*

(b) *The allowed lines of G are precisely those lines which are allowed in one of the elementary graphs used in one of the steps.*

(c) *Conversely, if G is any saturated graph, it can be built up using the Cathedral Construction starting with a saturated elementary graph G_0 and a collection of $|\mathcal{P}(G_0)|$ smaller saturated graphs (some perhaps empty) already constructed. The graph G_0 may be uniquely described as the subgraph of G induced by those points of G which, for each $x \in V(G)$, do not lie in $C(G-x)$.*

Before presenting the proof of this theorem, we direct the reader to Figure 5.3.3 for a better understanding of the Cathedral structure. Note that the large plus signs indicate that each point in the partition class below it is joined to *every* point in the tower above it. Of course in general we may have towers nested within towers in the Cathedral structure. Claims 2, 3, 4 and 5 below assert further important properties of this construction.

PROOF (of Theorem 5.3.8).

Claim 1. The Cathedral Construction yields a saturated graph.

Let S be any class in $\mathcal{P}(G_0)$. Clearly the graph G constructed has a perfect matching. Moreover, we cannot add any line joining two points of a G_S nor of G_0 without producing new perfect matchings.

Fix some $S_0 \in \mathcal{P}(G_0)$. Suppose $x \in V(G_{S_0})$ and $y \in V(G_0)$ are not adjacent in G. Then $y \notin S_0$. Choose a point $z \in S_0$ and let F_0 be a perfect matching in $G_0 - y - z$. For each $S \in \mathcal{P}(G)$ let F_S be a perfect matching of G_S. Let xu be the line of F_{S_0} incident with x. Then $((\bigcup_{S\in\mathcal{P}(G_0)} F_S) \cup F_0 \cup \{uz\}) - \{ux\}$ is a perfect matching of $G-x-y$. Hence $\Phi(G+xy) > \Phi(G)$.

Finally, suppose $x \in V(G_{S_1})$ and $y \in V(G_{S_2})$ for some S_1 and $S_2 \in \mathcal{P}(G_0)$. Let F_{S_i} be a perfect matching of G_{S_i} for each $S_i \in \mathcal{P}(G_0)$. Suppose xu_1 is the line of F_{S_1} covering x and yu_2 is the line of F_{S_2} covering y. Choose $z_1 \in S_1$ and $z_2 \in S_2$ arbitrarily. Let F_0 be a perfect

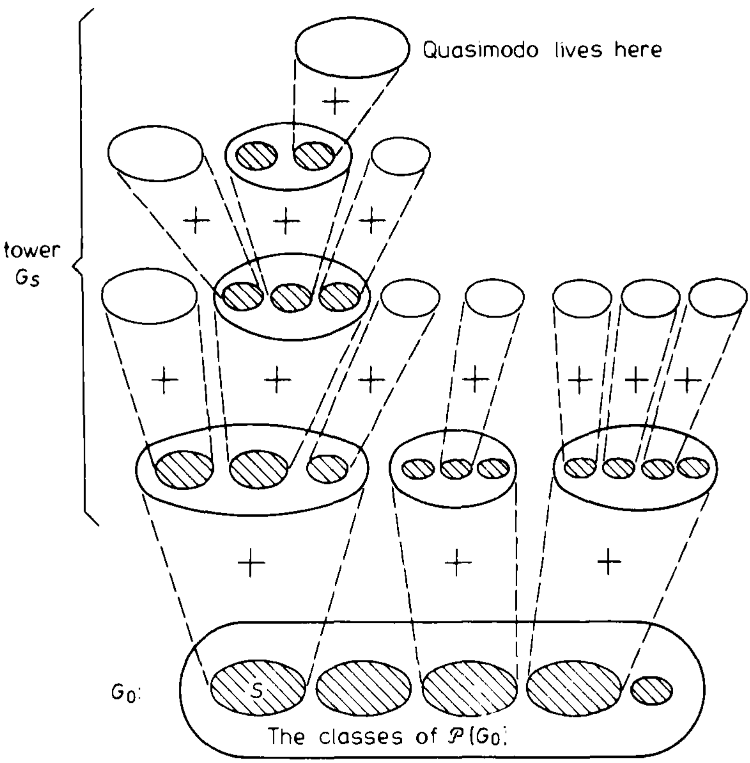


FIGURE 5.3.3. A Cathedral Construction

matching of $G_0 - z_1 - z_2$. Then $\bigcup_{S \in \mathcal{P}(G_0)} F_S - \{xu_1, yu_2\} \cup \{u_1z_1, u_2z_2\} \cup F_0$ is a perfect matching of $G - x - y$. So once again $\Phi(G + xy) > \Phi(G)$ and the proof of Claim 1 is complete. So (a) holds.

Claim 2. Every perfect matching of G consists of a perfect matching of G_0 and of a perfect matching of each G_S, $S \in \mathcal{P}(G)$.

To see this, let F be a perfect matching of G. Now $G-S$ consists of a component G_S and $|S|$ other components, each of which contains a component T of G_0-S and also contains any non-empty $G_{S'}$'s for which $S' \subseteq V(T)$. But all such towers $G_{S'}$ must have an even number of points (since they have perfect matchings). So each component of $G-S$, other than G_S, is odd and since there are $|S|$ of them, each is joined to S by a line of F. Therefore, these lines cover S and hence all lines joining S and G_S are forbidden. Thus Claim 2 is proved and so is part (b) of the theorem.

Claim 3. For each $x \in S \in \mathcal{P}(G_0)$ we have $A^*(G-x) = S$ and $C(G-x) = V(G_S)$.

By Claim 2, we know that all points of $V(G_S)$ and $S-\{x\}$ are joined to x by forbidden lines. Since, by Claim 1, G is saturated we may use Lemma 5.3.1(a) to conclude that all such points lie in $A(G-x) \cup C(G-x)$ and that all other points of $V(G)$ belong to $D(G-x)$. It follows then that $V(G_S) \cup (S-\{x\}) = A(G-x) \cup C(G-x)$. But now we know that the neighbors of $D(G-x)$ in G form the set S. Hence $A^*(G-x) = S$ and it then follows that $C(G-x) = V(G_S)$. Therefore Claim 3 is proved.

Claim 4. If $S \in \mathcal{P}(G_0)$ and $x \in V(G_S)$, then $A^*(G-x) = A^*(G_S - x) \cup S$ and $C(G-x) = C(G_S - x)$.

We know G_S is saturated and so by Lemma 5.3.1, $D(G_S - x)$ consists of all points of $V(G_S) - \{x\}$ not adjacent to x or joined to x by a line allowed in G_S. Let y be such a point. First suppose y is not adjacent to x in G_S. Then y is not adjacent to x in G either. On the other hand, if y is joined to x by a line allowed in G_S, this line is also allowed in G, by Claim 2. But by Claim 1, G is saturated and hence by Lemma 5.3.1 (applied to graph G), $y \in D(G-x)$; that is, $D(G_S - x) \subseteq D(G-x)$. So $D(G_S - x) \subseteq D(G-x) \cap V(G_S)$.

On the other hand, if $y \in D(G-x) \cap V(G_S)$, then either y is not adjacent to x in G_S or line xy is allowed in G, and hence allowed in G_S by Claim 2. So $y \in D(G_S - x)$ by Lemma 5.3.1. Hence $D(G-x) \cap V(G_S) \subseteq D(G_S - x)$ and therefore $D(G_S - x) = D(G-x) \cap V(G_S)$.

Also note that $D(G-x) \cap S = \emptyset$. But then $A^*(G-x) = A^*(G_S - x) \cup S$. Thus since $\{C(G-x), A^*(G-x), D(G-x)\}$ is a partition of $V(G)$, it follows that $C(G-x) = C(G_S - x)$ and Claim 4 is proved.

Claim 5. $V(G_0)$ is precisely the set of those points $y \in V(G)$ such that for all $x \in V(G)$, $y \notin C(G-x)$.

Claims 3 and 4 imply that if $y \in V(G_0)$, then $y \notin C(G-x)$. Conversely, suppose $y \notin V(G_0)$. So $y \in V(G_{S'})$ for some $S' \in \mathcal{P}(G_0)$. Let $x \in S'$. Then $C(G-x) = V(G_{S'})$ by Claim 3 and hence $y \in C(G-x)$.

Now we proceed to prove part (c) of the theorem. Let G be an arbitrary saturated graph. If G is elementary, there is nothing to prove; let $G = G_0$. So suppose G is not elementary. Motivated by Claim 5, define G_0 to be the subgraph of G induced by those points of G which are not contained in any $C(G-x)$, $x \in V(G)$. Let us agree to call the sets $C(G-x)$, for each $x \in V(G_0)$, **pseudo-towers** and the sets $A^*(G-x)$ the **pseudo-foundations**. (As the reader probably suspects, we will show every pseudo-tower to be a tower and each pseudo-foundation to be a foundation!)

Every point of a pseudo-tower is joined to every point of its pseudo-foundation by a forbidden line according to Lemma 5.3.1(b), but no other point is adjacent to any point of the pseudo-tower (since the remaining points all lie in $D(G-x)$). Thus a pseudo-tower defines its pseudo-foundation uniquely.

Conversely, each pseudo-foundation $A^*(G-x)$ of a pseudo-tower determines the pseudo-tower $C(G-x)$ by the Gallai-Edmonds Theorem. By Lemma 5.3.1(b), any pseudo-foundation $A^*(G-x)$ induces a complete graph with forbidden lines only and any pseudo-tower $C(G-x)$ induces a saturated (and thus connected) graph, by Lemma 5.3.1(c).

Claim 6. Different pseudo-towers have disjoint pseudo-foundations.

(Note that at this point we do not know if *different* pseudo-towers are in fact *disjoint*, but we shall see a bit further on in this proof that this is true.)

We shall show, in fact, that if $z \in A^*(G-x)$, $x \in V(G_0)$, then $A^*(G-z) = A^*(G-x)$. To see this, suppose $z \in A^*(G-x)$. If $z = x$ we are done, so suppose $z \neq x$. Then $z \in A(G-x)$. Suppose, however, that $x \notin A^*(G-z) \cup C(G-z)$. Thus $x \in D(G-z)$. But $D(G-z) \subseteq D(G-x)$ by Lemma 5.3.2(a) and so $x \in D(G-x)$, a contradiction. So $x \in A^*(G-z) \cup C(G-z)$. But $x \in V(G_0)$ and hence $x \notin C(G-z)$. Thus $x \in A^*(G-z)$. Thus, applying Lemma 5.3.2(a) twice, we obtain $A^*(G-z) \subseteq A^*(G-x)$ and $A^*(G-x) \subseteq A^*(G-z)$; that is, $A^*(G-x) = A^*(G-z)$. This proves Claim 6.

Claim 7. The pseudo-foundation of any pseudo-tower lies in $V(G_0)$.

To prove this, suppose $x \in V(G_0)$ and that $z \in A^*(G-x)$. Furthermore, suppose, by way of contradiction, that for some y, $z \in C(G-y)$. Then $A^*(G-z) \supseteq A^*(G-y)$ by Lemma 5.3.2(b), and so $y \in A^*(G-z)$. However, $A^*(G-z) = A^*(G-x)$ by Claim 6, so $y \in A^*(G-x)$. But now, by Lemma 5.3.2(a), $D(G-y) \subseteq D(G-x)$. Moreover, $x \notin D(G-x)$, so $x \notin D(G-y)$. Hence $x \in A^*(G-y) \cup C(G-y)$. But $x \in V(G_0)$ and, therefore, $x \notin C(G-y)$. So $x \in A^*(G-y)$. Now applying Lemma 5.3.2(a) yet again, we have $A^*(G-x) \subseteq A^*(G-y)$ and $D(G-x) \subseteq D(G-y)$. Thus $C(G-y) \cap (A^*(G-x) \cup D(G-x)) = \emptyset$ and hence $C(G-y) \subseteq C(G-x)$. But then we have $z \in C(G-x)$, a contradiction, and Claim 7 is proved.

At this point we may conclude that any two pseudo-towers are point-disjoint. This follows from the observation that a point in a pseudo-tower determines its pseudo-foundation as the set of those points of $V(G_0)$ which are adjacent to it.

Since $\operatorname{def}(G) = 0$ and G is not elementary, we know that $C(G-x) \neq \emptyset$ for some x by Theorem 5.1.3. We prove even more, however.

Claim 8. Every point y of $V(G) - V(G_0)$ lies within some pseudo-tower.

Let $z \in V(G)$ be such that $y \in C(G-z)$ and suppose $C(G-z)$ is inclusion-wise maximal with respect to all sets $C(G-x)$, $x \in V(G)$. That is, no $C(G-x)$ properly contains $C(G-z)$. We will show that $C(G-z)$ is a pseudo-tower; that is, $z \in V(G_0)$.

Suppose $z \notin V(G_0)$. Thus $z \in C(G-u)$ for some $u \in V(G)$. But then by Lemma 5.3.2(b), $A^*(G-z) \supseteq A^*(G-u)$ and $D(G-z) \supseteq D(G-u)$. But since $(D(G-z), A^*(G-z), C(G-z))$ and $(D(G-u), A^*(G-u), C(G-u))$ are both partitions of $V(G)$, it follows that $C(G-z) \subseteq C(G-u)$. But $z \in C(G-u) - C(G-z)$ and hence $C(G-z) \subset C(G-u)$, contradicting the maximality of $C(G-z)$. This proves Claim 8.

Note now that Claim 8, together with the fact that pseudo-towers are mutually disjoint, guarantees that the pseudo-towers are precisely the components of $G - V(G_0)$.

To complete the proof of (c), we must show that G_0 is a saturated elementary graph and that the pseudo-foundations of the pseudo-towers are precisely the classes of the partition $\mathcal{P}(G_0)$.

Since all lines between $V(G_0)$ and $V(G)-V(G_0)$ join a pseudo-tower to its pseudo-foundation, (i.e., a $C(G-x)$ to an $A^*(G-x)$), all such lines are forbidden by the Gallai-Edmonds Theorem. Hence since G has a perfect matching, so has G_0.

Claim 9. G_0 is saturated.

Choose any $x, y \in V(G_0)$ and suppose $xy \notin E(G_0)$. Since G is saturated, let F' be a perfect matching of $G + xy$ containing xy. Then $F = F' - xy$ is a perfect matching of $G - x - y$. Since all pseudo-towers of G have perfect matchings, such pseudo-towers are, in particular, even. So each pseudo-tower must have an even number of lines of F joining it to its pseudo-foundation in $G_0 - x - y$. But then $F_0 = F \cap E(G_0)$ can be extended to a perfect matching of $G_0 - x - y$, for we can match pairs of points in each pseudo-foundation, since said pseudo-foundations induce complete subgraphs in G_0. So $G_0 + xy$ has a perfect matching containing xy; that is, G_0 is saturated and Claim 9 is verified.

Note that, conversely, it is trivial that if $x, y \in V(G_0)$ and $G - x - y$ has no perfect matching, then neither does $G_0 - x - y$. Thus a line of G_0 is forbidden in G if and only if it is forbidden in G_0.

Claim 10. G_0 is elementary and the classes of $\mathcal{P}(G_0)$ are precisely the pseudo-foundations $A^*(G - x)$, for all $x \in V(G_0)$.

By the remark just before Claim 9, G_0 has a perfect matching, so by Theorem 5.1.3, it will suffice to show that $C(G_0 - x) = \emptyset$ for all $x \in V(G_0)$.

By the remark just prior to the statement of Claim 10 and by Lemma 5.3.1(a), we know that $D(G_0 - x) = D(G - x) \cap V(G_0)$. So $A^*(G_0 - x) \cup C(G_0 - x) \subseteq A^*(G - x) \subseteq V(G_0)$. But since each point y of $A^*(G - x)$ is adjacent to a point z of $D(G - x)$ via an allowed line, we must have $z \in D(G_0 - x)$ and hence each point y' of $A^*(G_0 - x) \cup C(G_0 - x)$ is adjacent to such a z. But then since no point of $C(G_0 - x)$ is adjacent to a point of $D(G_0 - x)$, we must have $C(G_0 - x) = \emptyset$. This completes the proof of Claim 10 and the theorem. ■

Let v be a point of a graph G with a perfect matching. Call v **totally covered** if every line incident with v is allowed. We can appeal to the Cathedral Construction to prove the following result.

5.3.9. THEOREM. *If G is a k-connected graph possessing a perfect matching and $k \geq 2$, then G has at least k totally covered points.*

PROOF. We may assume G is saturated for if not, saturate it, realizing that the set of allowed lines remains the same during the saturation process. Hence the saturated graph arising from G has no new totally covered points. Let G_0 be as defined in Theorem 5.3.8 and suppose $\{x\}$ is a singleton class in $\mathcal{P}(G_0)$. Then tower $G_{\{x\}}$ must be empty or else x would be a cutpoint of G. But now every line incident with x in G_0 is allowed in G_0, since $\{x\}$ is a class of $\mathcal{P}(G)$, and hence is also allowed in G; that is, x is totally covered in G.

Now if every class of $\mathcal{P}(G_0)$ is a singleton, $G = G_0$, and since a k-connected graph must have at least $k + 1$ points, the conclusion of the theorem follows immediately. So suppose S is a class of $\mathcal{P}(G_0)$ and that S has at least two members. Then, by Theorem 5.2.2(d), S is a cutset of G_0, and hence of G as well by Theorem 5.3.8. Thus $|S| \geq k$ by the hypothesis. But now by Exercise 5.3.5, there must be at least $|S|$ one-element classes in $\mathcal{P}(G_0)$ and the proof is complete. ■

Actually, we have proved a more general result. Namely, if G is a k-connected saturated graph, then G contains at least k points v such that $G - v$ is factor-critical.

The considerations of the present section will be used later in our discussion about the number of different perfect matchings a k-connected graph may have. (See Section 8.6.)

We can at this point, however, use our results on the Cathedral structure of saturated graphs to help us ascertain the structure of all graphs having *exactly one* perfect matching. Recall that this problem was investigated for bipartite graphs in Section 4.3. The next result is due to Kotzig (1959b).

5.3.10. THEOREM. *If a connected graph G has exactly one perfect matching, then it has a cutline belonging to the perfect matching.*

PROOF. As before, it is enough to prove the result for G saturated. If $G = K_2$, we are done. Otherwise, we may suppose G is not elementary. (For if it were elementary, it would have two adjacent allowed lines, a contradiction.) Thus G has a non-trivial Cathedral structure; that is, $G \neq G_0$ or equivalently, G contains at least one non-empty tower. Moreover, G_0 is elementary and it possesses only one perfect matching. Hence $G_0 = K_2$ and G_0 is a cutline of G. ■

5.3.11. EXERCISE. Prove Theorem 5.3.10 directly without appealing to the Cathedral structure.

5.3.12. COROLLARY. *A graph G has a unique perfect matching if and only if it can be constructed by iterating the following construction:*

Let G_1 and G_2 be two point-disjoint graphs, each with a unique perfect matching. (Either or both may be empty.) Let x_1 and x_2 be two new points. Join at least one point of G_i to x_i for $i = 1$ and 2 and join x_1 to x_2.

PROOF. First suppose G arises from G_1 and G_2 as described. Then x_1x_2 must be in every perfect matching of G and hence G has exactly one such matching.

Conversely, suppose G has a unique perfect matching. Then by Theorem 5.3.10, graph G has a cutline x_1x_2 belonging to this matching. Thus each component of $G - x_1 - x_2$ has a unique perfect matching. Let G_1 be any one of these components and G_2, the union of the rest. ■

The following two results bound the degree and the number of lines, respectively, in a graph with a unique perfect matching.

5.3.13. COROLLARY. *If graph G has p points and a unique perfect matching, then G contains a point of degree $\leq \lfloor \log_2(p+1) \rfloor$. For each even p, this bound is sharp.*

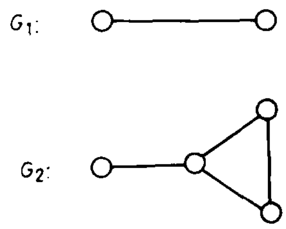


FIGURE 5.3.4.

PROOF. To see that the bound in Corollary 5.3.13 is best possible, the following family of graphs $\{G_k\}$, $k = 1, 2, 3, \ldots$, will suffice. Note that for each k, $|V(G_k)| = 2k$, $\text{mindeg}(G_k) = \lfloor \log_2(k+1) \rfloor$ and G_k has a unique perfect matching.

Let G_1 and G_2 be the graphs shown in Figure 5.3.4.

Now suppose $k \geq 3$ and $G_1, G_2, \ldots, G_{k-1}$ have been constructed. We construct G_k as follows.

First suppose $k-1$ is even, say $k-1 = 2s$. Form G_k by taking two copies of G_s, two new points x_1 and x_2, joining x_1 to all points of one G_s, x_2 to all points of the second G_s and then joining x_1 to x_2.

If $k-1$ is odd, say $k-1 = 2s+1$, form G_k by joining each point of a G_s to a new point x_1, each point of a G_{s+1} to a second new point x_2 and then join x_1 to x_2.

The existence of the bound claimed follows by induction on p and the details are left to the reader. ∎

5.3.14. COROLLARY. (*Hetyei, unpublished*). *If $f(n)$ denotes the maximum number of lines in a graph on $2n$ points having a unique perfect matching, then $f(n) = n^2$.*

PROOF. (See Lovász (1972a).) A set of extremal graphs $\{G_1, G_2, \ldots\}$ in this case may be constructed as follows. Note that for each n, G_n has $2n$ points, n^2 lines and a unique perfect matching. Let $G_1 = K_2$. Given G_{n-1} on $2n-2$ points, form G_n by adding two new points x and y, and then join x to all points of G_{n-1}, as well as to y.

It is easy to prove by induction that $f(n) \leq n^2$ and these details are also left to the reader. ■

5.4. Ear Structure 1-extendable Graphs

In Section 4.1 we saw that elementary bipartite graphs are always 1-extendable and that they can be constructed in a simple way by adding one "ear" at a time to smaller graphs of the same type. In the present section we shall see that we can construct general 1-extendable graphs in a similar way, but unfortunately we cannot always get by with adding a single ear at each step. To study ear structure in this more general (that is, non-bipartite) setting, we need to introduce several different versions of ear decompositions.

Let G be a graph and G' a subgraph of G. **An ear of G relative to G′** is any path in G having both endpoints — but no interior points — in G' or a cycle in G having exactly one point in G'. (The cycle, of course, may be viewed as the degenerate case of the path when the endpoints of the path coincide.) Ears having distinct endpoints are called **open**; otherwise, **closed**. In this section we shall deal only with open ears and hence we shall dispense with the word "open" here. (Closed ears will be encountered in Section 5.5.) Furthermore, all ears considered in this book will be of odd length. **An ear decomposition of G starting with G′** is a representation of G in the form $G = G' + P_1 + \cdots + P_k$, where P_1 is an ear of $G' + P_1$ relative to G' and P_i is an ear of $G' + P_1 + \cdots + P_i$ relative to $G' + P_1 + \cdots + P_{i-1}$ for $2 \leq i \leq k$.

Recall that a subgraph G' of any graph G is called nice if $G - V(G')$ has a perfect matching. We begin our study of ear decompositions with the following result, implicit in the work of Hetyei (1964).

5.4.1. THEOREM. *Let G be 1-extendable and G' a subgraph of G. Then G has an ear decomposition starting with G' if and only if G' is a nice subgraph of G.*

PROOF. Let G and G' be as given in the hypothesis and first suppose that G has an ear decomposition starting with G'. For each ear P_i, let $M(P_i)$ denote the set of lines of P_i situated at an odd distance from the endpoints of P_i (i.e., take the second, fourth, etc. lines of P_i starting at either end of P_i). Then $M(P_1) \cup \cdots \cup M(P_k)$ is a perfect matching of $G - V(G')$ and hence G' is nice.

Conversely, suppose G' is a nice subgraph. Let M be a perfect matching of $G - V(G')$. If G' spans G there is nothing to show, so suppose G' does not span G. G is connected so let e be any line joining $V(G')$ to $V(G) - V(G')$ and let F be a perfect matching of G containing e. If P_1 is the connected component of $M \cup F$ containing e, it must be a path beginning and ending on G'. So P_1 is an ear of G' and $G' + P_1$

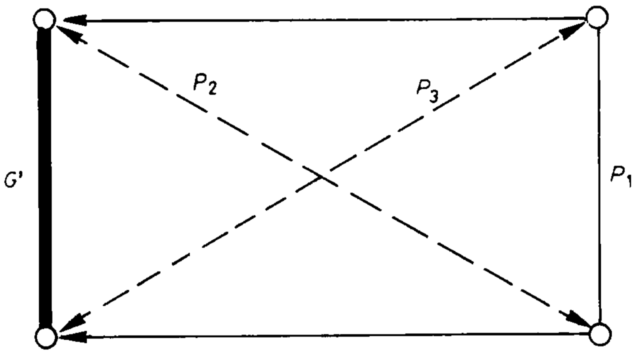


FIGURE 5.4.1.

is a nice subgraph of G. Continuing in this manner, we must eventually reach a first spanning subgraph of G, say $\hat{G}$. Then all remaining ears of $E(G) - E(\hat{G})$ are single lines. ∎

There is a subtle, but important, point to be made here. In the preceding proof we saw that we can get an ear decomposition of a 1-extendable graph G, starting with any nice subgraph and adding one ear at a time. But it is not necessarily true that each of the intermediate graphs formed between G' and G is also *itself* 1-extendable. This does hold for *bipartite* graphs as we saw in Chapter 4. But consider the simple non-bipartite example in Figure 5.4.1. Note that although G' and $G'+P_1$ are 1-extendable, neither $G'+P_1+P_2$ nor $G'+P_1+P_3$ is 1-extendable. In other words, to maintain 1-extendability at each stage, we may sometimes have to add more than one ear in a single step. We will see, however, as one of the main results of this section (the so-called "Two Ear Theorem", Theorem 5.4.6), that we can retain 1-extendability at each step by adding no more than *two* ears at a time. We start with several new definitions.

Again, let G be any graph and G' any subgraph of G. **An ear system of G relative to G$'$** is a set of point-disjoint paths of G of odd length each of which is openly disjoint from G', but has both endpoints in G'. Then a sequence of subgraphs of G, $(G_0, G_1, \ldots, G_m)$ is **a graded ear decomposition of G starting with G$_0$** if $G_m = G$, every G_i for $i = 0, \ldots, m$

is a nice 1-extendable subgraph of G and for each i, G_{i+1} is obtained from G_i by attaching an ear system relative to G_i. Note, then, that only 1-extendable graphs can have graded ear decompositions. Integer $m+1$ is said to be the **length** of the decomposition.

Sometimes the following notation for a graded ear decomposition is instructive:

$$G = G_0 + (P_1 + \cdots + P_{i_1}) + (P_{i_1+1} + \cdots + P_{i_2}) + \cdots + (P_{i_{r-1}+1} + \cdots + P_{i_r})$$

where $\{P_{i_{k-1}+1}, \ldots, P_{i_k}\}$ is the ear system to be added to G_{k-1} to obtain G_k. We call this ear system the **(k+1)st grade** of the decomposition of G. (G_0 is then the **first grade**.)

5.4.2. THEOREM. *Every* 1*-extendable graph* G *has a graded ear decomposition starting with any nice* 1*-extendable subgraph* G_0.

PROOF. Let G and G_0 be as given and suppose $G_i \neq G$ has been constructed. As in the proof of Theorem 5.4.1, choose any line e joining $V(G_i)$ to $V(G) - V(G_i)$, let M be any perfect matching of $G - G_i$ and let F be a perfect matching of G containing e. As before, $M \cup F$ consists of disjoint paths, alternating cycles and duplicated lines. Moreover, there is at least one path in $M \cup F$, namely that containing e. This time add *all* such paths to G_i to obtain G_{i+1}. The rest follows as before. ■

Let us pause for a few remarks about graded ear decompositions. Let G' be a nice 1-extendable subgraph of a 1-extendable graph G.

(1) There is a graded ear decomposition of G in which G' occurs starting with any line of G'. This is self-evident for G' has itself a graded ear decomposition which we may simply continue to get G. In particular, every 1-extendable graph has a graded ear decomposition starting with any given line.

(2) Furthermore, if F is *any* perfect matching of $G - V(G')$, then all ears after G' can be chosen to be alternating with respect to F.

(3) If $(G_0, \ldots, G_m)$ is any graded ear decomposition of G where G_0 is a single line, then there is exactly one perfect matching F in G such that $F \cap E(G_i)$ is a perfect matching of G_i for every i, $0 \leq i \leq m$. Clearly, all ears occurring in this decomposition alternate with respect to F. It is reasonable, therefore, to call F the perfect matching *associated with* the given ear decomposition.

(4) Just as in the case of bipartite graphs, the number of ears (but not grades!) in any ear-decomposition of graph G, starting with a single line, is always $|E(G)| - |V(G)| + 2$ (where the starting line is counted as the first ear.)

5.4.3. EXERCISE. Prove that every 1-extendable graph on $p \geq 6$ points contains a nice cycle of length ≥ 6.

Theorem 5.4.2 can be applied to give a simple proof of the following result due to Little (1974b).

5.4.4. THEOREM. *If G is 1-extendable and e_1 and e_2 are any two lines of G, then e_1 and e_2 lie on a nice cycle.*

PROOF. We proceed by induction on $|E(G)|$. First suppose both endpoints of e_2 are of degree 2. Then replace the path of length 3, having e_2 as its middle line, by a single line e_3. The new graph $\hat{G}$ resulting from this operation is 1-extendable and has $|E(\hat{G})| < |E(G)|$. So by the induction hypothesis applied to $\hat{G}$, there is a nice cycle in $\hat{G}$ containing e_1 and e_3 and this cycle clearly extends to a nice cycle in G.

So assume e_2 has an endpoint of degree at least 3. Let F_2 be a perfect matching in G containing e_2. We claim there is a cycle C containing e_1 and alternating with respect to F_2. If $e_1 \in F_2$, let e_3 be any line adjacent to e_1 and F_3 a perfect matching containing e_3. Then let C be the F_2-F_3 alternating cycle containing e_1. If $e_1 \notin F_2$, let F_1 be any perfect matching containing e_1 and let C be the $F_1 - F_2$ alternating cycle containing e_1. This proves the claim.

Now by Remark (1) above, we have a graded ear decomposition of G, $\{e_1\} = G_0 \subsetneq G_1 = C \subsetneq G_2 \subsetneq \cdots \subsetneq G_m = G$. Moreover, by Remark (2), we may choose this ear decomposition using C so that every ear alternates with respect to F_2. Now let G_i be the first graph in the sequence which includes e_2. Then $G_i \neq G$, for e_2 — being in F_2 — has both endpoints of degree 2 in G_i. But then $|E(G_i)| < |E(G)|$ and so by the induction hypothesis applied to G_i, there is an alternating cycle in G_i (and hence in G) containing e_1 and e_2. ■

A graded ear decomposition $(G_0, G_1, \ldots, G_m)$ of G is **longest** if m is maximum among all graded ear decompositions of G. Unfortunately, we do not know how to efficiently find a longest decomposition nor even to check if a given decomposition is longest! So we shall resort to some other types of decompositions which are related to this one, but are more tractable.

A graded ear decomposition $(G_0, \ldots, G_m)$ is **non-refinable** if given any $G_{i+1} = G_i + P_1 + \cdots + P_k$ in the sequence, the proper subgraph $G_i + P_{i_1} + \cdots + P_{i_r}$ is not 1-extendable for any r, $0 < r < k$. That is, there is no proper refinement of the grouping of the ears where each step remains 1-extendable. Figure 5.4.2 shows the essentially unique non-refinable ear-decomposition of the Petersen graph. Note that in the

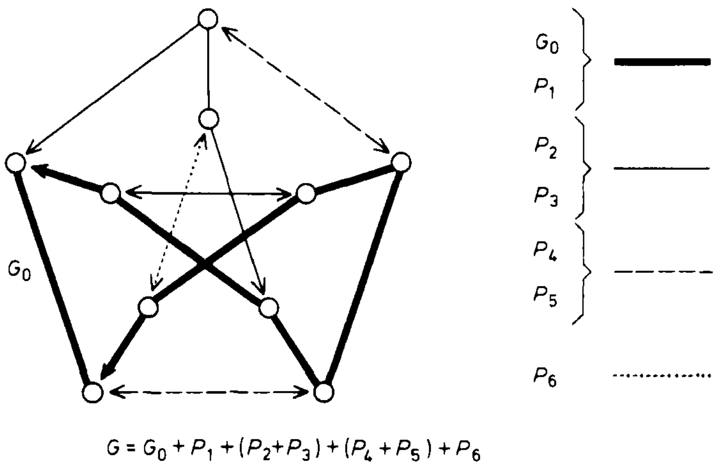


FIGURE 5.4.2. An ear decomposition of Petersen's graph

decomposition, stages G_1 and G_4 are one-ear additions, wheres G_2 and G_3 consist of two-ear additions. (The length of the decomposition is $m+1=5$.)

Of course longest decompositions must be non-refinable, but the converse need not hold. Just consider the two non-refinable ear decompositions of the graph in Figure 5.4.3, listed below, which have different lengths and only the longer of the two turns out to be a longest decomposition. More specifically, it is easy to see that $(G_0, G_1, G_2, G_3, G_4, G_5 = G)$ and $(H_0, H_1, H_2, H_3, H_4 = G)$ are non-refinable decompositions of different lengths where $G_0 = v_1v_2$, $G_1 = G_0 + v_2v_3v_4v_5v_6v_7v_8v_9v_{10}v_1$, $G_2 = G_1 + (v_3v_7 + v_6v_8)$, $G_3 = G_2 + v_4v_9$, $G_4 = G_3 + v_2v_8$ and $G_5 = G_4 + v_1v_5$, whereas $H_0 = v_2v_8$, $H_1 = H_0 + v_8v_6v_7v_3v_4v_5v_1v_2$, $H_2 = H_1 + (v_7v_8 + v_5v_6)$, $H_3 = H_2 + (v_1v_{10}v_9v_8 + v_2v_3)$ and $G = H_4 = H_3 + v_4v_9$. This example is due to Naddef and Pulleyblank (1982).

But why on earth should we be interested in finding longest ear decompositions? For one thing, the longer the graded ear decomposition, the fewer ears attached at each grade and sometimes, therefore, the easier it is to follow what happens to certain properties and parameters as the graph is built up. A more concrete reason is that the number of grades (including G_0) is always a lower bound on the number of perfect matchings (since clearly each new grade introduces at least one new perfect matching). In fact, this bound is sharp! Consider the graph G in Figure 5.4.4.

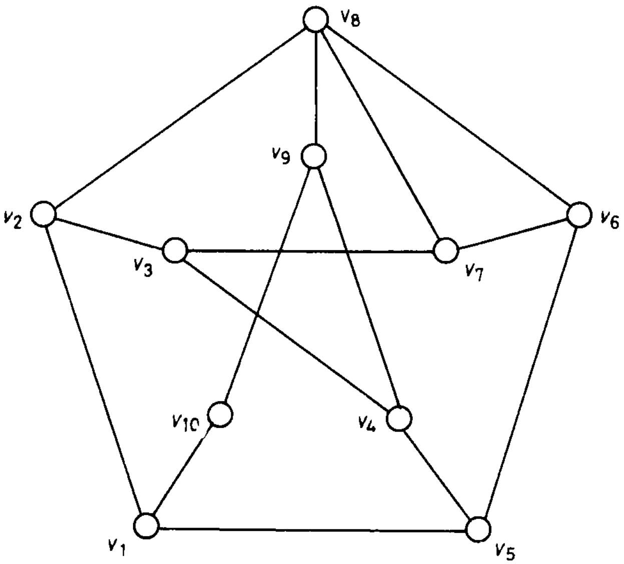


FIGURE 5.4.3.

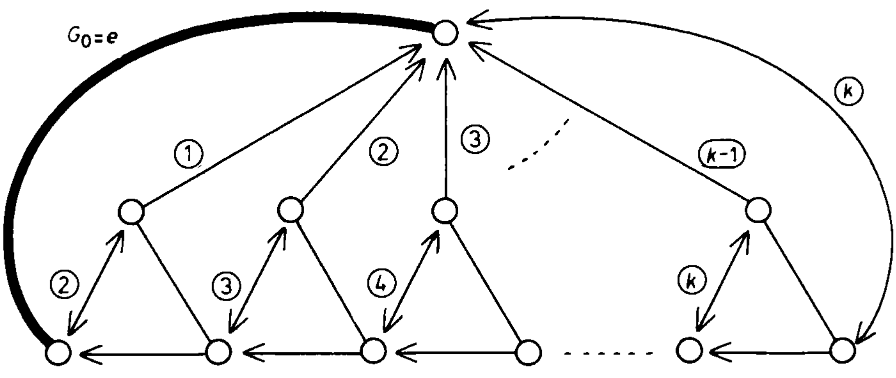


FIGURE 5.4.4.

If p denotes the number of points in G, the graded ear decomposition shown is non-refinable and has $k+1 = p/2+1$ grades. Could there possibly be another ear decomposition with more than $p/2$ grades? It is easy to see that the number of perfect matchings in this graph is $p/2+1$, so the answer is no! Note that this decomposition has two ears in each grade except the first two. In view of Theorem 5.4.6 below, it follows that in *any* non-refinable ear decomposition of this graph, all grades but the first two consist of exactly two ears.

Ear decompositions with varying special properties will prove useful to us in different situations to follow. For instance, in one case we shall add only one ear whenever possible; in another case we shall try to accomplish all 2-ear additions as soon as possible; in a third case we shall add an ear which brings in a new point whenever possible. Such selection rules turn out to guarantee various nice properties of the resulting ear decompositions.

At this point let us make a comment which will prove useful subsequently. In some applications of ear decompositions, we will have a graded ear decomposition $(G_0,\ldots,G_m)$ of some 1-extendable graph G and $G_{k+1}=G_k+P_1+\cdots+P_s$ at the k^{th} step of this decomposition. We shall want to decide if the addition of some subset $\{P_{i_1},\ldots,P_{i_r}\}$ of this ear system would also give a 1-extendable graph. Define a new graph $G'_{k+1} = G_k+e_1+\cdots+e_s$, where for each i, $1\le i\le s$, e_i is a new line connecting the endpoints of P_i. (Recall that the P_i's are openly disjoint.) Then clearly, as mentioned in Section 5.1, G'_{k+1} is 1-extendable and $G_k+P_{i_1}+\cdots+P_{i_r}$ is 1-extendable if and only if $G_k+e_{i_1}+\cdots+e_{i_r}$ is 1-extendable. So if there is a perfect matching F of G'_{k+1} such that $F\cap\{e_1,\ldots,e_s\}=\{e_{i_1},\ldots,e_{i_r}\}$, then $G_k+P_{i_1}+\cdots+P_{i_r}$ is 1-extendable.

Recall that $\Phi(G)$ denotes the number of perfect matchings in G.

5.4.5. LEMMA. *Let G be elementary and let $e_1,\ldots,e_s$ be lines not in $E(G)$, but having both endpoints in $V(G)$. Suppose $\Phi(G+e_1+\cdots+e_s)>\Phi(G)$. Then there exist i and j, $1\le i<j\le k$ such that $\Phi(G+e_i+e_j)>\Phi(G)$.*

PROOF. It suffices to consider the special case when $s=3$. If $\Phi(G+e_i)>\Phi(G)$ for some i, $1\le i\le 3$ we are done, so suppose $\Phi(G+e_i)=\Phi(G)$ for $i=1,2$ and 3. Then by Theorem 5.2.2(b), both endpoints of e_i belong to the same class of $\mathcal{P}(G)$, for $i=1,2$ and 3.

We claim, moreover, that no class $S\in\mathcal{P}(G)$ spans all of e_1, e_2 and e_3. For suppose to the contrary that all three lines have their endpoints in one class $S\in\mathcal{P}(G)$. By Theorem 5.2.2(d), $G-S$ consists of $|S|$ critical components. But $G-S=(G+e_1+e_2+e_3)-S$ and hence again by Theorem 5.2.2(d), $S\in\mathcal{P}(G+e_1+e_2+e_3)$. Thus e_1, e_2 and e_3 are forbidden in

$G+e_1+e_2+e_3$ and hence $\Phi(G+e_1+e_2+e_3)=\Phi(G)$, a contradiction. So our claim is verified.

Thus we can find a class $S \in \mathcal{P}(G)$ which spans exactly one of e_1, e_2 and e_3, say $e_3 = xy \in G[S]$. If $\Phi(G+e_1+e_3) > \Phi(G)$ we are done, so suppose $\Phi(G+e_1+e_3) = \Phi(G)$. Thus x and y belong to the same class $S_1 \in \mathcal{P}(G+e_1)$. Similarly, x and y belong to the same class $S_2 \in \mathcal{P}(G+e_2)$.

Now $(G+e_1)-S_1$ has $|S_1|$ critical components and hence $G-S_1$ has at least $|S_1|$ odd components. So $\operatorname{def}(G-S_1) \geq |S_1|$. On the other hand, we always have $\operatorname{def}(G-S_1) \leq \operatorname{def}(G)+|S_1| = |S_1|$, so $\operatorname{def}(G-S_1) = |S_1|$ and hence by Theorem 5.2.2(a), S_1 lies in a class of $\mathcal{P}(G)$. Since $\{x,y\} \subseteq S_1 \cap S$, it follows that $S_1 \subseteq S$, since maximal barriers are disjoint in an elementary graph. Moreover, since $G-S_1$ has at least $|S_1|$ odd components, it must have exactly $|S_1|$ odd components since G has a perfect matching, and so by definition, S_1 is a barrier in G.

Similarly $S_2 \subseteq S$ and S_2 is a barrier in G.

Next we claim that no line joins S_1-S_2 to S_2-S_1 in G. For suppose such a line e exists. Then e is forbidden in G since it joins two points of S, but it is allowed in $G+e_1$ since it joins a point of $S_1 \in \mathcal{P}(G+e_1)$ to a point not in S_1. Thus $\Phi(G+e_1) > \Phi(G)$, a contradiction.

But now we may apply Lemma 5.1.4 to conclude that $S_1 \cap S_2$ is a barrier in G.

Now consider again lines e_1 and e_2. By the choice of S, e_1 and e_2 have no points in S. Since S_1 is a barrier in G, we have, by Theorem 5.1.6, that $c(G-S_1) = c_0(G-S_1) = |S_1|$. On the other hand, since $S_1 \in \mathcal{P}(G+e_1)$ we have, by Theorem 5.2.2(d), that $c(G+e_1-S_1) = c_0(G+e_1-S_1) = |S_1|$. Thus e_1 must join two points of the same component of $G-S_1$. Since $S_1 \cap S_2 \subseteq S_1$, it follows that e_1 joins two points in the same component of $G-(S_1 \cap S_2)$.

Similarly, e_2 joins two points in the same component of $G-(S_1 \cap S_2)$. Since $S_1 \cap S_2$ is a barrier in G we have, by Theorem 5.1.6, that $G-(S_1 \cap S_2)$ has $|S_1 \cap S_2|$ components, all of which are odd. But $(G+e_1+e_2)-(S_1 \cap S_2)$ has precisely these same components. So $S_1 \cap S_2$ is also a barrier in $G+e_1+e_2$.

But x and y lie in $S_1 \cap S_2$ and hence must lie in the same *maximal* barrier of $G+e_1+e_2$, that is, in the same class of $\mathcal{P}(G+e_1+e_2)$. Thus by Remark 3(b) following Theorem 5.2.2, line $e_3 = xy$ is forbidden in graph $(G+e_1+e_2)+e_3$. Hence $\Phi(G+e_1+e_2) = \Phi(G+e_1+e_2+e_3) > \Phi(G)$ and the case when $s=3$ is finished. ∎

Now our target result follows easily.

5.4.6. THEOREM. (*The Two Ear Theorem*). *If* $(G_0, G_1, \ldots, G_m)$ *is a non-refinable graded ear decomposition of a* 1*-extendable graph* G*, then for each* i*,* $0 \le i < m$*,* G_{i+1} *arises from* G_i *by adding at most two ears.*

PROOF. If $G_{i+1} = G_i + P_1 + \cdots + P_s$ and $G'_{i+1} = G_i + e_1 + \cdots + e_s$, where e_j is a new line joining the endpoints of P_j, it is then enough to

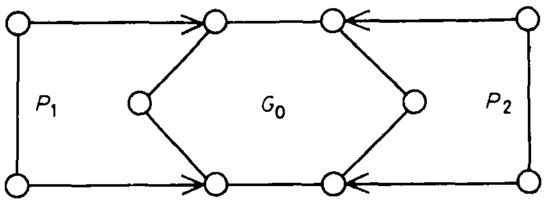


FIGURE 5.4.5. A non-elementary graph

show, by the comment preceding Lemma 5.4.5, that G'_{i+1} has a perfect matching containing at most two of the lines $e_1, \ldots, e_s$. But this follows from Lemma 5.4.5. ■

So Theorem 5.4.2, together with the preceding result, tells us that every 1-extendable graph has a non-refinable graded ear decomposition starting with an arbitrary line (or equivalently with an even cycle), in which each grade G_{i+1} results from the previous G_i through the addition of one or two ears. In the case of a 1-ear addition, the ear must join two different classes of $\mathcal{P}(G_i)$, whereas when a 2-ear addition is employed, one ear joins two points of one class and the other joins two points of a different class of $\mathcal{P}(G_i)$.

On the other hand, an arbitrary 2-ear addition to a 1-extendable graph G does not always result in a new 1-extendable graph $G + P_1 + P_2$. In fact, $G + P_1 + P_2$ may not even be elementary! (See Figure 5.4.5. Note that $G = G_0 + P_1 + P_2$, where G_0 is a 6-cycle and hence clearly a nice 1-extendable subgraph of G. But G is not elementary.)

We now know, on the basis of the Two Ear Theorem, that in constructing 1-extendable graphs, one- or two-ear steps will *suffice* to guarantee that the intermediate graphs in the construction will themselves be 1-extendable. But when is a two-ear step *necessary*? The next lemma will help in obtaining our answer to this question as formulated in Theorem 5.4.9 below.

5.4.7. LEMMA. (*The Ear Selection Lemma*). *Suppose G is a* 1*- extendable graph. Suppose H is a nice subgraph of G, M is a perfect matching of $G-V(H)$ and $V(H)=V_1 \cup V_2$ is a partition of $V(H)$. Suppose also that there is a path in $G-E(H)$ joining V_1 and V_2. Then G contains an ear relative to H which alternates with respect to M and joins V_1 and V_2.*

PROOF. We proceed by induction on $|E(G)-E(H)|$. If there is a line $e \notin E(H)$ joining V_1 to V_2 we are done, so assume there is no such line. Let M be the given perfect matching in $G-V(H)$. Let f be any line joining V_1 to $V(G)-V_1-V_2$. (For example, f may be the first line of any path in $G-E(H)$ joining V_1 and V_2). Let F be a perfect matching in G containing line f and let P be the connected component of $F \cup M$ containing f. Then P is an ear of H with at least one endpoint in V_1 and containing at least one point not in $V(H)$.

If the other endpoint of P is in V_2 we are done, so suppose it is in V_1. Then $H'=H+P$ is a nice subgraph of G. Let $V_1'=V_1 \cup V(P)$. By the induction hypothesis, H' has an ear Q joining V_1' and V_2. But then Q, together with a suitable subpath of P, forms an ear of H joining V_1 and V_2. ■

In Figure 5.4.6 we illustrate an example of a situation dealt with by the Ear Selection Lemma 5.4.7. Note that lines in perfect matching M of $G-V(H)$ are shown in bold face and the "selected ear" (which must be M-alternating) is indicated by a dotted path.

There is a bipartite analogue of the previous theorem which we present next. It will prove useful in the proof of Theorem 5.4.11 below.

5.4.8. LEMMA. (*The Bipartite Ear Selection Lemma*). *Assume $G=(U,W)$ is a* 1*-extendable bigraph, H is a nice subgraph of G, M a perfect matching of $G-V(H)$ and $V(H)=V_1 \cup V_2$ a partition of $V(H)$. Suppose also that $|V_1 \cap U| \geq |V_1 \cap W|$ and no line of H joins V_1 to V_2. Then G contains an ear relative to H which alternates with respect to M and joins $V_1 \cap U$ to $V_2 \cap W$.*

PROOF. This proof is by induction on $|V(G)-V_1-V_2|$. If $U \subseteq V_1$, then any line incident with V_2 forms an appropriate ear by itself. So suppose that $U \not\subseteq V_1$. Since G is elementary, it follows from Theorem 4.1.1 that $|\Gamma(V_1 \cap U)| > |V_1 \cap U| \geq |V_1 \cap W|$ and hence there must be a line e joining $V_1 \cap U$ to $W-V_1$.

Let F be a perfect matching of G containing line e. Then the connected component P of $F \cup M$ containing e is an odd path with one endpoint in $V_1 \cap U$. If the other endpoint of P lies in $V_2 \cap W$, we are finished. So suppose the other end of P lies in $V_1 \cap W$ and consider

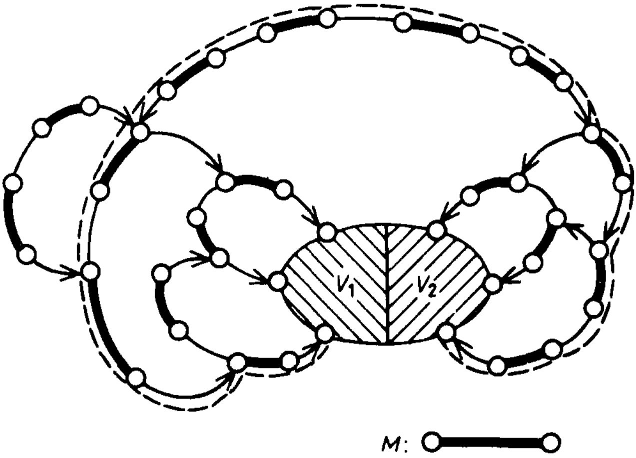


FIGURE 5.4.6.

subgraph $H' = H \cup P$ and the set $V_1' = V_1 \cup V(P)$. Note that H' is again nice, since $M' = M - E(P)$ is a perfect matching in $G - V(H')$. Since, moreover, $|V_1' \cap U| \geq |V_1' \cap W|$, we may apply the induction hypothesis to conclude that there is an M'-alternating path Q joining $V_1' \cap U$ to $V_2 \cap W$. But then Q, together with an appropriately chosen subpath of P, if necessary, forms an M-alternating path joining $V_1 \cap U$ to $V_2 \cap W$.■

Now we proceed to describe when two ears are needed in a single step in the non-bipartite case.

Let G be a graph and H an elementary subgraph of G. Call H a **splitting subgraph** of G if no connected component of $G - E(H)$ has points in more than one class of $\mathcal{P}(H)$; that is, no path of $G - E(H)$ joins two classes of $\mathcal{P}(H)$. The following theorem tells us that in an ear decomposition of a 1-extendable graph G, we are forced to add two ears to the subgraph H at hand only if H is splitting in G.

5.4.9. THEOREM. *Let G be a 1-extendable graph and H a nice 1-extendable subgraph of G. Then there is a single ear P in G relative to H, such that $H + P$ is 1-extendable if and only if H is non-splitting.*

PROOF. Suppose first that H is splitting and P is a single ear relative to H. Then by definition of splitting, P joins points in the same class of $\mathcal{P}(G)$ and then by Theorem 5.2.2(b) and by Remark 1 following it, $H + P$ is not 1-extendable.

Conversely, suppose H is non-splitting; that is, there is a path in $G - E(H)$ joining two different classes of $\mathcal{P}(H)$. Let V_1 be one of these classes and set $V_2 = V(H) - V_1$. Since H is nice, theie is a perfect matching M of $G - V(H)$ and, by the Ear Selection Lemma 5.4.7, there must be an ear P relative to H joining V_1 and V_2 which is M-alternating. Since H is 1-extendable, clearly the second, fourth, ... lines of P are allowed in $H + P$. On the other hand, since P joins different classes of $\mathcal{P}(H)$, $H + P$ must have a perfect matching, by Theorem 5.2.2(b), which uses the first, third, ... lines of P. It follows that $H+P$ is 1-extendable.∎

So now we know, in terms of the concept of a splitting subgraph, how long we may postpone the first two-ear step in the construction of a general 1-extendable graph.

It is a different matter to ask how *soon* a 2-ear addition can be made. We can answer this question for the first 2-ear step and this answer turns out to have some non-trivial applications.

In particular, suppose we have a non-refinable ear decomposition of an arbitrary 1-extendable graph starting with a line. This starts with some (at least one) 1-ear addition and all such 1-ear steps keep the resulting graph bipartite. The first 2-ear addition, however, results in a non-bipartite graph. So finding a nice 1-extendable non-bipartite subgraph with minimum cyclomatic number is equivalent to determining how *soon* a 2-ear addition step can occur in a non-refinable ear decomposition. The following theorem provides an answer.

5.4.10. THEOREM. *Every non-bipartite 1-extendable graph G has a non-refinable graded ear decomposition $(G_0, G_1, G_2, G_3, \ldots, G_m = G)$ (where G_0 is a single line and G_1, an even cycle) such that either G_2 or G_3 is a 2-ear addition.*

In Figure 5.4.7 we see K_4, which has a decomposition in which the final grade G_2 results from G_1 through the addition of two single lines as ears. On the other hand, the second graph R_3, the triangular prism, has no decomposition in which G_2 results from a 2-ear addition, but it does have one in which G_3 arises from G_2 via adding two ears.

The two graphs K_4 and R_3 of Figure 5.4.7 are, in a very real sense, the archetypal instances of the two possible types of behavior described in Theorem 5.4.10. We make this precise in the next result, first proved in Lovász (1982c) and from which Theorem 5.4.10 follows easily.

5.4.11. THEOREM. *Let G be a non-bipartite 1-extendable graph all of whose nice 1-extendable proper subgraphs are bipartite. Then G is isomorphic to an even subdivision of K_4 or of the triangular prism R_3.*

PROOF. The proof proceeds by induction on $|V(G)|$. Let us first show that it suffices to consider only 3-line-connected graphs.

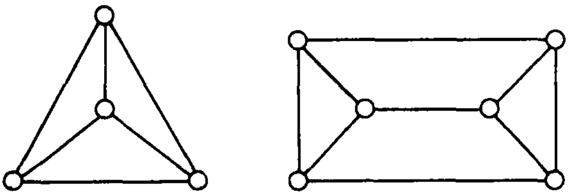


FIGURE 5.4.7.

Suppose G contains a point x of degree 2 with incident lines xy and xz. Shrink path yxz to a single point. The resulting graph G' is trivially non-bipartite and 1-extendable and hence contains a subgraph H' with these same properties, but having only bipartite nice 1-extendable proper subgraphs. So by the induction hypothesis applied to H', we see that G' contains an even subdivision G_1 of K_4 or R_3 as a nice subgraph. Reinserting path yxz, we see that G itself must contain an even subdivision of K_4 or of R_3 as a nice (non-bipartite) subgraph. But then by the hypothesis of this theorem, G itself must be an even subdivision of K_4 or R_3. So we may as well assume G has no points of degree 2.

Now suppose G has a pair $\{e, f\}$ of lines such that $G-e-f$ is disconnected. Since G is 2-connected (because every elementary graph is), it is clear that $G-e-f$ has precisely two components G_1 and G_2 and e and f each join points in different G_i's. Let $e = x_1x_2$ and $f = y_1y_2$ where x_1, $y_1 \in V(G_1)$ and x_2, $y_2 \in V(G_2)$.

Case 1. Suppose G_1 and G_2 are both even. Then, since G is non-bipartite, one of $G_1 + x_1y_1$ and $G_2 + x_2y_2$ is also non-bipartite, say $G_1+x_1y_1$. It is easy to see that $G_1+x_1y_1$ is 1-extendable and hence must contain an even subdivision $\hat{G}_1$ of K_4 or R_3 as a nice subgraph. If $x_1y_1 \notin$

$E(\hat{G}_1)$ we are done, so suppose $x_1y_1 \in E(\hat{G}_1)$. Then, since $G_2 + x_2y_2$ is 1-extendable and $\neq K_2$ (for G has no points of degree 2), it contains a nice cycle C through x_2y_2. But then $\hat{G}_1 - x_1y_1 + x_1x_2 + y_1y_2 + (E(C) - x_2y_2)$ is again an even subdivision of K_4 or R_3 contained in G.

Case 2. Suppose G_1 and G_2 are both odd. Then let G_1' and G_2' denote the graphs obtained from G by contracting G_2 and G_1, respectively, to a point. Clearly both G_1' and G_2' are 1-extendable and at least one of them is non-bipartite, say, G_1'. The argument is then analogous to that in Case 1.

So we may now assume G is 3-line-connected. Consider any non-refinable decomposition $(G_0, \ldots, G', G)$. The penultimate graph in this chain, G', must be bipartite or else we are finished by induction. Thus G, being non-bipartite, must arise from G by attaching two ears. But G has no point of degree two, so these two ears are single lines. Let us denote them by e_1 and e_2. Then it is clear that e_1 joins two points of the same class of the bipartition of G' and e_2, two points of the other class. Moreover, every nice cycle (in fact, every *even* cycle) in G containing one of them also must contain the other.

Let C be a nice cycle using e_1 (and hence e_2) and consider a non-refinable decomposition of G starting with C. As before, it follows that G arises from a bipartite subgraph G'' by attaching two single-line ears f_1 and f_2, one in each color class of the bipartition of G.

By Little's Theorem 5.4.4, there is a nice cycle C' in G containing e_1 and f_1. Hence by our remarks above, C' also contains e_2 and f_2. Consider now a decomposition of G starting with C'. Let G''' be the penultimate graph of this decomposition to which lines h_1 and h_2 are added to yield G. Notice that the deletion of any of the three pairs $\{e_1, e_2\}$, $\{f_1, f_2\}$ or $\{h_1, h_2\}$ from G leaves a bipartite graph.

Applying Theorem 5.4.4 again, let C_0 be a nice cycle in $G' = G - e_1 - e_2$ through both endpoints of e_1. The line e_1, together with the two paths of cycle C_0, forms two odd cycles C_1 and C_2 in $G - e_2$. Suppose $f_1 \notin E(C_0)$. Then one of C_1 and C_2 is odd and contains neither f_1 nor f_2, contradicting the fact that $G - f_1 - f_2$ is bipartite. So f_1 (and similarly f_2) are lines of cycle C_0. Let J_1 and J_2 be the two component paths of $C_0 - f_1 - f_2$. Then e_1 must join J_1 to J_2 or, again, C_0 and e_1 would form an odd cycle in $G - f_1 - f_2$. Moreover, J_1 and J_2 are both odd paths, since f_1 and f_2 are spanned by different color classes of the bipartite graph $G - f_1 - f_2$.

Now $J_1 \cup J_2$ is a nice subgraph of bigraph $G - f_1 - f_2$. Moreover, if the bipartition of $V(G - f_1 - f_2)$ is (U, W), we have $|V(J_1) \cap U| = |V(J_1) \cap W|$.

Interchanging the names of U and W if necessary, suppose all the points of J_1 lying at an odd distance on J_1 from the endpoint of e_1 are in U. Then by the Bipartite Ear Selection Lemma (Lemma 5.4.8), there is a path P in $G - f_1 - f_2$ joining J_1 to J_2, whose endpoint on J_1 is at an odd distance on J_1 from the endpoint of e_1 on J_1 and alternating with respect to some perfect matching M of $G - V(C_0)$. Thus P has odd length. Moreover, since $G - f_1 - f_2$ is bipartite, the endpoints of e_1 and P on J_2 are also an odd distance apart on J_2.

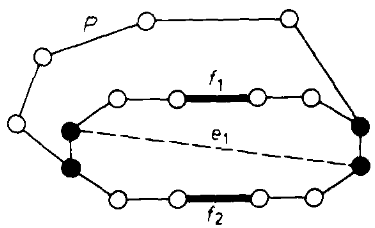


FIGURE 5.4.8.

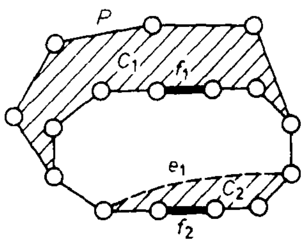


FIGURE 5.4.9.

If the endpoints of e_1 and P separate each other on cycle C_0, we have an even subdivision of K_4. (See Figure 5.4.8.). So suppose they do not. Then line e_1 and path P, together with appropriate paths of C_0, form two point-disjoint odd cycles C_1 and C_2. (See Figure 5.4.9.).

Furthermore, since $C_0 \cup P$ is a nice subgraph of G, it follows that $C_1 \cup C_2$ is a nice subgraph of G. Interchanging subscripts if necessary, we may assume $e_i \in E(C_i)$ for $i = 1, 2$, since $G - e_1 - e_2$ is bipartite. Similarly, we may assume f_i and $h_i \in E(C_i)$ for $i = 1, 2$.

Let E_i, F_i and H_i denote the component paths of $C_i - \{e_i, f_i, h_i\}$ so that E_i and e_i, F_i and f_i, as well as H_i and h_i, are disjoint. (See Figure 5.4.10.)

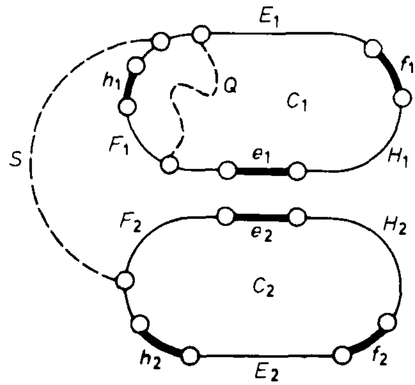


FIGURE 5.4.10.

Claim 1. No path in $G - \{e_1, e_2, f_1, f_2, h_1, h_2\}$ joins two points of the same C_i on two different paths among the three E_i, F_i and H_i.

Suppose, to the contrary, that there is a path Q joining E_1 to F_1, say. Then Q, together with one of the two pieces of C_1 joining its endpoints, forms an odd cycle C_1'. But then odd cycle C_1' misses at least one of the pairs $\{e_1, e_2\}$, $\{f_1, f_2\}$ or $\{h_1, h_2\}$, a contradiction.

Claim 2. If any path openly disjoint from $C_1 \cup C_2$ joins C_1 to C_2, then it joins E_1 to E_2, F_1 to F_2 or H_1 to H_2. (See Figure 5.4.10 once more.)

Again, suppose to the contrary that path S joins E_1 to F_2, say. Since G is 3-line-connected, $G - \{e_1, f_1\}$ contains a path T joining H_1 to $(C_1 \cup C_2 \cup S) - V(H_1)$ and openly disjoint from the two. The endpoint of T on $(C_1 \cup C_2 \cup S) - V(H_1)$ cannot lie on $S \cup F_2$ or on C_1, by Claim 1. So it must lie on $E_2 \cup H_2$. Let R_1 be the path of C_1 joining the endpoints of S and T which contains line f_1, but not e_1 nor h_1. Let R_2 be that path of C_2 joining endpoints of S and T such that cycle $\hat{C} = R_1 \cup S \cup R_2 \cup T$ is odd. Then $\hat{C}$ misses either both e_1 and e_2 or both h_1 and h_2. But either results in a contradiction.

Now since C_1 is odd, at least one of E_1, F_1, and H_1 is of even length, say E_1.

Claim 3. If M is a perfect matching of $G - V(C_1 \cup C_2)$, then there exists an M-alternating path P joining a point $x_1 \in V(E_1)$ to a point $x_2 \in V(E_2)$. Moreover, x_1 divides E_1 into two even paths.

To prove this, consider the 1-extendable bipartite graph $G - e_1 - e_2$ and the two subgraphs $G_1 = E_1$ and $G_2 = E_2 \cup F_1 \cup F_2 \cup H_1 \cup H_2$. Let us name the bipartition (U, W) of $G - e_1 - e_2$ in such a way that the midpoint of E_1 lies in U. Then of course $|V(G_1) \cap U| \geq |V(G_1) \cap W|$ and, by the Bipartite Ear Selection Lemma 5.4.8, there is an M-alternating (odd) path P joining a point x_1 in $V(G_1) \cap U = V(E_1) \cap U$ to a point x_2 in $V(G_2) \cap W = W \cap (V(E_2) \cup V(F_1) \cup V(F_2) \cup V(H_1) \cup V(H_2))$. By Claims 1 and 2, P must meet G_2 in $W \cap V(E_2) \subseteq E_2$ and the claim is proved.

Now either both F_1 and H_1 are of even length or both are of odd length. But in either case, with at most a slight modification of the preceding argument, we can find an M-alternating path Q joining a point $y_1 \in V(F_1)$ to a point $y_2 \in V(F_2)$, where y_1 is at an even distance on F_1 from h_1. Similarly, we can find an M-alternating path R joining a point $z_1 \in V(H_1)$ to a point $z_2 \in V(H_2)$, where z_1 is at an even distance from f_1 on H_1.

By Claim 1 the three paths P, Q and R are mutually point-disjoint. (See Figure 5.4.11.) Remembering that P, Q and R are odd and since $G - e_1 - e_2$ is bipartite, we see that x_2, y_2 and z_2 must divide C_2 into three odd paths. But then $C_1 \cup C_2 \cup P \cup Q \cup R$ must be an even subdivision of R_3 and the proof is complete. ■

As promised, the proof of Theorem 5.4.10 follows immediately.

PROOF (of Theorem 5.4.10). Let G be a non-bipartite 1-extendable graph and among all nice non-bipartite 1-extendable subgraphs of G, choose one, H, all of whose nice 1-extendable proper subgraphs are bipartite. Then by Theorem 5.4.11, H is an even subdivision of K_4 or R_3. Moreover, by Theorem 5.4.2, G has a graded ear decomposition starting with H, $(H, G_1, \ldots, G_m = G)$. So if $H = \hat{K}_4$ is an even subdivision of K_4, G has a *non-refinable* graded ear decomposition $(G_0, G_1, G_2 = \hat{K}_4, \ldots, G)$, where $G_2 = \hat{K}_4$ arises from even cycle G_1 by adding two ears. If $H = \hat{R}_3$ is an even subdivision of R_3, G admits a *non-refinable* graded ear decomposition $(G_0, G_1, G_2, G_3 = \hat{R}_3, \ldots, G)$, where G_1 is an even cycle, G_2 is obtained by adding a single ear to G_1 and $G_3 = \hat{R}_3$ arises from G_2 by adding two ears. ■

In the remainder of this section we apply ear decomposition techniques to study properties of *minimal* elementary graphs. An elementary graph is **minimal elementary** if $G - e$ is not elementary for each line $e \in E(G)$.

The following simple property of such graphs was observed by Hetyei (1964).

5.4.12. EXERCISE. Every minimal elementary graph contains at least two adjacent points of degree two.

We proceed now to derive an upper bound on the number of lines in a minimal elementary graph. The proof will not use the full strength of

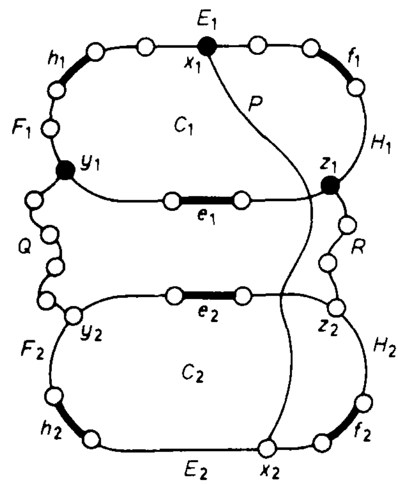


FIGURE 5.4.11.

the minimal elementary property and we shall formulate our result for a wider class of graphs.

An elementary graph G is **weakly minimal elementary** if for each line e in $E(G)$ either $G - e$ is not elementary or $\mathcal{P}(G - e) \neq \mathcal{P}(G)$. Minimal elementary graphs are clearly weakly minimal, but the converse is not true. Just consider $K_4 - e$ which is not even 1-extendable! An infinite class of 1-extendable counterexamples is also provided by the minimal bicritical graphs to be studied in Section 5.5. Such a graph G is weakly minimal elementary. However, G cannot contain any points of degree two, being bicritical. But then, in turn, it cannot be minimal elementary by Exercise 5.4.12.

5.4.13. LEMMA. *Let G be elementary, suppose $e_1, \ldots, e_k \notin E(G)$, but suppose $\bigcup_{i=1}^{k} V(e_i) \subseteq V(G)$. Let $G' = G \cup \{e_1, \ldots, e_k\}$ and suppose for each i, $|\mathcal{P}(G')| > |\mathcal{P}(G' - e_i)|$. Then $|\mathcal{P}(G)| \leq |\mathcal{P}(G')| - k$.*

PROOF. It is enough to show that there exists an i, $1 \leq i \leq k$, such that $|\mathcal{P}(G + e_i)| > |\mathcal{P}(G)|$, since the inequality in the hypothesis of this lemma then follows by induction.

We know that $|\mathcal{P}(G)| \leq |\mathcal{P}(G_i)|$ since $G_i = G + e_i$. (See Remark 2 following Theorem 5.2.2.) If $|\mathcal{P}(G)| < |\mathcal{P}(G_i)|$ we are finished, so suppose $|\mathcal{P}(G)| = |\mathcal{P}(G_i)|$ for each $i = 1, \ldots, k$. But $\mathcal{P}(G_i)$ is always a refinement of $\mathcal{P}(G)$ for each i (see Remark 3 after Theorem 5.2.2), so we must have $\mathcal{P}(G_i) = \mathcal{P}(G)$ for each i.

Let S be one of the classes of partition $\mathcal{P}(G)$. Then by Theorem 5.2.2(d), the graph $G - S$ has exactly $|S|$ components and all are factor-critical. But S is also a class of $\mathcal{P}(G_i)$ for each i. So $G_i - S$ has $|S|$ critical components. Hence e_i cannot join two components of $G - S$. Since this holds for each i, it follows that $G' - S$ has $|S|$ critical components. So S is a class of $\mathcal{P}(G')$, again by Theorem 5.2.2(d). Hence $\mathcal{P}(G) = \mathcal{P}(G')$ which is a contradiction, since $|\mathcal{P}(G)| \leq |\mathcal{P}(G' - e_1)| < |\mathcal{P}(G')|$. ■

An ear is called **improper** if it consists of a single line, and otherwise it is said to be **proper**. The preceding lemma now helps us prove the following result.

5.4.14. LEMMA. *Let G be a 1-extendable graph and suppose H is a nice 1-extendable, but non-spanning, subgraph of G. Let M be a perfect matching of $G - V(H)$. Then there exists an ear system $\{P_1, \ldots, P_m\}$ relative to H in G, such that $H' = H + P_1 + \cdots + P_m$ is 1-extendable, one or two of the P_i's are proper and at most $|\mathcal{P}(H')| - |\mathcal{P}(H)|$ of the P_i's are improper.*

PROOF. First we show that there is an ear system $\{P_1, \ldots, P_m\}$ such that $H' = H + P_1 + \cdots + P_m$ is nice, 1-extendable and at least one of the P_i's is proper. To see this, let e be any line with exactly one endpoint in H and let M_e be a perfect matching of G containing e. As we saw in the proof of Theorem 5.4.2, $M \cup M_e$ consists, in general, of paths (one of which contains e), alternating cycles and duplicated lines. Append all of the alternating paths (some of which may be single lines) to H and the resulting graph is not only nice, but 1-*extendable*. Moreover, the ear containing e is proper.

Now among all such ear systems (that is, those of the form $H+P_1+\cdots+P_m$ which are nice, 1-extendable and at least one P_i is proper) select one $\{P_1,\ldots,P_m\}$ which is inclusion-wise minimal. (That is, no proper subset of $\{P_1,\ldots,P_m\}$ has the properties that at least one is a proper ear and if all are attached to H, the resulting graph is nice and 1-extendable.) Let us denote the resulting graph by $H'=H+P_1+\cdots+P_s+P_{s+1}+\cdots+P_{s+r}$, where $P_1,\ldots,P_s$ are proper and $P_{s+1},\ldots,P_{s+r}$ are improper (that is, single lines). Let e_j be the single line of P_j, $s+1\le j\le s+r$. We must show $s\le 2$ and $r\le|\mathcal{P}(H')|-|\mathcal{P}(H)|$.

Let $H_0=H+P_{s+1}+\cdots+P_{s+r}$. Graph H_0 is then an elementary graph and $H'=H_0+P_1+\cdots+P_s$ is 1-extendable.

From H' build a new graph H'' by replacing each proper ear relative to H by a single line. Let us denote these lines, the endpoints of which may or may not already be adjacent in H, by $e_1,\ldots e_s$. Now by Remark 2 at the beginning of Section 5.1, each e_j is allowed in H''. Also H'' is 1-extendable. Hence we must have $\Phi(H_0+e_1+\ldots+e_s)>\Phi(H_0)$. But then by Lemma 5.4.5, there exist i, j, $1\le i$, $j\le s$ such that $\Phi(H_0+e_i+e_j)>\Phi(H_0)$. But then $\Phi(H+e_i+e_j+e_{s+1}+\cdots+e_{s+r})\ge\Phi(H_0+e_i+e_j)>\Phi(H_0)\ge\Phi(H)$. Thus there is a perfect matching F of $H+e_i+e_j+e_{s+1}+\cdots+e_{s+r}$, such that $F\cap\{e_i,e_j\}\ne\emptyset$. Let $F\cap\{e_i,e_j,e_{s+1},\ldots,e_{s+r}\}$ be denoted by $\{e_{i_1},\ldots,e_{i_k}\}$. Then $H_k=H+P_{i_1}+\cdots P_{i_k}$ is trivially 1-extendable and it contains either one or two proper ears, namely those corresponding to e_i and e_j. By the inclusion-wise minimality of H', we have $H_k=H'$ and so $s\le 2$.

It remains to show that $r\le|\mathcal{P}(H')|-|\mathcal{P}(H)|$. For every i, $s+1\le i\le s+r$, $H'-e_i$ is not elementary by the inclusion-wise minimality used in the choice of H'. On the other hand, $H''-e_i$ *is* elementary because it has an elementary spanning subgraph H. So, again by Remark 2 in Section 5.1, $H''-e_i$ will contain at least one $e_j(1\le j\le s)$ not allowed in $H''-e_i$. But since this e_j was allowed in H'', the endpoints of e_j lie in *different* classes of $\mathcal{P}(H'')$, but in the *same* class of $\mathcal{P}(H''-e_i)$. This follows from Theorem 5.2.2(b). But since $\mathcal{P}(H'')$ refines $\mathcal{P}(H''-e_i)$ by Remark 2 after Theorem 5.2.2, it follows that $|\mathcal{P}(H''-e_i)|<|\mathcal{P}(H'')|$. Now $H''-e_s+1-\cdots-e_s+r$ is elementary, since it too contains H as an elementary spanning subgraph. Moreover, $\mathcal{P}(H''-e_s+1-\cdots-e_s+r)$ refines $\mathcal{P}(H)$. So by Lemma 5.4.13, $|\mathcal{P}(H'')|\ge|\mathcal{P}(H''-e_s+1-\cdots-e_s+r)|+r\ge|\mathcal{P}(H)|+r$. But $|\mathcal{P}(H')|\ge|\mathcal{P}(H'')|$, since by Remark 1 after Theorem 5.2.2, the partition $\mathcal{P}(H')$ restricted to $V(H'')$ is just $\mathcal{P}(H'')$. The proof of the lemma is thus complete. ■

5.4.15. THEOREM. *If G is weakly minimal elementary and contains a nice cycle of length l, then* $|E(G)| \le \frac{3}{2}|V(G)| - l/2 + |\mathcal{P}(G)| - 2$.

PROOF. Consider an ear decomposition for G, say $G_1, \ldots, G_m = G$, where G_1 is a nice l-cycle and G_{j+1} arises from G_j by attaching at least one proper ear as long as G_j is non-spanning, and as few additional ears as possible. Let G_k be the first spanning subgraph in this sequence.

First we will show by induction on j, $0 < j \le k$, that G_j satisfies the inequality stated in the theorem; that is,

$$|E(G_j)| \le |\mathcal{P}(G_j)| - 2 + \frac{3}{2}|V(G_j)| - \frac{\ell}{2}. \tag{5.4.2}$$

For $j = 1$ the inequality is trivial. Suppose it holds for j and let G_{j+1} arise from G_j by attaching s proper ears and r improper ears. Then

$$\begin{aligned}
|E(G_{j+1})| &= |E(G_j)| + (|E(G_{j+1})| - |E(G_j)|)\\
&= |E(G_j)| + (|V(G_{j+1})| - |V(G_j)| + r + s)
\end{aligned}$$

(by induction hypothesis)

$$\le |\mathcal{P}(G_j)| - 2 + \frac{3}{2}|V(G_j)| - \frac{\ell}{2} + (|V(G_{j+1})| - |V(G_j)| + r + s)$$

(using Lemma 5.4.14)

$$\begin{aligned}
&\le |\mathcal{P}(G_j)| - 2 + \frac{3}{2}|V(G_j)| - \frac{\ell}{2} + |V(G_{j+1})| - |V(G_j)|\\
&\quad + |\mathcal{P}(G_{j+1})| - |\mathcal{P}(G_j)| + s
\end{aligned}$$

(since each proper ear contains at least two new points)

$$\begin{aligned}
&\le |\mathcal{P}(G_j)| - 2 + \frac{3}{2}|V(G_j)| - \frac{\ell}{2} + |V(G_{j+1})| - |V(G_j)|\\
&\quad + |\mathcal{P}(G_{j+1})| - |\mathcal{P}(G_j)| + \frac{\ell}{2}\big(|V(G_{j+1})| - |V(G_j)|\big)\\
&= |\mathcal{P}(G_{j+1})| - 2 + \frac{3}{2}|V(G_{j+1})| - \frac{\ell}{2},
\end{aligned}$$

as desired.

Next let $E(G) - E(G_k) = \{e_1, \ldots, e_t\}$. Now G_k is a spanning elementary subgraph of $G - e_i$ for each $i = 1, \ldots, t$, so $G - e_i$ is elementary for $i = 1, \ldots, t$. But then by the definition of weakly minimal elementary, we must have $|\mathcal{P}(G - e_i)| < |\mathcal{P}(G)|$ for each i. Hence by Lemma 5.4.13, we have $|\mathcal{P}(G_k)| \le |\mathcal{P}(G)| - t$. So using inequality (5.4.2), we obtain

$$\begin{aligned}|E(G)| &= |E(G_k)| + t \\ &\leq |\mathcal{P}(G_k)| - 2 + \frac{3}{2}|V(G_k)| - \frac{\ell}{2} + t \\ &\leq |\mathcal{P}(G)| - 2 + \frac{3}{2}|V(G_k)| - \frac{\ell}{2}.\end{aligned}$$ ■

5.4.16. COROLLARY. *If G is a minimal bicritical graph with $p \geq 6$ points, then $|E(G)| \leq 5(p-2)/2$. Moreover, equality holds if and only if G is a member of the family $\mathcal{M}$ shown in Figure* 5.4.12.

PROOF. By Exercise 5.4.3 we know G contains a cycle of length at

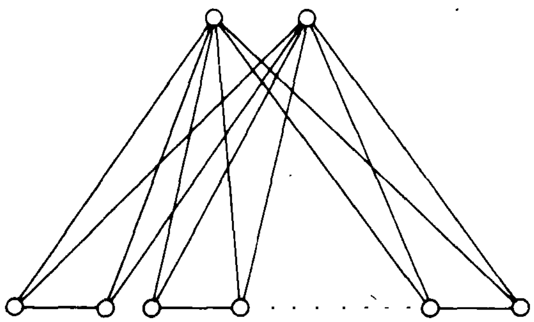


FIGURE 5.4.12. The extremal family $\mathcal{M}$

least 6. The inequality now follows from Theorem 5.4.15. The uniqueness of $\mathcal{M}$ is left to the reader. ■

The same upper bound — $5(p-2)/2$ — also follows for the number of lines in a minimal elementary graph ($\neq K_2$), but it is *never* sharp in this case. The best possible upper bound is not known.

5.5. More about Factor-critical and Bicritical Graphs

All of the results on ear structure of 1-extendable graphs developed in the earlier sections of this chapter carry over to bicritical graphs of course. But what can be done along these lines with respect to *factor-critical* graphs? Note that while all 1-extendable graphs must be 2-connected, factor-critical graphs, although necessarily 2-line-connected, may have cutpoints. So we must extend our notion of ears to allow the two endpoints of an ear to coincide. (See the so-called "closed ears" mentioned — but never used — in Section 5.4.) It is then an easy matter to obtain an ear decomposition of any factor-critical graph. (See Lovász (1972d).)

5.5.1. THEOREM. *Every factor-critical graph G can be represented as $P_0+P_1+\cdots+P_r$ where $P_0=K_1$ and for each i, P_{i+1} is an odd path having only its two endpoints in common with $P_0+\cdots+P_i$ or P_{i+1} is an odd cycle with precisely one point in common with $P_0+\cdots+P_i$.*

PROOF. Let $P_0 = v$ be any point of G and suppose u and w are two other points adjacent to v. Now $G-v$ has a perfect matching M_v and $G-u$, a perfect matching M_u. There must then be an $M_u - M_v$ alternating path P in G joining u and v and its length is necessarily even. Take odd cycle $P_1 = P+uv$ as our first ear. Let us agree to call M_v our "reference matching". If $P+uv$ spans $V(G)$ we simply add all remaining lines of G one at a time as single-line ears. Clearly each intermediate graph in the sequence is factor-critical.

So suppose $P_1 = P+uv$ does not span $V(G)$. Then there must be a line ab in G having exactly one endpoint — say a — on ear P_1. Let M_b be a perfect matching of $G-b$. Then there is an $M_v - M_b$ alternating path Q joining b and v. Traverse Q from b to the first point of P_1 encountered which we will call c. Then define $P_2 = Q[b,c]+ab$ to be our second ear. We simply continue this procedure, keeping M_v as our reference matching throughout, until all lines of G belong to an ear. ∎

Simple enough. Too simple, in fact, to help us very much! However, if we refine our techniques a bit, we can obtain a better "ear theorem" which, in fact, will help provide us with a lower bound for the number of near-perfect matchings for factor-critical graphs. By Lemma 5.5.1 we may concentrate on 2-connected factor-critical graphs. We now show that any such graph has an ear decomposition where each ear added is *open*. Let us call such a decomposition an **open ear decomposition**.

5.5.2. THEOREM. *Let G be any 2-connected factor-critical graph. Then G can be decomposed as $P_0+P_1+\cdots+P_r$ where $P_0=K_1, P_1$ is an odd cycle and P_i is an open odd path for all i. Hence $G_i = P_0+\cdots+P_i$ is a 2-connected factor-critical graph for $i=1,\ldots,r$.*

PROOF. Consider the ear decompositions for G guaranteed to exist by Theorem 5.5.1. Among all such, choose one in which $P_2, \ldots P_i$ are all open and i is maximum with respect to this property. As usual, we denote $P_0 + \cdots + P_i$ by G_i.

Now G_i is a block, but G_{i+1} is not, and since G itself is a block, there is a j such that j is the maximum l such that G_i is a block of G_l. Now consider ear P_{j+1}. We know it is open by the maximality of j, so let us consider the several cases corresponding to where its endpoints may lie.

Case 1. Both endpoints of P_{j+1} lie in G_i. Then

$$P_0 + \cdots + P_i + P_{j+1} + P_{i+1} + \cdots + P_j + P_{j+2} + \cdots + P_r$$

is a valid ear decomposition for G in which the first $i+2$ ears are open, contradicting the maximality of i.

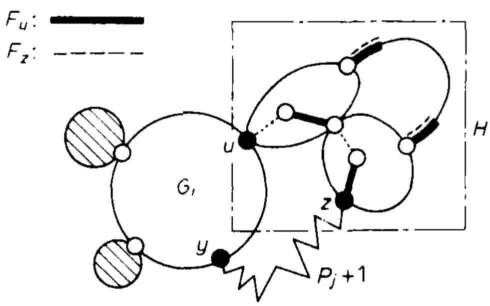


FIGURE 5.5.1.

Case 2. Ear P_{j+1} has one endpoint y in G_i and the other endpoint z on an ear P_{i+k}, for some $k > 0$. (So $z \notin V(G_i)$). Now P_{j+1} joins a block H at u ($H \neq G_i$) to a point $y \in V(G_i) - V(H)$. (See Figure 5.5.1.) Note that $y \neq u$ by the maximality of j.

Moreover, H is factor-critical, so let F_u (respectively, F_z) be a perfect matching of $H - u$ (respectively, $H - z$). Then u and z are joined by an $F_u \cup F_z$ alternating path Q. But if we let $G''_{i+1} = P_0 + \cdots + P_i + (P_{j+1} + Q)$, then graph G''_{i+1} can be completed to an ear decomposition of G (since its complement contains a perfect matching of G). But this contradicts the maximality of i, since $Q + P_{j+1}$ is an open ear.

Case 3. Suppose there are two different blocks H, H' attached to G_i at points u, u' $(u \neq u')$ and P_{j+1} joins points y and y', $y \in V(H)-V(G_i)$, $y' \in V(H')-V(G_i)$, $y \neq u$, and $y' \neq u'$. (See Figure 5.5.2.)

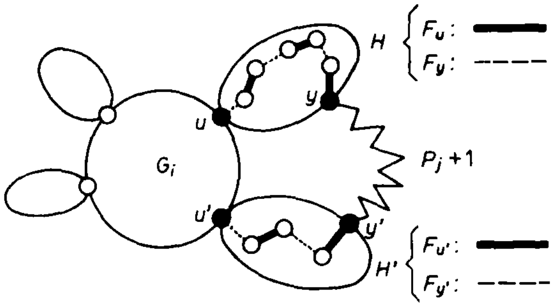


Figure 5.5.2.

Now H and H' are both factor-critical, so let F_u, $F_{u'}$, F_y and $F_{y'}$ be perfect matchings of $H-u$, $H'-u'$, $H-y$ and $H'-y'$ respectively. Then there is an $F_u \cup F_y$ (respectively, $F_{u'} \cup F_{y'}$) alternating path in H (respectively, H') joining u and y (respectively, u' and y'); call it Q (respectively, Q'). Then $Q+P_{j+1}+Q'$ is an open odd path. But then $P_0+\cdots+P_i+(Q+P_{j+1}+Q')$ can be completed to an ear decomposition of G, again contradicting the maximality of i. ■

Let us note in passing that ear decompositions of factor-critical graphs, be they open or not, are certainly not unique in general. However, the *number* of ears r in any such decomposition is invariant and, in fact, we always have $r = |E(G)|-|V(G)|$. Moreover, life is simpler here than it was for bicritical graphs (see Section 5.4) in that we can build up graph G *one ear at a time*, while preserving 2-connectedness and factor-criticality at each step.

Several other observations are in order here as well. Recall that a subgraph H of any graph G is said to be **nice** if $G-V(H)$ has a perfect matching. Now if factor-critical graph G has an ear decomposition $P_0+\cdots+P_r$, then for each i, $0 \leq i \leq r$, $P_0+\cdots+P_i$ is nice. This is immediate by observing that each ear P_i, be it open or closed, contains a matching which covers all of its points except the endpoint, or endpoints.

Conversely, any ear decomposition of a nice factor-critical subgraph extends to an ear decomposition of G.

Now let us return to the problem of finding a lower bound for the number of near-perfect matchings in a factor-critical graph G. We will now show that if G is any 2-connected factor-critical graph, then G always has at least $|E(G)|$ near-perfect matchings. This result is due to Pulleyblank (1973) who, in fact, proved that such a graph has $|E(G)|$ near-perfect matchings whose incidence vectors are *linearly independent* over $\Re$. We postpone the proof of this stronger fact until Chapter 7, because there its significance will be clear and its proof will be quite straightforward. However, the elementary proof of the following result can be refined to yield the linear independence. (See Exercise 5.5.4.)

5.5.3. THEOREM. *Every 2-connected factor-critical graph G contains $|E(G)|$ near-perfect matchings.*

PROOF. From the ear structure result for 2-connected factor-critical graphs (Theorem 5.5.2), we can write $G = P_0 + \cdots + P_r$, where $r \geq 0$, P_0 is an odd cycle and each ear P_{i+1} is an open path of odd length joining two different points of $P_0 + \cdots + P_i$.

We proceed by induction on r, the number of ears. If $r = 0$ the result is clear. So suppose that the conclusion holds true for $i \geq 0$ and let G be a 2-connected factor-critical graph with ear decomposition $P_0 + \cdots + P_{i+1}$. Let $G_i = P_0 + \cdots + P_i$. Further, suppose that $P_{i+1} = xu_1 \cdots u_{l-1}y$ for l odd and $l \geq 3$, or $P_{i+1} = xy$ (that is, $l = 1$).

If F is any near-perfect matching of G_i, then

$$F' = F \cup \{u_1u_2, \ldots, u_{l-2}u_{l-1}\}$$

is a near-perfect matching of G. So by the induction hypothesis we find $|E(G_i)| = |E(G)| - l$ near-perfect matchings in G.

If F_{u_j} is a perfect matching of $G - u_j$, $j = 1, \ldots, l-1$, we obtain $l-1$ further near-perfect matchings of G.

We now construct yet another near-perfect matching in G different from all the F''s and F_{u_j}'s formed above. In particular, it will cover points x and y using lines of ear P_{i+1}. Let F_x be any perfect matching of $G_i - x$. Matching F_x must cover y with a line $yz \in E(G_i)$. Then $F_0 = (F_x - \{yz\}) \cup \{xu_1, u_2u_3, \ldots, u_{l-1}y\}$ is a matching which is perfect in $G - z$ and hence near-perfect in G. ■

5.5.4. EXERCISE. Prove that the incidence vectors of the near-perfect matchings constructed above are linearly independent over $\Re$.

A different proof of this result will follow from Theorem 7.3.1. (See Exercise 7.3.2.) In fact, factor-critical graphs will play an important role in that section.

Next we turn our attention to factor-critical and bicritical graphs which are *minimal* with respect to line deletion. This approach has proved quite successful for elementary and positive surplus graphs. Theorem 4.2.11 tells us that it can be decided algorithmically whether a given graph is a subgraph of a minimal elementary bipartite graph. While no analogous result is known for minimal bicritical graphs, several excluded subgraphs for these graphs are known. (See Lovász and Plummer (1975a).) Here we only treat one such class in detail to illustrate the techniques. Yet another excluded subgraph is treated in Exercise 5.5.10. Similar questions for minimal factor-critical graphs will be considered in Exercises 5.5.12 through 5.5.19.

5.5.5. LEMMA. *If G is minimal bicritical and x is any line in G, then there is a set $S_x \subseteq V(G)$ such that $G - S_x - x$ has $|S_x|$ odd components and line x must join two of these components.*

PROOF. By minimality, $G - x$ is not bicritical and hence by Theorem 5.2.5, $G - x$ contains a set $S_x \subseteq V(G - x)$ such that $|S_x| \geq 2$ and $c_0(G - x - S_x) \geq |S_x| - 1$. Since $|S_x|$ and $c_0(G - x - S_x)$ have the same parity this implies that $c_0(G - x - S_x) \geq |S_x|$. But also by Theorem 5.2.5, we have $c_0(G - S_x) \leq |S_x| - 2$ and hence it follows that x must join two odd components of $G - S_x$. ■

There are several useful observations related to this theorem which deserve to be mentioned. If x is any line in a minimal bicritical graph G, there is a set of points S_x such that

(1) any cycle through x meets S_x;

(2) if u and v are points of S_x, then every perfect matching of $G - u - v$ contains x;

(3) any perfect matching of $G - x$ contains exactly one line joining each odd component of $G - S_x - x$ to S_x, and no line joining the even components (if any) of $G - S_x - x$ to S_x.

We shall need the next several results for our march toward excluded subgraphs.

5.5.6. COROLLARY. *If G is minimal bicritical and contains the subgraph shown in Figure 5.5.3, then every perfect matching of $G - a - d$ contains line bc.*

PROOF. Let $x = bc$. Then by the above theorem, $G - x$ has a set S_x such that $G - S_x - x$ has $|S_x|$ odd components. But then by remark (1) above, points a and d lie in S_x, and so by remark (2) above, every perfect matching of $G - a - d$ contains line x. ■

The proof of the next corollary is left to the reader.

5.5.7. COROLLARY. *Let G' be a subgraph of a minimal bicritical graph G and let x be a line of G'. Suppose for any set $S' \subseteq V(G')$ which separates*

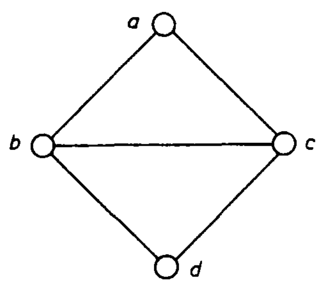


FIGURE 5.5.3.

the endpoints of x, $G' - S' - x$ has less than $|S'|$ odd components. Then no perfect matching of $G' - x$ extends to a perfect matching of G. ■

5.5.8. LEMMA. *If G is a minimal bicritical graph and contains the configuration of Figure* 5.5.4, *then every perfect matching of G containing ae also contains bc.*

PROOF. Suppose F is a perfect matching of G which contains ae, but not bc. Consider $G-b-d$. By Corollary 5.5.6, there is a perfect matching F' of $G-b-d$ which contains ce. But then there is an alternating $F \cup F'$ path P joining b and d.

Suppose P does not contain path aec. Then aec lies on an alternating $F \cup F'$ cycle. But then we may simply interchange lines on this cycle to get a new perfect matching F'' of $G-b-d$ which does not contain ce, contradicting Corollary 5.5.6.

Thus we may suppose P contains path aec. In this case traverse P starting at b. Denote the configuration of Figure 5.5.4 by H and let $H \cup P = G'$. Note that F contains a perfect matching of G' which extends to G. Also note that if $x = be$ or if $x = bc$, then with this choice of G' and x, the conclusion of Corollary 5.5.7 does not hold.

First suppose in our traversal of P, that a is encountered before c. Then let $P_1 = P[b,a], P_2 = P[c,d]$, and let $x = be$. On the other hand, let S' be any subset of $V(G')$ which separates b and e, the endpoints of x. So S' contains a and c together with k_i other points of P_i, $k_i \geq 0$, $i = 1, 2$. Obviously, if m is the number of odd components of $G' - S' - x$, then $m \leq k_1 + k_2 < |S'| = k_1 + k_2 + 2$. Thus Corollary 5.5.7 is contradicted.

H:

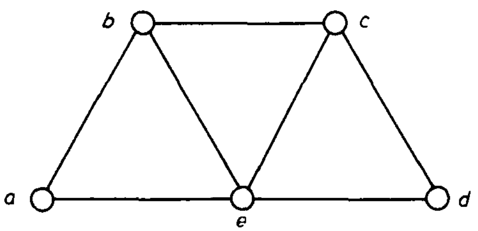


FIGURE 5.5.4.

Now suppose c is encountered before a. Let $x = bc$. Again let S' be any subset of G' separating b and c, the endpoints of line x. Then $e \in S'$ and in addition, at least $k \geq 2$ points of the odd cycle $G' - e - x$ belong to S'. So S' separates $G' - x$ into no more than $k - 1$ odd components, but $|S'| = k + 1$. Hence the hypothesis of Corollary 5.5.7 is again satisfied and again the corollary is contradicted. ■

Recall that a graph G is called a **wheel** if G consists of a cycle (called the **rim**) every point of which is joined to a single common point (called the **hub**) by a line (called a **spoke**). It is easy to check that any wheel on an even number of points is minimal bicritical. On the other hand, the reader may find the following result somewhat surprising!

5.5.9. THEOREM. *If G is minimal bicritical and is not a wheel, then G contains no wheel as a subgraph.*

PROOF. Suppose G properly contains a wheel subgraph W. Let F_0 be a perfect matching of G containing a spoke of W.

First let us suppose that W has an even number of points on its rim (and hence at least 4 such). Label W as in Figure 5.5.5 below and suppose the spoke in F_0 is ab. Then by Lemma 5.5.8, line $cd \in F_0$. Form a new perfect matching of G, $F' = F_0 - ab - cd + bc + ad$. Then again by Lemma 5.5.8, the rim line of W at e, other than de, is in F' and hence in F_0. Continuing in this way, we obtain the result that every second rim line of W is in F_0, a contradiction since F_0 contains a spoke.

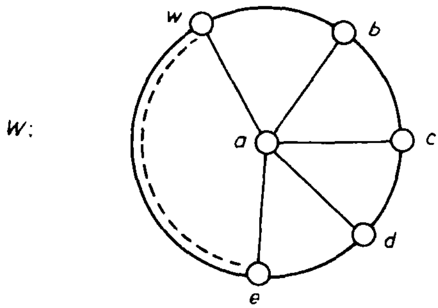


FIGURE 5.5.5.

Now suppose W contains an odd number of points on its rim, hence at least 3. Again let its hub be a. Since $G \neq W$, there is a line $x = bv$, say, with $b \in V(W)$, but $v \notin V(W)$. Moreover, if all such lines left W at hub a, G would have a cutpoint at a, a contradiction. So we may assume b is on the rim of W. Now let F' be a perfect matching of G containing x. Let F be any perfect matching of G containing a spoke of W. By the argument in the preceding paragraph of this proof, F contains a perfect matching of W. Thus we have

(*) If F is a perfect matching of G containing a spoke of W, then F contains a perfect matching of W.

Now $F'' = F_0 - E(W)$ is a perfect matching of $G - W$, so there is an alternating $F' \cup F''$ path P starting with line x which must return to W at some point r. If r is on the rim of W, let F''' be a perfect matching of $W - r - b$. But then $F''' \cup F'' - E(P) \cup (F' \cap E(P))$ is a perfect matching of G which contains a spoke of W, but does not contain a perfect matching of W. This contradicts statement (*) above.

Finally, suppose $r = a$, the hub of the wheel. Then P, acb, and awb are, respectively, one odd and two even paths joining the endpoints of line $x = ab$. Let F be a perfect matching of W containing ac (and so also de) and let $F_1 = F \cup F''$. Also let $G' = P + awb + acb + ab$. Then this G', x and perfect matching F_1 contradict Corollary 5.5.7, completing the proof of the theorem. ∎

5.5.10. EXERCISE. If G is minimal bicritical, then G does not contain $K_{3,3}$ as a subgraph. (Hint: Suppose, to the contrary, that G does contain a $K_{3,3}$ and that line x is a line of the $K_{3,3}$. Let x be as in Lemma 5.5.5.)

Now let us "switch gears" and return to the study of minimal *factor-critical* graphs.

Recall that in contrast to bicritical graphs, factor-critical graphs may have cutpoints. However, we saw earlier that we may reduce many questions regarding factor-critical graphs to the 2-connected case. The same is true for *minimal* factor-critical graphs.

5.5.11. EXERCISE. A graph is minimal factor-critical if and only if each of its blocks is minimal factor-critical.

We have seen that factor-critical graphs have an ear decomposition theory quite analogous to that of elementary bipartite graphs. This technique can be applied to the study of *minimal* factor-critical graphs as well and the results obtained are quite analogous to the results concerning minimal elementary bipartite graphs. Therefore we shall only state these results, relegating the proofs to exercises for the reader. Note, however, that the analogy with elementary graphs is not complete, for see Exercise 5.5.19.

5.5.12. EXERCISE. Every nice factor-critical subgraph of a minimal factor-critical graph is also minimal factor-critical.

5.5.13. EXERCISE. In every ear decomposition of a minimal factor-critical graph all ears are proper.

5.5.14. EXERCISE. Show that in a minimal factor-critical graph no nice cycle has a chord. Give an example of a minimal factor-critical graph G in which a (non-nice) cycle *does* have a chord.

5.5.15. EXERCISE. A minimal factor-critical graph cannot contain a cycle on 4 points.

5.5.16. EXERCISE. If a minimal factor-critical graph G contains a K_3, then this K_3 must be a block of G.

The reader is invited to show by example that a minimal factor-critical graph may indeed contain a cycle of any length greater than 4.

It is easy to find a minimal *bicritical* graph containing 3-cycles and 4-cycles. Two simple examples are K_4 and R_3, the triangular prism, the two "basis graphs" of Theorem 5.4.11. On the other hand, we have the following result.

5.5.17. EXERCISE. If H is an excluded subgraph in all minimal bicritical graphs, it is also excluded in all minimal factor-critical graphs.

Now clearly any graph, all the blocks of which are K_3's, is minimal factor-critical. (Recall that no factor-critical graph can contain a cutline.) These graphs all have $(3p-3)/2$ lines, where p is the number of points, and the next result shows that they are the densest minimal factor-critical graphs of all.

5.5.18. EXERCISE. If G is minimal factor-critical then $q \leq \frac{3}{2}(p-1)$ and equality holds if and only if G is a connected graph each block of which is a K_3.

The next result should be contrasted with Little's result, Theorem 5.4.4.

5.5.19. EXERCISE. Produce a 2-connected minimal factor-critical graph G containing two lines e and e' which do not lie on a nice cycle.

Several interesting families of factor-critical and bicritical graphs have been discovered and studied. We call a graph G a **Halin graph** if G can be drawn in the plane as a tree, with all non-endpoints having minimum degree 3, together with a cycle C passing through the endpoints of T. The proof of part (b) of the next theorem may be found in Lovász and Plummer (1975b).

5.5.20. THEOREM. *Let G be a Halin graph on p points. Then*

(a) *if p is odd, G is factor-critical, while*

(b) *if p is even, G is minimal bicritical.* ■

For further study of the matching structure of even Halin graphs see Pulleyblank (1980).

Recall that a graph G is said to be **cyclically k-line-connected** if G cannot be separated into two components, each containing a cycle, by the deletion of fewer than k lines. The proof of the next result is a straightforward application of Tutte's Theorem.

5.5.21. EXERCISE. If, for some $k \geq 3$, G is k-regular, cyclically $k+1$-line-connected and has an even number of points, then G is bicritical or elementary bipartite.

The final two main results of this chapter serve, we think, to further highlight the fundamental nature of the building block graphs studied in Chapters 3, 4 and 5; that is to say, the elementary bipartite, factor-critical and bicritical graphs. In particular, we will discuss 2-extendable graphs and point-transitive graphs.

A natural extension of the concept of 1-extendable graphs, which were studied extensively in Section 5.4, is the notion of an n-extendable graph, for $n > 1$. Let us define a graph G to be **n-extendable** $(n \geq 1)$ if it is connected, has a set of n independent lines and every set of n independent lines extends to a perfect matching. (See Plummer (1980).)

It is a trifle vexing that there are graphs which are n-extendable, but not $(n-1)$-extendable. For $n \geq 2$, a trivial example is the path of length $2n-1$. Here of course $|V(G)| = 2n$. However, we do have the following.

5.5.22. LEMMA. *Let p and n be integers with $n \geq 2$, p even and $p \geq 2n+2$. Let G be a graph with p points. Then if G is n-extendable, it is also $(n-1)$-extendable.*

PROOF. Suppose n, p and G satisfy the hypotheses of the lemma, but suppose G is not $(n-1)$-extendable. In particular, let X be a set of $n-1$ independent lines which do not extend to a perfect matching and let M be any perfect matching of G. Then $M \oplus X$ consists of some number of even cycles together with at least two alternating paths, each of which has both its first and last lines in M. Let P be the line set of one such path. Then $P \oplus X$ is a set of n independent lines which can be extended to a perfect matching. Moreover, this perfect matching will contain at least one line e not in $P \oplus X$, since $|P \oplus X| = n$ and $p \geq 2n+2$. But then $X \cup e$ is a set of n independent lines which extends to a perfect matching containing X, a contradiction. ∎

The class of 2-extendable graphs partitions nicely into two of our building block families mentioned above.

5.5.23. THEOREM. *Let G be 2-extendable and have $p \geq 6$ points. Then G is either bicritical or elementary bipartite.*

PROOF. Graph G is elementary by Lemma 5.5.22. Suppose it is not bicritical. Then by Theorem 5.2.5, there is a set $S \subseteq V(G)$, such that $|S| \geq 2$, $|S| \leq c_0(G-S)+1$ and hence by parity, $|S| \leq c_0(G-S)$. But

G contains a perfect matching and hence by Tutte's Theorem, $|S| = c_0(G-S)$. It is then clear that every perfect matching of G matches a point of S to a different odd component of $G-S$. But since G has no forbidden lines, S is independent and $G-S$ has no *even* components.

It remains only to show that each component of $G-S$ is a singleton. Suppose, to the contrary, that G has an odd component N having at least three points. Again, since G is 2-connected, there must be two independent lines x' and y' joining N to S. Since G is 2-extendable, $\{x', y'\}$ extends to a perfect matching M' of G. (Actually, by parity, M' must contain at least three lines joining N to S.) But each of the other $|S|-1$ odd components must have a line of M' joining it to a point of S. Hence M' contains at least $3 + c_0(G-S) - 1 = 3 + |S| - 1 = |S| + 2$ lines incident with S, a contradiction.

Thus G is bipartite with bipartition (S, T) where $|S| = |T|$. ∎

Finally we turn to the family of point-transitive graphs. These graphs have been the object of considerable study in graph theory. For a general survey of this work we refer the reader to Biggs (1974) and the bibliography therein.

We call a graph G **point-transitive** (respectively, **line-transitive**) if it has a point-transitive (respectively, line-transitive) automorphism group. (Note that neither type of transitivity implies the other.) Recall that, by Exercise 3.2.5, a connected point-transitive graph is either factor-critical or has a perfect matching. The next theorem improves this result.

5.5.24. THEOREM. *If G is connected, has p points and is point-transitive, then*

(a) *if p is odd, G is factor-critical, while*

(b) *if p is even, G is either elementary bipartite or bicritical.*

We shall need some preliminary results before proving this theorem. Let G be any k-line-connected graph and let $X \subseteq V(G)$. Recall that we denote by $\nabla(X)$ the set of all lines of G having exactly one endpoint in X. ($\nabla(X)$ is frequently called either the **cocycle** of X or the **coboundary** of X.) Also recall that a set of lines S in G is a **minimum line cut** if there is a set of points X ($\emptyset \neq X \neq V(G)$) such that $S = \nabla(X)$ and $|\nabla(X)| = k$. Let us denote by $r(X, Y)$ the number of lines of G joining point sets $X - Y$ and $Y - X$.

5.5.25. LEMMA. *If $\nabla(X)$ and $\nabla(Y)$ are minimum line cuts in a graph G, $X \cap Y \neq \emptyset$ and $X \cup Y \neq V(G)$, then $\nabla(X \cap Y)$ and $\nabla(X \cup Y)$ are also minimum line cuts and $r(X, Y) = 0$.*

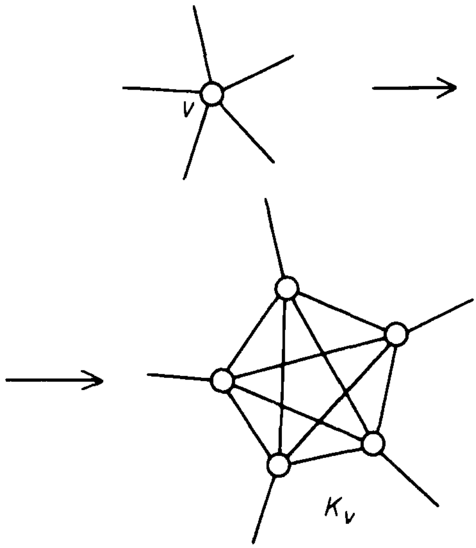


FIGURE 5.5.6.

PROOF. Simply by counting lines it is immediate that $|\nabla(X \cup Y)| + |\nabla(X \cap Y)| = |\nabla(X)| + |\nabla(Y)| - 2r(X,Y)$. Now since $\emptyset \neq X \cap Y \subseteq X \cup Y \neq V(G)$, both $\nabla(X \cap Y)$ and $\nabla(X \cup Y)$ are line cuts. Thus by minimality, if $|\nabla(X)| = |\nabla(Y)| = k$, we have:

$$2k \leq |\nabla(X \cup Y)| + |\nabla(X \cap Y)| = |\nabla(X)| + |\nabla(Y)| - 2r(X,Y) \leq 2k.$$

Hence equality holds throughout and the lemma follows. ∎

Now suppose v is a point of degree r in a graph G. If G' is the graph arising from G by replacing point v with a complete r-graph (as is illustrated in Figure 5.5.6), we shall say G' arises from G by an **r-clique insertion** at v. Of course if G is regular of degree r, then after any number of r-clique insertions the resulting larger graph will also be r-regular.

This construction is needed in the next lemma. The proof of r-line-connectivity was first obtained by Mader (1971).

5.5.26. LEMMA. *Let G be a simple, connected, point-transitive r-regular graph with $r \geq 3$. Then G is r-line-connected and either*

(a) *every minimum line cut of G is the star of a point, or*

(b) *G arises from a (not necessarily simple) point- and line-transitive r-regular graph G_0 by an r-clique insertion at each point of G_0. Moreover, every minimum line cut of G is the star of a point of G or a minimum line cut of G_0.*

PROOF. Let $\mathrm{Aut}(G)$ denote the automorphism group of G. Suppose $X \subseteq V(G)$, $\varnothing \neq X \neq V(G)$, $|\nabla(X)| = k$ is minimum over all line cuts and moreover, $|X|$ is minimum with respect to all these properties.

Claim 1. For every automorphism α in $\mathrm{Aut}(G)$, $\alpha(X) = X$ or $\alpha(X) \cap X = \varnothing$.

Suppose $\alpha(X) \cap X \neq \varnothing$. Then by the minimality of $|X|$, $|X| = |\alpha(X)| \leq p/2$, so $|\alpha(X) \cup X| = |\alpha(X)| + |X| - |\alpha(X) \cap X| \leq p/2 + p/2 - |\alpha(X) \cap X| < p$. Thus $\alpha(X) \cup X \neq V(G)$. Moreover, since α is an automorphism, $|\nabla(\alpha(X))| = |\nabla(X)| = k$ and thus $\alpha(X)$ is a minimum line cut. Thus by Lemma 5.5.25, $|\nabla(\alpha(X) \cap X)| = k$; that is, $\alpha(X) \cap X$ is a minimum line cut. Hence by minimality of $|X|$, $\alpha(X) \cap X = X$; that is, $X = \alpha(X)$ and Claim 1 is proved.

Now suppose $x \in X$ and $|\nabla(\{x\}) \cap \nabla(X)| = s$; that is, suppose x is adjacent to s points not in X and thus to $r - s$ points in X. Since G is point-transitive we may assume $s > 0$. Choose $y \in X$ and α in $\mathrm{Aut}(G)$ such that $\alpha(x) = y$. Then $\alpha(X) = X$ by Claim 1, and so $|\nabla(\{y\}) \cap \nabla(X)| = s$; that is, every point of X is adjacent to s points of $V(G) - X$. Thus $|\nabla(X)| \geq s \cdot |X| \geq s(r - s + 1) \geq r$. Now recalling that $|\nabla(X)| = k$ was chosen to be minimum over all line cuts and that G is r-regular, it follows immediately that $k = r$; that is, G has line connectivity r.

Suppose next that there is a set $X \subseteq V(G)$ with $1 < |X| \leq p/2$ such that $|\nabla(X)| = r$. Let X_0 be a minimal such X. Note that $G[X_0]$ is connected by the r-line-connectivity of G.

Claim 2. If $\alpha \in \mathrm{Aut}(G)$, then $\alpha(X_0) = X_0$ or $\alpha(X_0) \cap X_0 = \varnothing$.

Suppose, on the contrary, that $\alpha(X_0) \neq X_0$ and $\alpha(X_0) \cap X_0 \neq \varnothing$. As in Claim 1, we have $|\alpha(X_0) \cup X_0| = |\alpha(X_0)| + |X_0| - |\alpha(X_0) \cap X_0| \leq p/2 + p/2 - |\alpha(X_0) \cap X_0| < p$, so $\alpha(X_0) \cup X_0 \neq V(G)$ and so by Lemma 5.5.25, $\nabla(\alpha(X_0) \cap X_0)$ and $\nabla(\alpha(X_0) \cup X_0)$ are minimum line cuts and $q(X_0, \alpha(X_0)) = 0$. Moreover, $|\alpha(X_0) \cap X_0| < |X_0|$ and so by the minimality of X_0, $|\alpha(X_0) \cap X_0| = 1$.

To see that $\alpha(X_0) - X_0$ also determines a minimum line cut, we proceed as follows. We know that $|X_0| > 1$ and hence $|\alpha(X_0)| > 1$, so $\alpha(X_0) - X_0$ and $X_0 - \alpha(X_0)$ are both non-empty. Hence by Lemma 5.5.25, the set $(V(G) - X_0) \cap \alpha(X_0) = \alpha(X_0) - X_0$ does indeed determine a minimum line cut. But again by minimality, $|\alpha(X_0) - X_0| = 1$. So $|X_0| = 2$. But $G[X_0]$ is connected and is thus a line. On the one hand then, since G is r-regular, $|\nabla(X_0)| = 2r - 2$ and on the other, by hypothesis, $|\nabla(X_0)| = r$. So we obtain $2r - 2 = r$ or $r = 2$, contrary to one of the hypotheses of this lemma, and Claim 2 is proved.

Claim 3. $G[X_0] = K_r$. Let x be any point of X_0 such that x is adjacent to some s ($s > 0$) points of $\nabla(G) - X_0$. Choose any $y \in X_0$ and any α in $\mathrm{Aut}(G)$ such that $\alpha(x) = y$. Then $\alpha(X_0) = X_0$, by Claim 2, and so y too is adjacent to s points of $V(G) - X_0$. So $|\nabla(X_0)| = s \cdot |X_0| \geq s(r - s + 1) \geq r$. But recall X_0 was chosen so that $|\nabla(X_0)| = r$. Hence equality holds here and, since $|X_0| > 1$, it follows that $s = 1$ and $|X_0| = r$. Since G is r-regular the Claim follows.

Now let $X_0, X_1, \ldots, X_t$ be the images of X_0 under $\mathrm{Aut}(G)$. By point-transitivity, $X_0 \cup \cdots \cup X_t = V(G)$ and by Claim 2, point sets $X_0, \ldots, X_t$ are disjoint. Thus every automorphism of G permutes the X_i's, $i = 1, \ldots, t$.

For $i = 0, \ldots, t$ shrink each X_i to a single point v_i and call the resulting graph G_0. Clearly G_0 has a point-transitive automorphism group. To see that $\mathrm{Aut}(G_0)$ is also *line*-transitive, let uv and xy be lines in G_0. Let $\overline{uv}$ and $\overline{xy}$ be corresponding lines of G and suppose $\overline{u} \in X_i$ and $\overline{x} \in X_j$. Choose α in $\mathrm{Aut}(G)$ such that $\alpha(\overline{u}) = \overline{x}$. Then, of course, $\alpha(X_i) = X_j$. We must show that if $\overline{v} \in X_k$, $\overline{y} \in X_l$ and $\alpha(\overline{v}) \in X_m$, that $l = m$. But if $l \neq m$, then $\overline{x}$ is adjacent to two points, $\overline{y}$ and $\alpha(\overline{v})$, not in X_j and hence $\overline{x}$ has degree $r - 2$ in X_j, contradicting the fact that $G[X_j]$ is the clique K_r. So $l = m$ and thus $\alpha(X_k) = X_l$. But then α induces an automorphism α_0 of G_0 such that $\alpha_0(uv) = xy$ and hence G_0 is line-transitive as claimed.

Finally, let $\nabla(Y)$ be a minimal line cut in G. By symmetry we may suppose that $|Y| \leq p/2$. Suppose also that $|Y| > 1$. We want to show that Y is the union of some of the X_i's in $\{X_0, \ldots, X_t\}$. Suppose not. Then there is an X_i such that $Y \cap X_i \neq \emptyset$ and $(V(G) - Y) \cap X_i \neq \emptyset$. Clearly $|X_i| \leq p/2$ and so $Y \cup X_i \neq V(G)$. But then by Lemma 5.5.25, $X_i \cap Y$ has $|\nabla(X_i \cap Y)| = r$ and by the minimality of X_0, $|Y \cap X_i| = 1$. Then, however, $(V(G) - Y) \cup X_i \neq V(G)$ and again by Lemma 5.5.25, $|\nabla((V(G) - Y) \cap X_i)| = r$, and again by the minimality of X_0, we have $|(V(G) - Y) \cap X_i| = 1$. So $r = |X_i| = 2$, a contradiction. This completes the proof of Lemma 5.5.26. ∎

We are now prepared to prove Theorem 5.5.24.

PROOF (of Theorem 5.5.24). First note that if $p = |V(G)|$ is odd then G is critical by the Gallai-Edmonds Theorem and transitivity. So we may suppose p to be even.

Since G is connected, it must be r-regular for some $r \geq 2$. If $r = 2$, G is an even cycle and hence trivially elementary bipartite. So let us assume that G is regular of degree $r \geq 3$.

Choose any set of points X, $\emptyset \neq X \subseteq V(G)$ with $|X| \leq p/2$. We will show that either $|X| = p/2$, $\{X, V(G)-X\}$ is a bipartition and G is elementary bipartite, or $c_0(G-X) < |X|$. If this second case occurs for all X, graph G is bicritical.

Choose any set of points X, $\emptyset \neq X \subseteq V(G)$ and suppose $c(G-X) \geq |X|$. Let these components be $G_1, \ldots, G_k$, where $k \geq |X|$. Graph G is r-line-connected by Lemma 5.5.26, so there are at least kr lines with one endpoint in $\bigcup_{i=1}^{k} G_i$ and the other in X. On the other hand, since G is r-regular these same lines are no more than $r \cdot |X|$ in number. Hence we may conclude that $|X| = k$, X is an independent set and each $\nabla(G_i)$ is a minimum cut for $i = 1, \ldots, k$.

If every G_i is a singleton, then $\{X, V(G)-X\}$ is a bipartition and G is elementary as well, by the regularity assumption. So suppose some G_i is not a singleton. Thus again by Lemma 5.5.26, G must arise from G_0 by r-clique insertions at each point of G_0. Suppose $v \in X$ and let K_v be the r-clique containing v. Then since X is independent, $K_v \cap X = \{v\}$ and hence $K_v - v \subseteq G_i$ for some i. But $\nabla(G_i)$ is a minimum cut which is not the star of a point, so it corresponds to a minimum cut of G_0 and hence is an independent set of lines. But this contradicts the fact that $\nabla(G_i)$ contains the $r-1$ mutually adjacent lines of K_v incident with v.■

5.5.27. EXERCISE. (a) Find a graph which is 2-extendable, but not point-transitive.

(b) Find a point-transitive graph (with an even number of points) which is not 2-extendable.

This Page Intentionally Left Blank

6

Some Graph-theoretical Problems Related to Matchings

6.0. Introduction

In this chapter we begin by discussing an important analogue of the matching problem, the so-called 2-matching problem. It has much in common with the matching problem, particularly from the point of view of linear programming (which will be taken up in Chapter 7). On the other hand, the 2-matching problem turns out to be much simpler than the matching problem in many respects; in fact, it is more or less equivalent to the bipartite matching problem, as will be shown by some of the proofs to follow. It deserves to be treated here for several reasons. (1) It may be viewed as a relaxation of the general matching problem, and some non-trivial connections between the two may be exploited in the solution of the matching problem. (2) It may be viewed as a relaxation of the minimum point cover problem, which is an NP-complete problem, and 2-matchings play an important role in some investigations into the point packing problem as well. (This will be briefly discussed in Chapter 12.) (3) It may also be viewed as a relaxation of the Hamilton cycle problem.

Some interesting results toward bridging the gap between the 2-matching and Hamilton cycle problems are discussed in Section 4. Then after treating the famous Chinese Postman Problem, we venture further afield to discuss some of the more distant relatives on the family tree of matching. These include various problems involving optimality of paths, cycles, joins and cuts.

6.1. 2-matchings and 2-covers

A **2-matching** of a graph G is an assignment of weights 0, 1 or 2 to the lines of G such that the sum of weights of lines incident with any given point is at most 2.

We shall sometimes say that those lines of weight 1 or 2 "lie in the 2-matching", whereas the lines of weight 0 do not. The sum of weights in a 2-matching ω will be called the **size** of ω and will be denoted by $|\omega|$. The maximum size of a 2-matching in G will be denoted by $\nu_2(G)$.

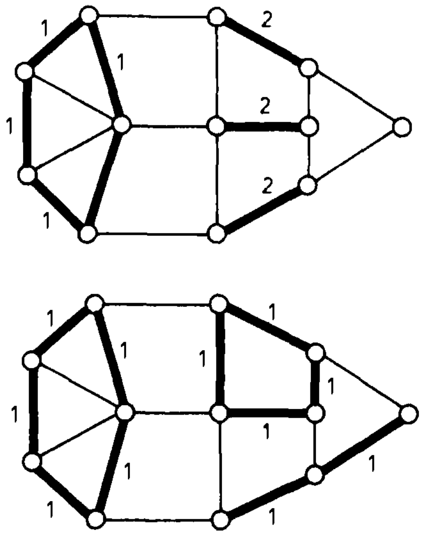


FIGURE 6.1.1. A basic and a non-basic 2-matching

Often it will be convenient to view a 2-matching as a collection of lines of G (in which a line may occur more than once) such that at most 2 lines in this collection contain any point. The **sum** of two matchings (i.e. their union in which the common lines occur twice) is a 2-matching. Hence it follows that $\nu_2(G) \geq 2\nu(G)$.

It is clear that in every 2-matching the lines with weight 2 are disjoint from each other and from the lines with weight 1. The lines with weight 1 form point-disjoint paths and cycles. We say that the 2-matching is **basic** if its lines with weight 1 form only odd cycles. (Such odd cycles must clearly be point-disjoint.) See Figure 6.1.1.

If we have a 2-matching which contains a path P of lines with weight 1, we may double the weight of the $1^{st}, 3^{rd}, \ldots$ line of the path (starting from either end), and reduce the weight of the remaining lines to 0. In this way we get another 2-matching whose size is at least as large as that of the original. Similarly, if there is an even cycle of lines of weight 1 in a 2-matching we can construct a new 2-matching of the same size by increasing and decreasing the weights of the lines of the cycle alternately. Thus we have proved the following simple fact:

6.1.1. LEMMA. *Every graph contains a maximum 2-matching which is basic.* ∎

The reader can easily show that in fact we have a bit stronger result:

6.1.2. EXERCISE. Suppose we are given a graph G and a line xy in G. Then if xy lies in some maximum 2-matching of G, it lies in a basic maximum 2-matching.

Just as König's Minimax Theorem for bipartite graphs relates matchings to covers, 2-matchings can be related to 2-covers in general graphs. A **2-cover** of a graph G is an assignment of weights 0, 1 and 2 to the points such that the sum of weights of the two endpoints of any line is at least 2. The sum of all weights is called the **size** of the 2-cover. The minimum size of any 2-cover of G will be denoted by $\tau_2(G)$. Since assigning weight 1 to all points results in a 2-cover, we have $\tau_2(G) \leq p$.

The **sum** of two point covers is just the union of the two sets in which the common points occur twice. Thus the sum of two covers is a 2-cover. Hence it follows that $\tau_2(G) \leq 2\tau(G)$. A 2-cover is said to be **basic** if every component of the graph spanned by the points of weight 1 in the cover is non-bipartite. It is easy to see that there is always a 2-cover of minimum size which is basic.

If A denotes the set of points with weight 0 in a 2-cover, then trivially A must be an independent set of points and the neighbors of A must therefore have weight 2. The remaining points may have weight 1 or 2, but if we find a point non-adjacent to A with weight 2, we may lower its weight to 1 and still have a 2-cover. So, for example, in a minimum 2-cover the points not adjacent to A will have weight 1. Conversely, if A is any independent set of points (including $A = \emptyset$), then by assigning weight 0 to the points in A, weight 2 to the points in $\Gamma(A)$ and weight 1 to all remaining points, we obtain a 2-cover, which we shall call **the 2-cover determined by the independent set** A. Thus we have proved:

6.1.3. LEMMA. *Every minimum 2-cover is the 2-cover determined by the set of its independent points. Consequently,* $\tau_2(G) = \min\{p - |A| + |\Gamma(A)| \mid A \subseteq V(G),\ A \text{ independent}\}$. ∎

Now we come to the main theorem relating ν_2 and τ_2. The result, though not stated in this form, is essentially due to Tutte (1953).

6.1.4. THEOREM. *If G is any graph, then* $\nu_2(G) = \tau_2(G)$.

PROOF. Let us construct a bipartite graph G^b as follows. We replace every point v of G by two new points v' and v'', and for every line uv of G we construct the two lines $u'v''$ and $u''v'$. The statement of the theorem will follow from König's Theorem 1.1.1 if we can verify the following two equations:

$$\nu_2(G) = \nu(G^b). \tag{6.1.1}$$

$$\tau_2(G) = \tau(G^b). \tag{6.1.2}$$

To prove equation 6.1.1 let M be a maximum matching in G^b. Then letting the weight of uv be the number $|M \cap \{u'v'', u''v'\}|$, we obtain a 2-matching of G of the same size as M. This proves that $\nu_2(G) \geq \nu(G^b)$. Second, let $\omega : E(G) \to \{0, 1, 2\}$ be a maximum 2-matching of G. By Lemma 6.1.1, we may assume that ω is basic. Define a set M^b of lines of G^b as follows. If $\omega(uv) = 2$, then put both of $u'v''$ and $u''v'$ into M^b. If C is a cycle formed by lines of weight 1, and $v_1, \ldots, v_r$ are the points of C in one of the two routings of C, then put the lines $v_1'v_2'', v_2'v_3'', \ldots, v_n'v_1''$ into M^b. The set M^b defined in this way is a matching of G^b of the same size as ω. Thus equation 6.1.1 is verified.

The verification of equation 6.1.2 is quite similar and is left to the reader. ∎

The construction used in this proof also shows that a maximum 2-matching as well as a minimum 2-cover can be computed in polynomial time.

A 2-matching is called **perfect** if the sum of weights of lines incident with any point is exactly 2. Equivalently, a 2-matching is perfect if and only if its size is p. Note that the existence of a perfect 2-matching in a graph G means that G contains a system of point-disjoint cycles and lines which cover all points. Such a system of cycles and lines was called a **q-factor** by Tutte (1953), who proved the following result.

6.1.5. COROLLARY. *A graph G has a perfect 2-matching (i.e., a q-factor) if and only if $|\Gamma(A)| \geq |A|$ for every independent set A of points.*

PROOF. First suppose G has a perfect 2-matching. Then $\nu_2(G) = p = \tau_2(G)$ by Theorem 6.1.4 and hence for all independent sets $A \subseteq V(G)$, $|\Gamma(A)| - |A| \geq 0$ by Lemma 6.1.3.

Conversely, if for all independent sets $A \subseteq V(G)$, $|\Gamma(A)| - |A| \geq 0$, then $p \leq \tau_2(G)$ by Lemma 6.1.3. Since, as remarked above $\tau_2(G) \leq p$ always holds, we have $\tau_2(G) = p$ and so by Theorem 6.1.4, $\nu_2(G) = p$. ∎

It is immediate that in a bipartite graph every basic 2-matching arises by taking every line of some matching with weight 2. So a bipartite graph has a perfect 2-matching if and only if it has a perfect matching. In fact, Corollary 6.1.5 is essentially equivalent to Frobenius' Marriage Theorem (Corollary 1.1.4).

6.1.6. EXERCISE. Derive the theorem of Fulkerson, Hoffman and McAndrew (see Exercise 3.4.16) from Corollary 6.1.5.

Let us consider for a moment the following two properties of graphs:
(1) G has a perfect matching;
(2) for every independent set $A \subseteq V(G)$, $|\Gamma(A)| \geq |A|$.
For bipartite graphs, these two properties are equivalent and this fact provides a good characterization of both. For non-bipartite graphs they are not equivalent, and so to obtain good characterizations for each, two separate results are needed. "The" Tutte Matching Theorem 3.1.1 yields a good characterization of property (1). On the other hand, Corollary 6.1.5 is a good characterization of property (2).

6.2. 2-bicritical and Regularizable Graphs

In this section we discuss briefly the 2-matching analogues of elementary and bicritical graphs. The main reason for doing so is that they give an answer to some other problems which have been studied in a different setting. Since the proof of Theorem 6.1.4 shows that the 2-matching problem can be reduced to the bipartite matching problem in a very simple way, we do not attempt to give a full structure theory analogous to the one developed in Chapter 4, but only point out some simplifications and applications.

A graph G will be called **2-bicritical** if $|V(G)| \geq 2$ and $G-v$ contains a perfect 2-matching for every point $v \in V(G)$. (One might wonder why we do not call these graphs 2-*critical*. The reason is that their properties turn out to be more analogous to the properties of bicritical graphs. For example, a 2-bicritical graph will always contain a perfect 2-matching.)

The following characterization of 2-bicritical graphs (see Pulleyblank (1979) is an immediate consequence of the definition and of Theorem 6.1.5.

6.2.1. LEMMA. *A graph G is 2-bicritical if and only if $|\Gamma(A)| > |A|$ holds for every non-empty independent set A of points.* ∎

6.2.2. COROLLARY. *Every 2-bicritical graph contains a perfect 2-matching.* ∎

6.2.3. EXERCISE. Every factor-critical graph with more than one point is 2-bicritical.

A graph G will be called **regularizable** if it is possible to replace every line of G by some non-empty set of parallel lines so as to obtain a regular graph. Regularizable graphs were introduced and studied by Berge (1978a, 1978b, 1978c, 1978d, 1979, 1981). Here we discuss them only from the point of view of the theory of 2-matchings.

We present two characterizations of regularizable graphs. The proof of the first will depend upon a classical theorem of Petersen (1891). A set F of lines of G is called a **2-factor** if every point is incident with exactly two lines of F. It is obvious that 2-factors are in a natural one-to-one correspondence with those perfect 2-matchings in which every weight is 0 or 1.

In Chapter 10 we shall state a necessary and sufficient condition for the existence of a 2-factor. Here, however, the following much simpler result of Petersen (1891) will suffice.

6.2.4. THEOREM. *Let G be a $(2r)$-regular graph. Then $E(G)$ can be decomposed into the union of r line-disjoint 2-factors.*

PROOF. We shall use a construction similar to, but different from, the method used in the proof of Theorem 6.1.4. First we remark that it suffices to consider connected graphs. A connected graph with even degrees has an Euler trail. Traversing this Euler trail we get an orientation D of G such that every point has indegree and outdegree $= r$.

Next we replace every point $v \in V(D)$ by two points v', v'', and for every directed line $uv \in E(D)$ we draw in *one* line from u' to v''. Since D has in- and outdegrees equal to r the resulting bigraph G' is r-regular. By Lemma 1.3.12, the lines of G' can be decomposed into r perfect matchings. Now if we identify v' and v'' for every v, we recover the graph G, and these r perfect matchings of G' will be mapped onto r 2-factors of G which partition the lines. ∎

6.2.5. THEOREM. *A graph G is regularizable if and only if for each line e of G, there exists a perfect 2-matching of G in which e has weight 1 or 2.*

PROOF. Assume first that G is regularizable, and let $e \in E(G)$. By definition we can replace every line of G by parallel lines and obtain a regular graph G_1. Trivially we may assume that G_1 has even degrees, since this can be achieved by doubling every line if necessary. By Petersen's Theorem 6.2.4, G_1 can be decomposed into 2-factors. Let F be a 2-factor of G_1 containing any line parallel to e. For $f \in E(G)$, let $\omega(f)$ denote the number of parallels of f in F. Then $\omega(f)$ is a perfect

2-matching of G in which the line e has positive weight. Thus each line of G occurs in some perfect 2-matching.

Conversely, assume that each line e of G is contained in a perfect 2-matching ω_e. For $f \in E(G)$, let $\omega(f) = \sum_e \omega_e(f)$. Replace every line f by $\omega(f)$ parallel lines. By assumption $\omega(f) > 0$ for every f, so at least one new line is substituted for every old line. The resulting graph is clearly $(2q)$-regular. This proves that G is regularizable. ∎

6.2.6. EXERCISE. Let xy be a line in graph G. Then the following inequalities hold:
(i) $\nu_2(G) \geq \nu_2(G - x - y) + 2$, and
(ii) $\nu_2(G) \geq \frac{1}{2}(\nu_2(G - x) + \nu_2(G - y)) + 1$.

6.2.7. EXERCISE. Let xy be a line in graph G. Prove that xy lies in some maximum 2-matching in G if and only if equality holds in at least one of (i) and (ii) in the preceding exercise.

6.2.8. EXERCISE. Show that regularizability of a graph can always be checked in polynomial time.

Our second characterization of regularizable graphs relates them to 2-bicritical graphs and elementary bipartite graphs.

6.2.9. THEOREM. *A graph G is regularizable if and only if every connected component of G is either an elementary bipartite graph or a 2-bicritical graph.*

PROOF. Assume first that G is regularizable. We may assume that G is connected. Suppose that G is not 2-bicritical. Then there exists a non-empty independent set A of points such that $|\Gamma(A)| \leq |A|$. (Of course by Theorem 6.2.5 and Corollary 6.1.5 we know equality holds here.) We may assume that A is a minimal set with this property, so that $|\Gamma(X)| > |X|$ holds for every non-empty proper subset X of A. Let G_1 be a regular graph obtained from G by multiplying its lines and let d be the degree of G_1. Then there are exactly $d|A|$ lines of G_1 incident with A. Each of these is also incident with $\Gamma(A)$, which is only possible if $|A| = |\Gamma(A)|$ and no further line is incident with $\Gamma(A)$. Hence G is bipartite with bipartition $(A, \Gamma(A))$. Since by hypothesis $|\Gamma(X)| > |X|$ holds for every non-empty proper subset X of A, it also follows by Theorem 4.1.1 that G is elementary.

Conversely, let G be a graph such that every connected component of G is either 2-bicritical or elementary bipartite. We want to show that

G is regularizable. It suffices to show that every connected component of G is regularizable, since from the regularizations of the components a regularization of G is easily constructed. (Just multiply every line by the product of degrees of the regularizations of the other components.)

So it suffices to show that 2-bicritical graphs and elementary bipartite graphs are regularizable. If G is 2-bicritical, then let ω_v be a perfect 2-matching of $G - v$ and let

$$f(e) = \left(\sum_{v \in V(G)} \omega_v(e) \deg(v) \right) + 2.$$

Then

$$\sum_{\{e | v_0 \in e\}} f(e) = \left(\sum_{v \neq v_0} 2 \deg(v) \right) + 2 \deg(v_0) = 4q,$$

and thus if we replace every line of G by $f(e)$ parallel lines, we obtain a regular graph. So G is regularizable.

If G is elementary bipartite then each line of G occurs in a perfect matching by Theorem 4.1.1. So each line of G occurs in a perfect 2-matching, and thus by Theorem 6.2.5, G is again regularizable. ∎

6.3. Matchings, 2-matchings and the König Property

We have seen in the previous sections that 2-matchings are much easier to handle than matchings. So it is sensible to try to apply the theory of 2-matchings to the theory of matchings. A natural question which arises is the following. Assuming that we have found a maximum 2-matching of a graph G, does this help in any way to find a maximum matching? This problem was addressed by Balas (1981), but results remain far from complete.

A natural approach to try is the following. Suppose ω is a basic 2-matching. Then we can construct a matching by taking those lines which have weight 2, together with a near-perfect matching from every odd cycle formed by the lines with weight 1. We shall say that the matching M resulting in this way is "contained in the 2-matching ω".

Is every matching M contained in a 2-matching? The answer is trivially *yes;* just take every line of M with weight 2. Is every matching M contained in a *maximum* 2-matching? The answer is no, even if M is maximum as is shown by the graph in Figure 6.3.1.

A non-trivial assertion about the connection of matchings and 2-matchings is the following result (Lovász (1979c, problem 7.3.8), Balas (1981)).

6.3.1. THEOREM. *Every graph has a maximum matching which is contained in a maximum 2-matching.*

PROOF. Consider the Gallai-Edmonds decomposition of the givengraph G. Let U be the set of those components of the subgraph spanned by $D(G)$ which are singletons, and let $\delta_U = \max\{|X| - |\Gamma(X)| \mid X \subseteq U\}$. Let G' denote the bipartite graph obtained from G by deleting the points in $C(G)$ and the lines spanned by $A(G)$, and by shrinking every connected component of $D(G)$ to a single point.

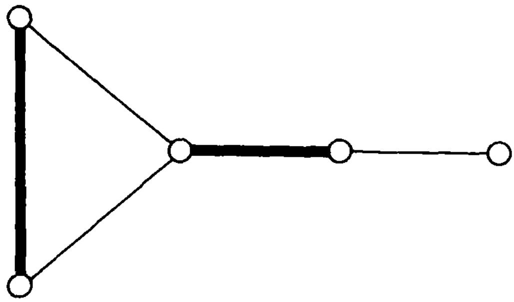


FIGURE 6.3.1.

Then applying Theorem 1.3.1 to the subgraph of G' induced by $U \cup A(G)$, we find that this subgraph contains a matching of size $|U| - \delta_U$. Thus G' has a matching which covers $|U| - \delta_U$ points of U. By Corollary 3.1.5, G' has a maximum matching which also covers at least $|U| - \delta_U$ points of U. Let M' be such a matching in G'. By statement (c) of the Gallai-Edmonds Structure Theorem, $|M'| = |A(G)|$. Let M be the matching in G corresponding to M'.

Now we can construct the maximum matching and 2-matching as claimed. Let M_0 be a perfect matching of the subgraph of G spanned by $C(G)$. Let $G_1, \ldots, G_t$ be the connected components of the subgraph spanned by $D(G)$, and let v_i be the endpoint of the line of M incident with G_i, if any such line exists; otherwise, let $v_i \in V(G_i)$ be arbitrary. Let M_i be a perfect matching of $G_i - v_i$. Let $N = M \cup M_0 \cup M_1 \cup \cdots \cup M_t$. Then, as we have seen in the proof of the Gallai-Edmonds Structure Theorem, the matching N is a maximum matching in G.

To construct a 2-matching ω containing N, we proceed as follows. For every component G_i, such that no line of M meets G_i and $|V(G_i)| > 1$, consider an odd cycle C_i which goes through v_i and which meets M_i in a near-perfect matching of C_i. (Such an odd cycle exists by the construction in Theorem 5.5.3.) Let the lines of the cycles C_i have weight 1, the other lines of N have weight 2, and all other lines of G have weight 0. We claim this 2-matching is maximum. In fact, the sum of weights of lines incident with any point is 2, except for δ_U points in U, where this sum is 0. So

$$|\omega| = p - \delta_U.$$

Since by Theorem 6.1.4 and Lemma 6.1.3, $\nu_2 = \tau_2 \leq p - \delta_U$, it follows that ω is indeed maximum. Since ω contains N, the proof is complete.■

6.3.2. COROLLARY. *Let $\pi(G)$ denote the minimum number of odd cycles in any maximum 2-matching of G. Then*

$$\nu(G) = (\nu_2(G) - \pi(G))/2.$$

■

We derive one more simple relation between matchings, 2-matchings and point covers.

6.3.3. THEOREM. *For every graph G,*

$$\nu_2(G) \leq \nu(G) + \tau(G).$$

PROOF. Let T be a minimum point cover of G, and let G' denote the bipartite subgraph of G formed by those lines which connect T to $V(G) - T$. Let S be a minimum point cover of G'. Then since G' is bipartite, we have $|S| = \nu(G')$. Furthermore, the sum of T and S is a 2-cover of G and hence

$$\nu_2(G) = \tau_2(G) \leq |S| + |T| = \tau(G) + \nu(G') \leq \tau(G) + \nu(G).$$

■

At this point we may return to a question posed in Chapter 1: which graphs besides those which are bipartite satisfy the relation $\nu(G) = \tau(G)$? Let us say that such a graph has the **König Property**. What makes this question interesting is that the computation of $\tau(G)$ is in general a difficult (i.e., NP-complete) problem (cf. Box 6A), while $\nu(G)$ can be computed in polynomial time (cf. Chapter 9). It turns out that the more special problem of determining whether $\tau(G)$ equals its obvious lower bound $\nu(G)$ in a given graph G is polynomially solvable.

We saw earlier in Section 6.1 that we always have $\tau_2(G) \leq 2\tau(G)$. It is interesting to know that those graphs for which equality holds are precisely those with the König Property.

The proof of this result follows immediately from Lemma 6.3.3 and Theorem 6.1.4.

6.3.4. COROLLARY. *A graph has the König Property if and only if it satisfies* $\tau_2(G) = 2\tau(G)$. ∎

We hasten to point out, however, that Corollary 6.3.4 does *not* provide a good characterization of the König Property! It is easy to show that a graph G has the König Property (if it does); just exhibit a matching M and a point cover T such that $|M| = |T|$. But if G does not satisfy the König Property, how do we exhibit this? Corollary 6.3.4 does not help in this respect. However, recently a good characterization of the König Property was found independently by Deming (1979) and Sterboul (1979). A refinement of this characterization was given more recently by Korach (1982). Here we shall state two other good characterizations which follow more directly from our preceding studies.

6.3.5. THEOREM. *A graph G has the König Property if and only if those lines of G which belong to at least one maximum 2-matching form a bipartite graph.*

PROOF. Suppose first that G has the König Property and let T be ε minimum point cover of G. Let ω be any maximum 2-matching in G Then $|\omega| = 2|T|$, and hence a simple computation shows that every lin‹ which occurs with positive weight in ω must connect T to $V(G) - T$. S‹ the bipartite graph formed by those lines of G connecting T to $V(G) - 7$ contains all lines which occur in any maximum 2-matching, and therefore the lines occurring in maximum 2-matchings form a bipartite graph.

Assume now that those lines of G which occur in maximum 2-matchings form a bipartite graph. We may assume that the theorem is valid for every graph with fewer points than G.

Case 1. Suppose G does not have a perfect 2-matching. Let $d = p - \tau_2(G)$. By Lemma 6.1.3, G contains an independent set X of points such that $|\Gamma(X)| = |X| - d$. Notice that every maximum 2-matching of G must contain $2(|X| - d)$ lines connecting X to $\Gamma(X)$ (with multiplicity allowed). Thus it contains no line induced by $\Gamma(X)$ or connecting $\Gamma(X)$ to $V(G) - X - \Gamma(X)$. Let $H = G - X - \Gamma(X)$. It follows that every maximum 2-matching of G contains a perfect 2-matching of H, and conversely, every perfect 2-matching of H is contained in a maximum 2-matching of G. Hence those lines of H which occur in a perfect 2-matching of H form a bipartite graph. So by the induction hypothesis, $\nu(H) = \tau(H) = \tau_2(H)/2 = |V(H)|/2$. If we let T be a minimum point cover of H, then $T \cup \Gamma(X)$ is a point cover of G of size $|T| + |\Gamma(X)| \leq |V(H)|/2 + |\Gamma(X)| = (|V(G)| + |\Gamma(X)| - |X|)/2 = (|V(G)| - d)/2 = \tau_2(G)/2$. So $\tau(G) \leq \tau_2(G)/2$ and hence $\tau(G) = \tau_2(G)/2$. By Corollary 6.3.4 this implies that G has the König Property.

Case 2. Suppose G has a perfect 2-matching, but is not 2-bicritical. Then by Lemma 6.2.1, there exists a non-empty independent set X of points of G such that $|\Gamma(X)| = |X|$. The proof in this case is identical with the proof in Case 1 (with $d = 0$).

Case 3. Suppose G is 2-bicritical. Then by Theorem 6.2.9, G is regularizable and so by Theorem 6.2.5, each line of G occurs in a perfect 2-matching. Thus by hypothesis G is bipartite. But then G cannot be 2-bicritical, a contradiction. ■

Theorem 6.3.5, together with Exercise 6.2.7, shows that we can check the König Property in polynomial time.

Before stating the next theorem concerning the König Property, we prove a simple lemma which will allow some simplification in the formulation of this theorem.

6.3.6. LEMMA. *A graph G has the König Property if and only if $D(G)$ is an independent set and the subgraph H induced by $C(G)$ has the König Property.*

PROOF. Assume first that G has the König Property. Let T be any minimum point cover of G. Then for each maximum matching M of G, every line of the matching is incident with one point of T and vice versa. Hence $T \subseteq A(G) \cup C(G)$ by the definition of the Gallai-Edmonds decomposition. So $D(G)$ cannot induce any line. Moreover, $T \cap C(G) = 1/2|C(G)|$ by a simple calculation and so $G[C(G)]$ also has the König Property.

Conversely, if H has the König Property then it has a point cover T consisting of $|V(H)|/2$ elements. If, in addition, $D(G)$ is independent, then $T \cup A(G)$ is a point cover of G of size $\nu(G)$, and so G has the König Property. ■

This lemma implies we need only concern ourselves with those graphs having perfect matchings.

Before stating our next result, recall that a subgraph H of a graph G is **nice** if $G - V(H)$ has a perfect matching.

6.3.7. THEOREM. *Let G be a graph with a perfect matching. Then G has the König Property if and only if it does not contain a nice subgraph which is an even subdivision of one of the two graphs in Figure* 6.3.2.

PROOF. It is straightforward to verify that if G contains an even subdivision of either one of the graphs in Figure 6.3.2 as a nice subgraph then $\tau(G) > p/2$ and so G cannot have the König Property.

Conversely, assume that G does not contain any even subdivision of the graphs in Figure 6.3.2 as a nice subgraph. We may assume that G is connected, since if every connected component of G has the König Property, then so does G.

Assume first that G has a perfect 2-matching which contains an odd cycle. Let ω be such a 2-matching with a minimum number of odd cycles. We may assume that ω is basic. Let M_0 be the matching formed by those lines having weight 2 in ω, and let M_1 be a perfect matching in G. Then $M_0 \cup M_1$ consists of the lines of $M_0 \cap M_1$, some alternating cycles and some paths connecting points in odd cycles of ω. These paths start and end with lines of M_1. At least one such path must connect points in different odd cycles in ω, since they cannot pair off the odd number of points on one cycle. So let P be a path-component of $M_0 \cup M_1$ connecting the odd cycles C_1 and C_2 of ω. If C_1 and C_2 are the only odd cycles of ω, then $P \cup C_1 \cup C_2$ is an even subdivision of the first graph in Figure 6.3.2, and it is also a nice subgraph of G. This contradicts the hypothesis of the theorem. If ω has more than two odd cycles then, deleting the lines of C_1 and C_2 from ω and adding the lines of a perfect matching of $C_1 \cup C_2 \cup P$ with weight 2, we obtain a perfect 2-matching which has fewer odd cycles than ω, but still at least one. This contradicts the choice of ω. Thus we have proved that no perfect 2-matching of G contains an odd cycle.

Let G_1 be the subgraph of G formed by those lines which occur in perfect 2-matchings. Since by the above argument every perfect 2-matching is the sum of two perfect matchings, G_1 is the graph formed by those lines which occur in perfect matchings. Hence every connected component of G_1 is an elementary graph. If G_1 has a connected component G_2 which is non-bipartite, then G_2 contains an even subdivision of K_4 or of R_3 as a nice subgraph, by Theorem 5.4.11. But then this even subdivision of K_4 or R_3 is a nice subgraph of G as well, which contradicts the hypotheses. So G_1 is bipartite. By Theorem 6.3.5, it follows that G has the König Property. ∎

Let us conclude with the remark that the study of the König Property in this section relates to the problem of characterizing τ-critical graphs, which will be discussed in Chapter 12. (See also Lovász (1982c).)

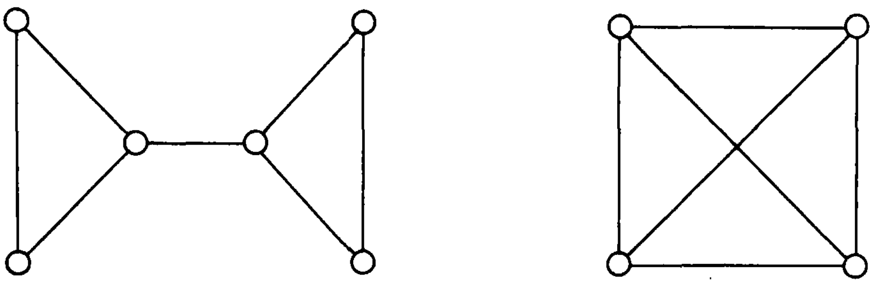


FIGURE 6.3.2. Excluded subgraphs for the König Property

BOX 6A. Reducibility Problems and NP-completeness

Why do we want to characterize graphs with the König Property? Would it not be more natural to seek a minimax formula for $\tau(G)$ (i.e., some kind of expression $\tau(G) = \max t$), analogous to the minimax formula (namely Berge's formula) for $\nu(G)$? If we could find such a formula for $\tau(G)$ then the relation $\tau(G) = \nu(G)$ would be equivalent to $\max t = \min(p - c_0(G - X) + |X|)/2$. Since the relation $\leq$ is always valid, this would be equivalent to saying that for every $X \subseteq V(G)$ and every feasible value of t, $2t \leq p - c_0(G - X) + |X|$. This would clearly yield a good characterization of the relation $\tau(G) = \nu(G)$.

The trouble is that no such minimax formula for $\tau(G)$ is known! Even worse, there is evidence that probably no such minimax formula exists at all! This assertion has not been proved, but the theory of *NP-completeness* seems to indicate that it is true. (The reader may judge for himself how strong he feels this evidence is. There are prominent mathematicians who believe that a good characterization of $\tau(G)$ *can* be found in spite of the results discussed below.)

We have already introduced the class of NP-properties (Box 1A) and mentioned their connection to polynomial algorithms (Box 1B). The next important observation in this area is that various NP-problems are reducible one to the other. By this we mean (somewhat loosely) that there is a polynomial algorithm which constructs, for every instance of problem A, an instance of problem B such that the answer to this instance of problem A is "yes" if and only if it is "yes" to the corresponding instance of problem B. There are various ways to make this notion more precise, but this does not change the fundamental fact that such a reduction does indeed reduce the solution of problem A to the solution of problem B. Such reductions are used quite frequently and we have already proved several theorems in this book this way. For example, we reduced the bipartite f-factor problem to a flow problem in Section 1.3 and we reduced the 2-matching problem for general graphs to the matching problem for bigraphs in Section 6.1. It is clear that if problem A can be reduced to problem B, then problem B cannot be essentially simpler than problem A. For example, if there exists a polynomial algorithm to solve problem B then there must also exist one to solve problem A.

It is a very surprising fact, discovered by S. Cook (1971) and Levin (1973) that there exist NP-problems to which *every* other NP-problem can be reduced. Such an NP-problem is called an **NP-complete problem.** (Sometimes it is also called a **universal search problem**). Historically the first problem to be proved NP-complete was SATISFIABILITY, defined in Box 1A. It is even more surprising that many "everyday" problems in graph theory are NP-complete. The first and most important such problems were discovered by Karp (1972, 1975), but since then the list of such problems has grown to several hundred. (See Garey and Johnson (1979).) For example, it is known that the following problems are NP-complete: (1) 3-colorability of a graph and, more generally, k-colorability of graph for every $k \geq 3$; (2) the property that a graph has k independent points; (There is a difference between this and k-colorability: k-colorability is NP-complete for every particular $k \geq 3$; on the other hand, $\alpha(G) \geq k$ can be trivially checked in polynomial (i.e., in $O(n^k)$) time for every particular value of k. Thus the latter problem is only NP-complete as a property of the pair (G, k).) (3) the existence of a Hamilton cycle and (4) 2-colorability of a hypergraph.

We should also mention that there are problems to which all problems in NP may be polynomially reduced, but whose membership in NP is not settled. These are known as **NP-hard** problems. (Actually, we shall be a bit more informal with this notion and simply call a problem NP-hard if all problems in NP polynomially reduce to it.) For a more precise treatment of these terms, as well as an interesting history of the terminology in complexity theory, again we send the reader in search of Garey and Johnson (1979).

It is generally (but not universally!) believed that no NP-complete problem can be solved in polynomial time or has a good characterization. If any one of them had such, then — using the reductions which show that it is NP-complete — we could obtain a polynomial time algorithm and/or a good characterization for *every* NP-problem automatically. Considering that the existing algorithms for (non-complete) NP-problems are based on very different and often quite ingenious ideas, it seems unlikely that one solution would work for every NP-problem.

The problem of deciding whether or not $\tau(G) = k$ for arbitrary k is known to be NP-complete. This justifies the fact that in the previous section we were content with solving a special case of it, namely when $k = \nu(G)$.

Let us hasten to point out that we do *not* share the point of view (which some people working in NP-completeness theory are often accused of having) that if a problem is shown NP-complete, then there is no point in working on it any longer! Quite the contrary, NP-completeness of a problem implies that (probably) no complete solution can be found (in the sense in which the König-Hall Theorem and the Hungarian Method solve the bipartite matching problem). So alternate approaches such as finding relatively large special classes for which the problem can be solved, finding approximate solutions and estimates as well as speeding up non-polynomial algorithms by heuristics, clever data structures, and so on, gain even more importance. In Chapter 12 we shall survey some interesting results in this vein on vertex packing (or equivalently, point covering).

6.4. Hamilton Cycles and 2-matchings

A Hamilton cycle is a special perfect 2-matching. Can the theory of 2-matchings help decide whether or not a graph has a Hamilton cycle? Since the latter is a difficult (NP-complete) problem, no complete solution can be expected. But recent results of Cornuéjols and Pulleyblank (1983) discussed below do shed some light in this connection. They also yield a nice unification of 1-matching and 2-matching theory, which is simpler and more elegant than the general f-factor theory which we shall treat in Chapter 10.

Let K be any set of natural numbers. We say that a 2-matching ω in a graph G is **K-gon-free**, if G does not contain any k-gon with $k \in K$ in which all lines have positive weight. We shall be interested in the problem of deciding which graphs have a K-gon-free perfect 2-matching. Let us point out immediately that only the odd numbers in K matter; the even cycles can be eliminated from any perfect 2-matching without creating any new cycles.

If $K = Z_+$, the set of all positive integers, then a K-gon-free perfect 2-matching is just a perfect matching with every line taken twice. So this problem includes the matching problem. For $K = \emptyset$, the problem includes the 2-matching problem. An important special case is when $K = \{1, 2, \ldots, p-1\}$ where $p = |V(G)|$ is odd. Then a K-gon-free 2-matching is just a Hamilton cycle. It is easy to see that the Hamilton cycle problem is NP-hard, even for graphs with odd cardinality.

A connected graph G will be called **K-gon-critical**, if it does not contain a K-gon-free perfect 2-matching, but $G - v$ has a K-gon-free perfect 2-matching for every $v \in V(G)$. If $K = Z_+$ then Z_+-gon-critical graphs are just the factor-critical graphs. If $K = \emptyset$, then the one-point graph is $\emptyset$-gon-critical, and there are no other $\emptyset$-gon-critical graphs. In fact, a $\emptyset$-gon-critical graph with at least two points would be 2-bicritical, but then it has a perfect 2-matching itself by Corollary 6.2.2, contradicting the definition of K-gon-critical. So such components are K-gon-critical.

The following theorem of Cornuéjols and Pulleyblank (1983) is a common generalization of Tutte's Theorems 3.1.1 and 6.1.5.

6.4.1. THEOREM. *A graph G has a perfect K-gon-free 2-matching if and only if for every $X \subseteq V(G)$, the number of K-gon-critical components of $G-X$ is at most $|X|$.*

PROOF. The necessity of the condition on the right hand side is easy; it follows in the usual way by counting lines from $G-X$ to X. To prove sufficiency, consider the Gallai-Edmonds decomposition of G, and the bipartite graph G' introduced in part (c) of the Gallai-Edmonds Structure Theorem 3.2.1. Let U be the set of those points of G' which are the shrunken copies of those components of $D(G)$ which have no K-gon-free perfect 2-matching. Let us point out immediately that if G_i is such a component then $G_i - v$ has a K-gon-free (even Z_+-gon-free) perfect 2-matching for every $v \in V(G)$, since G_i is factor-critical.

We claim that G' has a matching which covers U. Suppose not. Then by P. Hall's Theorem there is a set $W \subseteq U$ such that $|\Gamma_{G'}(W)| < |W|$. But then $G - \Gamma_{G'}(W)$ has at least $|W| > |\Gamma_{G'}(W)|$ K-gon-critical components, which is a contradiction of the assumption.

So G' has a matching covering U. Thus by Corollary 3.1.5, it also has a maximum matching M which covers U. By property (c) of the Gallai-Edmonds Theorem, M covers every point of $A(G)$. Let M_0 be a perfect matching of the graph induced by $C(G)$. Let $G_1, \ldots, G_k$ be those connected components of $D(G)$ which are met by M, and $G_{k+1}, \ldots, G_t$ those which are not. Let v_i be the point of G_i covered by M $(i = 1, \ldots, k)$ and let M_i be a perfect matching of $G_i - v_i$. Finally, let ω_i be a K-gon-free perfect 2-matching of G_i for $i = k+1, \ldots, t$. Then taking the lines in $M \cup M_0 \cup M_1 \cup \cdots \cup M_k$ with weight 2, together with the 2-matching $\omega_{k+1}, \ldots, \omega_t$, we obtain a K-gon-free perfect 2-matching of G. ∎

This theorem does not quite yield a good characterization of graphs *not* having a K-gon-free perfect 2-matching. To exhibit that a graph does not have such a 2-matching, we would have to prove that more than $|X|$ components of $G-X$ are K-gon-critical for some $X \subseteq V(G)$. But how can we check (or at least exhibit) that a component is K-gon-critical? Although there does not exist any general method to do this, some results on K-gon-critical graphs, as well as characterizations of certain special classes of such graphs, have been obtained by Cornuéjols and Pulleyblank (1983) and by Cornuéjols, Hartvigsen and Pulleyblank (1982).

6.4.2. THEOREM. *A graph G is K-gon-critical if and only if it is factor-critical and $G - V(C)$ has no perfect matching for any cycle $C \subseteq G$, where $|V(C)| \notin K$.*

PROOF. The proof of this theorem is very similar to the proof of Theorem 6.4.1, and is left to the reader. ∎

We mentioned above that the structure of K-gon-critical graphs is known when $K = Z_+$ (then they are just the factor-critical graphs) and when $K = \emptyset$ (the only one is the singleton). Note that Theorem 6.4.2 gives a good characterization of these graphs, if $Z_+ - K$ is finite. A further case which has been solved is when $K = \{3\}$ (Cornuéjols and Pulleyblank (1980a, 1982)).

6.4.3. THEOREM. *A connected graph G is $\{3\}$-gon-critical if and only if every block of G is a triangle.*

PROOF. It is immediate that if every block of a connected graph is a triangle, then the graph does not contain a triangle-free 2-matching. Since,
furthermore, it is obviously factor-critical, it is also $\{3\}$-gon-critical.

Assume now that G is $\{3\}$-gon-critical. Then by Theorem 6.4.2, it is factor-critical and so by Lemma 5.2.9, every block of G is factor-critical. Suppose, to the contrary, that G has a block B with more than three points. Consider the ear representation of this block B (Theorem 5.5.2). We may assume that it starts with an odd cycle longer than 3, since if the first cycle is a triangle, then there must be a second ear and so we could have started with the odd cycle consisting of this ear and two lines of the triangle.

So let C be the odd cycle of length greater than 3, which starts the ear representation. Then C is a nice subgraph of B, and hence of G. Taking the lines of C with weight 1 and the lines of a perfect matching of $G - V(C)$ with weight 2, we obtain a triangle-free 2-matching. Thus G is not $\{3\}$-gon-critical, a contradiction. ∎

6.4.4. EXERCISE. Give a good characterization of those graphs the points of which can be partitioned into complete graphs with at least two points.

One might argue that it is more natural to ask about the existence of K-gon-free 2-factors; more exactly, to view lines with weight 2 as 2-gons and so exclude these if $2 \in K$. This approach, however, leads to much more difficult problems which to date remain mostly intractable. For example, Papadimitriou proved that the existence of a $\{3,5\}$-gon-free 2-factor is an NP-complete problem. For a proof of this result, see Cornuéjols and Pulleyblank (1980b).

A further generalization of these problems is the following. Let $\mathcal{H}$ be a family of connected graphs and let G be any graph. An **$\mathcal{H}$-factor** of G is a spanning subgraph H of G, such that each component of H is isomorphic to a member of $\mathcal{H}$. The problem is then to decide if G has an $\mathcal{H}$-factor. Every graph has a $\{K_1\}$-factor, so assume that $K_1 \notin \mathcal{H}$. Of course if $\mathcal{H} = \{K_2\}$, we just have the ordinary perfect matching problem. It has recently been shown, however, that if $\mathcal{H}$ is any other single element family, then the $\mathcal{H}$-factor problem is NP-complete. (See Kirkpatrick and Hell (1978, 1983).) On the other hand, the problem is polynomial in the case where $\mathcal{H} = \{K_2\} \cup \mathcal{H}'$, where $\mathcal{H}'$ is any finite collection of factor-critical graphs. (See Cornuéjols, Hartvigsen and Pulleyblank (1982) and Hell and Kirkpatrick (1983).)

6.5. The Chinese Postman Problem

Let us suppose a postman on his daily route must pass along every street in a village and then return to his starting point. What route should he follow in order to minimize the total distance traversed?

In spite of the striking similarity between this problem (called the Chinese Postman Problem) and the Travelling Salesman Problem, the Postman Problem is easier to deal with and is closely related to matching theory. (Incidently, the "Chinese" adjective is present because this problem was first raised by the Chinese mathematician Meigu Guan (1962).)

Let us start by observing that if at every street intersection an even number of streets meet; that is, if the corresponding graph G is Eulerian, then the postman may follow an Eulerian trail and pass every line exactly once. Since he has to pass every line at least once, this route is certainly optimal. So in this case the minimum length of a route is equal to the number of lines.

Suppose now that the graph G is not Eulerian. We shall assume throughout for obvious reasons that it is connected. Let us consider a **Chinese Postman Tour**, that is, a closed walk which uses every line. Replace every line by a set of parallel copies equal in cardinality to the

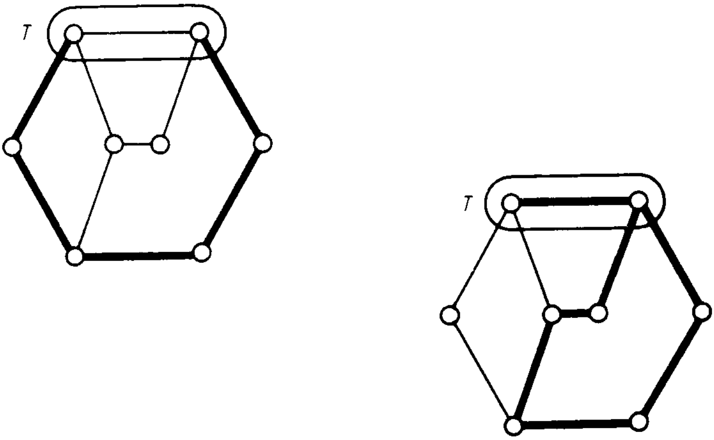


FIGURE 6.5.1. Two T-joins

number of times the Chinese Postman Tour uses the line. This results in an Eulerian graph. Conversely, if an Eulerian graph arises from G by multiplying lines, then an Eulerian trail of this graph yields a Chinese Postman Tour of the original.

It is then clear that an optimal Chinese Postman Tour does not use any line more than twice, for if it did, then from the corresponding Eulerian graph two copies of this line could be removed. But then an Eulerian trail in the resulting graph would correspond to a shorter Chinese Postman Tour in the original G, contradicting optimality.

So the Chinese Postman Problem is equivalent to the following: given a graph G, find the minimum number of lines whose doubling results in an Eulerian graph. If we let H be the subgraph formed by these lines, we must have

$$\deg_H(x) \equiv \deg_G(x) \pmod 2 \tag{6.5.1}$$

for every point x, and conversely, if a subgraph H satisfies (6.5.1), then doubling the lines in H results in a graph all components of which are Eulerian.

This problem may be generalized somewhat further by arbitrarily specifying the parity of $\deg_H(x)$ at the points of G: More exactly, let $T \subseteq V(G)$ be given with $|T|$ even. A set $H \subseteq E(G)$ is called a T**-join**, if

$$\deg_H(x) \equiv \begin{cases} 1 \pmod 2, & \text{if } x \in T \\ 0 \pmod 2, & \text{if } x \in V(G) - T. \end{cases}$$

The problem then is to find a T-join having minimum cardinality.

See Figure 6.5.1 for some examples. In each drawing the lines of H are bold-face.

Now again let T be an even subset of $V(G)$ and let $S \subseteq V(G)$ be arbitrary. The set of lines $C = \nabla(S)$ is a T-cut provided $S \cap T$ is odd. The sets S and $V(G) - S$ are called the **shores** of C. Obviously the star of a point in T is a T-cut; such T-cuts will be called **trivial**. All other T-cuts will be **non-trivial**. We will also call a $V(G)$-cut an **odd cut**.

6.5.1. EXAMPLE. let $T = \{u, v\}$. Then the T-cuts are just the cuts separating u and v. The minimal T-joins are precisely the $u - v$ paths. Every other T-join arises as the disjoint union of a $u - v$ path and an Eulerian subgraph.

6.5.2. EXAMPLE. Let $T = \{u, v, w, t\}$. Then a minimal T-join is either the union of two point-disjoint paths pairing up the four points of T, or the union of three openly disjoint paths connecting one of u, v, w, t to the other three, or the union of four openly disjoint paths joining a fifth point to u, v, w and t, respectively.

6.5.3. EXAMPLE. Let T consist of those points of G which have odd degree in G. Then G itself is a T-join and every other T-join can be obtained by deleting the lines of an Eulerian subgraph. The T-cuts are exactly those cuts which contain an odd number of lines.

6.5.4. EXAMPLE. Let $T = V(G)$. Then every perfect matching of G is a T-join, and these are the only T-joins with exactly $p/2$ lines. There may, however, exist other minimal (but not *minimum*) T-joins.

6.5.5. EXERCISE. Prove that in a simple graph every minimum T-join is a forest, such that every component of the forest is an induced subgraph. From this, prove Sumner's and Las Vergnas' Theorem 3.3.20.

6.5.6. EXERCISE. Let J be a T-join and C, a cut which intersects J in an odd number of lines. Prove that C is a T-cut.

Example 6.5.4 above shows that the matching problem is a special case of the problem of finding a minimum T-join. Conversely, Edmonds and Johnson (1973) showed that the T-join problem can be reduced to the *weighted* matching problem, which we shall study in Chapter 7 and from an algorithmic point of view in Chapter 9. The idea of this reduction is the following. For every pair of points u, v in T compute the distance $d(u, v)$ in G. Consider the complete graph H with point set T, with the lines of H weighted by the corresponding $d(u,v)$. Let M be a minimum weight perfect matching in H and, for each line $uv \in M$, let P_{uv} be a minimum length $u-v$ path in G. Then it is easy to prove that the P_{uv}'s are mutually line-disjoint and hence that $\bigcup_{uv \in M} P_{uv}$ is a minimum T-join. So to solve the Chinese Postman Problem, we take T to be the set of all points of odd degree, solve a shortest path problem for all pairs of points in T, and then solve a minimum weight perfect matching problem for the complete graph on T with weights equal to the distances computed.

A different reduction of the Chinese Postman Problem to a matching problem is described in Schrijver (1983a). Here, however, we do not use either one of these reductions, but we treat the T-join problem more directly, following Lovász (1975a) and Seymour (1981b).

It is interesting to point out that the weighted T-join problem, when every line has a (non-negative integral) length, is not essentially more general than the unweighted problem, since we may always replace a line having weight k with a path of length k, whose interior points are not in T. This constitutes a significant difference from the matching problem, where the weighted version turns out to be fundamentally more general (in the sense that no similar simple construction is known for reducing the weighted version to the unweighted). This treatment provides additional information on the "dual" problem; that is, it gives a simpler form of the "max" side of the minimax theorem characterizing the minimum length of a T-join.

We shall need to perform some manipulations with pairs (G,T), where T is any even subset of $V(G)$. First, if we say that we **subdivide** a line of G, we shall mean that the new point will not be in T; that is, T remains unchanged. Now let xy be a line in G. Contract xy in the usual way to get a graph G', but modify T as follows. The new point will belong to the new set T' if and only if exactly one of x and y belong to T. We shall say that the pair (G',T') arises from the pair (G,T) by **contraction** of xy. It is straightforward to check that T' is even and that the T'-cuts of G' correspond in the natural way to those T-cuts of G which do not contain the line xy.

The behavior of *T-joins* under contraction is only slightly more complicated. If J is any T-join, then its image under this contraction of xy is a T'-join. Conversely, let J' be a T'-join in G'. The lines in J' form a subgraph of G which satisfies the degree constraints for T-joins, with the possible exception of the two points x and y. Furthermore, either both x and y satisfy the degree constraints, or both violate them, by reasons of parity. Hence either J' or $J' + xy$ is a T-join.

One more preliminary observation we need is the fact that the "uncrossing procedure" used in Section 2.3 can be applied to T-cuts. To be more precise, recall that two cuts $C_1 = \nabla(X_1)$ and $C_2 = \nabla(X_2)$ are called **crossing** if each of the four sets $X_1 \cap X_2$, $X_1 - X_2$, $X_2 - X_1$ and $V(G) - X_1 - X_2$ is non-empty. The following procedure can be applied to crossing cuts.

UNCROSSING PROCEDURE. Let C_1 and C_2 be two crossing T-cuts. Let S_1 be either shore of C_1. Then $S_1 \cap T$ is odd by definition and C_2 splits S_1 into two sets, one of which again meets T in an odd set. Let S_2 be the shore of C_2 for which $S_1 \cap S_2 \cap T$ is odd. It follows then that $(S_1 \cup S_2) \cap T$ is also odd. Let C' and C'' be the T-cuts determined by the sets $S_1 \cap S_2$ and $S_1 \cup S_2$. The procedure just outlined is called the **uncrossing** of C_1 and C_2.

Besides being non-crossing, the pair C', C'' has the important property that every line of G belongs to at most as many of the cuts C', C'' as to the cuts C_1, C_2. More precisely, the lines between $S_1 - S_2$ and $S_2 - S_1$ belong to both C_1 and C_2, but not to C' or C''. Every other line is contained in the same number of the cuts C_1, C_2 as of the cuts C', C''.

Note too that if any cut C_3 crosses one of C' and C'', then it also crosses at least one of C_1 and C_2, and if C_3 crosses both C' and C'', then it crosses both C_1 and C_2. Roughly speaking, uncrossing does not introduce new crossings with other cuts.

6.5.7. LEMMA. *Let C be a T-cut and J be a T-join. Then $J \cap C$ is odd.*

PROOF. The sum of degrees of J on one shore S of C contains an odd number of odd terms, and so it is odd. This sum is equal to $|J \cap C|$ + twice the number of lines of J spanned by S. Hence $J \cap C$ must be odd. ■

6.5.8. LEMMA. *If G is a connected graph, then for every even subset $T \subseteq V(G)$ there exists a T-join in G.*

PROOF. Let $T = \{t_1, \ldots, t_{2k}\}$, and let P_i be a path connecting t_i to t_{k+i} $(i = 1, \ldots, k)$. Let J be the symmetric difference (mod 2 sum) of $E(P_1), \ldots, E(P_k)$. Then J is a T-join. ■

6.5.9. LEMMA. *A set of lines meets every T-join if and only if it contains a T-cut. A set of lines meets every T-cut if and only if it contains a T-join.*

PROOF. A moment's reflection here shows that the second assertion is just a reformulation of the first.

The "if" part of the first assertion is trivial by Lemma 6.5.8.

Conversely, assume that X is a set of lines which meets every T-join. We want to show that X contains a T-cut. Let $G_1, \ldots, G_k$ be the connected components of $G-X$. It cannot happen that $V(G_i) \cap T$ is even for every i, for then each G_i contains a $T \cap V(G_i)$-join J_i by Lemma 6.5.7, and therefore $J = J_1 \cup \cdots \cup J_k$ is a T-join in $G-X$, contradicting the choice of X. But if $T \cap V(G_i)$ is odd for some i then the cut determined by $V(G_i)$ is a T-cut contained in X. ■

In view of Lemma 6.5.9, it is natural to relate the minimum size of T-joins to "packings" of T-cuts. A **k-packing** of T-cuts is a list $C_1, \ldots, C_k$ of T-cuts, such that every line of G belongs to at most k of them. We denote the maximum number of T-cuts in a k-packing by $\nu_k(G,T)$, and the minimum size of a T-join by $\tau(G,T)$. Furthermore, we set $\nu = \nu_1$.

The following theorem is due to Edmonds and Johnson (1973):

6.5.10. THEOREM. *For every graph G and every even subset T of $V(G)$, $\tau(G,T) = \nu_2(G,T)/2$.*

Several comments and results are in order prior to the proof of this theorem.

First of all, by the discussion at the beginning of this section, a good characterization of the minimum length of a Chinese Postman Tour follows, if we take T to be the set of all points of odd degree.

6.5.11. COROLLARY. (*The Chinese Postman Theorem*). *Let G be a connected graph and let m denote the maximum size of a 2-packing of odd cardinality cuts in G. Then the minimum length of a Chinese Postman Tour is $q + m/2$.* ■

The following sharper relation for bipartite graphs was proved by Seymour (1981a). Strangely enough, in contrast to the case of matchings, here the bipartite graph theorem is the deeper, and Theorem 6.5.10 will follow easily from Theorem 6.5.12.

6.5.12. THEOREM. *For every bipartite graph G and every even subset T of $V(G)$, $\tau(G,T) = \nu(G,T)$.*

The proof of Theorem 6.5.12, in turn, will depend upon the following result.

6.5.13. LEMMA. *If G is a bipartite graph and T an even subset of $V(G)$, then $\nu_2(G,T) = 2\nu(G,T)$.*

PROOF. We use induction on the number of points in G. The statement is trivial if this number is 0 or 1.

So let $\mathcal{L}$ be a maximum 2-packing of T-cuts in G and assume that the desired conclusion holds for all graphs having fewer points than G. By repeatedly uncrossing pairs of cuts in $\mathcal{L}$, we may replace $\mathcal{L}$ by another 2-packing $\mathcal{L}'$ of the same size in which no two cuts cross. (Innocent though it may seem, this procedure of uncrossing is the heart of the matter. This technique has applications to many other problems in graph theory, such as flow problems (see Section 2.3), connectivity problems, etc. See the next section and also Box 6B.)

Let $a \in V(G)$ and for each cut C, let $S(C)$ denote the shore of C not containing a. A cut C such that $|S(C)| = 1$ will be called a **star-cut**. Note that every star-cut is trivial, but the converse is not true, since the star of a is not necessarily a star-cut.

I. First assume that there exists a star-cut which occurs twice in the list $\mathcal{L}'$. Then contracting the lines of this star-cut, only two cuts of $\mathcal{L}'$ are destroyed. So the resulting pair (G',T') has $\nu_2(G',T') \geq \nu(G,T) - 2$. By the induction hypothesis, $\nu_2(G',T') = 2\nu(G',T')$. Furthermore, any 1-packing of T'-cuts in G' corresponds to a 1-packing of T-cuts in G, and this 1-packing can be augmented by the star-cut C. Thus

$$\nu_2(G,T) \leq \nu_2(G',T') + 2 = 2\nu(G',T') + 2 \leq 2\nu(G,T).$$

Since the reverse inequality is trivial, the assertion is proved for case I. So in what follows we shall assume that every star-cut occurs at most once.

II. Suppose each cut in $\mathcal{L}'$ is a star-cut. Let (A,B) be a 2-coloration of G and assume, without loss of generality, that $|A \cap T| \geq |B \cap T|$, so that $|A \cap T| \geq |T|/2$. The star-cuts determined by the points in $A \cap T$ are line-disjoint and so

$$\nu(G,T) \geq |A \cap T| \geq |T|/2 \geq |\mathcal{L}'|/2 = \nu_2(G,T)/2.$$

This proves the assertion in this case. So in the following we may assume that there exists a cut $C \in \mathcal{L}'$ which is not a star-cut. Choose such a C where the cardinality of $S(C)$ is as small as possible.

IIIa. Suppose every point of $T \cap S(C)$ is the center of a star-cut in $\mathcal{L}'$. Without loss of generality, we may assume that $|A \cap T \cap S(C)| \geq |B \cap T \cap S(C)|$, so that $|A \cap T \cap S(C)| \geq |T \cap S(C)|/2$. Since C is a T-cut, $|T \cap S(C)|$ is odd and so it also follows that $|A \cap T \cap S(C)| \geq (|T \cap S(C)| + 1)/2$.

Delete from $\mathcal{L}'$ the cut C and also all cuts C' such that $S(C') \subset S(C)$. By the choice of C, every such C' is a star-cut, and clearly its center must belong to T. Furthermore, every such cut occurs in the list exactly once, by assumption. So we have deleted exactly $|T \cap S(C)| + 1$ cuts from $\mathcal{L}'$.

Add to the collection the stars of points in $A \cap T \cap S(C)$, each with multiplicity 2. This results in another 2-packing (check this!), whose size is

$$|\mathcal{L}'| - (|T \cap S(C)| + 1) + 2|A \cap T \cap S(C)| \geq |\mathcal{L}'|.$$

Thus we are back to I and so case IIIa is also settled.

IIIb. So, finally, we may suppose that there exists a point $v \in T \cap S(C)$ whose star is not in $\mathcal{L}'$. Then delete C from $\mathcal{L}'$, but add the star of v. This results in a 2-packing (check again!) of the same size, also consisting of non-crossing cuts, which contains more star-cuts than $\mathcal{L}'$. Repeating this if necessary, we must eventually arrive at case II or IIIa.■

The preceding lemma is in a sense a "self-refining" result:

6.5.14. COROLLARY. *If G is a bipartite graph and T an even subset of $V(G)$, then for every $k \geq 1$, $\nu_k(G,T) = k\nu(G,T)$.*

PROOF. By Lemma 6.5.13, $\nu_2(G,T) = 2\nu(G,T)$. Applying this lemma to the subdivision of G, we get that $\nu_4(G,T) = 2\nu_2(G,T)$. Continuing in a similar way, we get that $\nu_{2^m}(G,T) = 2^m\nu(G,T)$ for every $m \geq 0$.

On the other hand, if $k \geq 1$ is arbitrary, then let m be such that $2^m > k$. Using the trivial inequality, $\nu_{a+b} \geq \nu_a + \nu_b$, repeatedly, we get

$$\begin{aligned}\nu_{2^m}(G,T) &\geq \nu_{2^m-k}(G,T) + \nu_k(G,T)\\ &\geq (2^m - k)\nu(G,T) + k\nu(G,T) = 2^m\nu(G,T).\end{aligned}$$

But we know that the first and last quantities are equal. So equality must hold throughout and, in particular, $\nu_k(G,T) = k\nu(G,T)$. ■

Lemma 6.5.13 implies that Theorems 6.5.10 and 6.5.12 are in fact equivalent, and so it would suffice to prove Theorem 6.5.10. We shall, however, give a proof of Theorem 6.5.12 instead, and in the course of this we shall use Corollary 6.5.14.

PROOF (of Theorem 6.5.12). We use induction on $|E(G)|$. If $E(G) = \emptyset$ the assertion is trivial. So suppose that $E(G) \neq \emptyset$. We may assume that no T-cut is empty, since then $\nu(G,T) = \tau(G,T) = \infty$, and so the assertion is true.

The inequality $\nu(G,T) \leq \tau(G,T)$ is a trivial consequence of Lemma 6.5.8, and so it suffices to prove that G contains a T-join with no more than $\nu = \nu(G,T)$ lines.

If G has a line e such that $\nu(G-e,T) = \nu(G,T)$, then by the induction hypothesis, $G-e$ has a T-join with ν lines and this T-join is also a T-join in G. So we may assume that $\nu(G-e,T) > \nu$ for every line e.

If G has a T-cut C such that $|C| = 1$, then contract the line in C. The resulting pair (G',T') has $\nu(G',T') \leq \nu - 1$, since every 1-packing of T-cuts in G' can be augmented by C. Thus by the induction hypothesis, G' contains a T'-join J' of size $\leq \nu - 1$, and so either J' or $J' \cup C$ is a T-join in G of size $\leq \nu$.

If G has a T-cut C such that $|C| = 2$, then we argue in a similar fashion. Contracting the two lines e and f of C in (G,T), we get a pair (G'',T'') such that $\nu(G'',T'') \leq \nu - 1$. So G'' has a T''-join J'' with at most $\nu - 1$ lines. It follows from our discussion of contraction, that one of the sets J'', $J'' + e$, $J'' + f$ and $J'' + e + f$ is a T-join. But the last of these is certainly not a T-join, since it meets the T-cut C in an even number of lines. Thus we have found a T-join with at most ν lines.

It remains to show — and this is the major part of the proof — that it cannot happen that every T-cut of G has at least three lines and yet $\nu(G-e,T) > \nu(G,T)$ for every line e. Suppose, to the contrary, that this is the case. We shall derive a contradiction.

Let e be any line of G. Then $G-e$ contains $\nu+1$ line-disjoint T-cuts. In G, these T-cuts are almost line-disjoint; only the line e is contained in some (at most $\nu+1$) of them. So it follows that if we subdivide e by more than ν points, the resulting graph has $\nu+1$ line-disjoint T-cuts.

Let us subdivide a line of G by two new points if we can do so without producing $\nu+1$ line-disjoint T-cuts, and repeat this as long as we can. Since, by the observation above, no line will be subdivided by more than ν points, it follows that in a finite number of steps we must stop with a graph G' having no $\nu+1$ line-disjoint T-cuts. Note G' is also bipartite and that every T-cut in G' has at least two lines.

By hypothesis, if we subdivide any line e of G' by two points we produce $\nu+1$ line-disjoint T-cuts. This means that G' contains $\nu+1$ T-cuts which have the property that every line other than e belongs to at most one of them, while e itself belongs to at most 3 of them. Let $\mathcal{L}_e$

be such a list of $\nu+1$ T-cuts of G', and form the list $\mathcal{L} = \sum_{e\in E(G')} \mathcal{L}_e$. (Note here that multiple membership of cuts is to be allowed.) Then $\mathcal{L}$ is an $(m+2)$-packing of T-cuts, where $m = |E(G')|$. Thus

$$\nu_{m+2}(G', T') \geq m(\nu+1).$$

By Corollary 6.5.14 we have

$$\nu_{m+2}(G', T') = (m+2)\nu,$$

and hence it follows that $m \leq 2\nu$. Thus

$$|E(G)| \leq m \leq 2\nu.$$

But G contains ν line-disjoint T-cuts each of which has, by hypothesis, at least 3 lines. This contradiction proves the theorem. ■

PROOF (of Theorem 6.5.10). Subdivide each line of G by one point and apply Theorem 6.5.12 to the resulting bipartite graph. ■

6.5.15. EXERCISE. Prove that if G is a 2-line-connected graph with q lines and $T \subseteq V(G)$ with $|T|$ odd, then it contains a T-join with at most $q/2$ lines.

6.5.16. EXERCISE. Prove that if G is a graph without loops and with q lines, then it contains a bipartite subgraph with at least $q/2$ lines (Erdős (1967)). What is the connection between this and the previous exercise?

6.5.17. EXERCISE. By Theorem 6.5.10, a graph G has a perfect matching if and only if it does not contain a 2-packing of more than p $V(G)$-cuts. Show that this condition is equivalent to that of Tutte. (This is lengthy! See Lovász (1975a)).

Just as for matchings and point covers, we may ask also in the case of T-cuts and T-joins: for which graphs, besides those which are bipartite, does the relation $\nu = \tau$ hold? A general result due to Seymour (1977) yields one such class:

6.5.18. THEOREM. *Let G be a graph and T an even subset of G such that (G,T) cannot be contracted onto $(K_4, V(K_4))$. Then $\nu(G,T) = \tau(G,T)$.* ■

Note that in the special case when $|T| = 2$, the condition of the preceding theorem is trivially fulfilled. Hence we obtain the well-known result that *the minimum length of a path joining two points u and v of a graph G is equal to the maximum number of line-disjoint cuts separating u and v.*

A graph is called **series-parallel** if it can be obtained from K_2 by the repeated application of two operations: subdivision ("series extension") and multiplication of lines ("parallel extension"). It is easy to see that series-parallel graphs cannot be contracted onto K_4. (This property in fact characterizes them.) So Seymour's theorem above implies:

6.5.19. COROLLARY. *For every series-parallel graph G and every even subset T of $V(G)$, $\nu(G,T) = \tau(G,T)$.* ∎

(Tardos has pointed out that this corollary also follows via an easy induction. We leave the details of this proof to the reader as an exercise.)

The class of series-parallel graphs is independent of the class of bipartite graphs, in the sense that neither of them contains the other. It is not known in general which graphs satisfy the relation $\nu(G,T) = \tau(G,T)$ for every even subset T.

We conclude this section with a discussion of the weighted version of the T-join problem. Let ω be a weighting of the lines of G by non-negative integers. An **ω-packing** of T-cuts (cycles, etc.) is a collection of T-cuts (cycles, etc.) such that every line e is contained in at most $\omega(e)$ of them. Let $\nu(G,T,\omega)$ denote the maximum cardinality of an ω-packing of T-cuts in G. Let $\tau(G,T,\omega)$ denote the minimum weight of a T-join. As has already been pointed out, the case when every weight is a natural number can be reduced to the unweighted case by subdividing the lines. A line with weight 0 can be simply contracted (check the details!). In this way Theorem 6.5.10 yields the following.

6.5.20. THEOREM. *For each graph G, every even subset T of $V(G)$ and every non-negative weighting ω of the lines of G,*

$$\tau(G,T,\omega) = \nu(G,T,2\omega)/2.$$ ∎

We leave it to the reader to formulate the corresponding generalization of Theorem 6.5.12.

A very nice feature of the T-join problem is that it is "self-refining" in the sense that the problem of finding a *maximum* T-join can be reduced to the problem of finding a *minimum* T-join. More generally, the problem of finding a minimum weight T-join can be solved even if negative weights are allowed. In particular, if all weights are -1, we get the maximum cardinality T-join problem.

Let G be a graph and ϕ a weighting of its lines by (not necessarily non-negative) integers. Set

$$E^+ = \{e \mid \phi(e) \geq 0\},\ E^- = \{e \mid \phi(e) < 0\}, \text{ and } \psi(e) = |\phi(e)|.$$

Let T be an even subset of $V(G)$. As before, let $r(G,T,\phi)$ denote the minimum weight of a T-join. That is, let

$$r(G,T,\phi) = \min\{\phi(J) \mid J \text{ is a } T\text{-join}\}.$$

Here $\phi(J)$ is the sum of the weights of the lines in J. Let T_ϕ denote the set of those points which are incident with an odd number of negative lines.

6.5.21. LEMMA. *For each graph G, every even subset T of $V(G)$ and every integral weighting ϕ of the lines of G,*

$$r(G,T,\phi) = r(G, T \oplus T_\phi, \psi) + \phi(E^-).$$

PROOF. Let J be any subset of $E(G)$. Then

$$\begin{aligned}
\phi(J) &= \phi(J \cap E^+) + \phi(J \cap E^-) \\
&= \phi(J \cap E^+) + \phi(J \cap E^-) + \phi(E^- - J) + \psi(E^- - J) \\
&= \psi(J \oplus E^-) + \phi(E^-).
\end{aligned}$$

Now J is a T-join if and only if $J \oplus E^-$ is a $T \oplus T_\phi$-join. Hence we see that

$$\begin{aligned}
r(G,T,\phi) &= \min\{\phi(J) \mid J \text{ is a } T\text{-join}\} \\
&= \min\{\psi(J \oplus E^-) \mid J \text{ is a } T\text{-join}\} + \phi(E^-) \\
&= \min\{\psi(J') \mid J' \text{ is a } T \oplus T_\phi\text{-join}\} + \phi(E^-) \\
&= r(G, T \oplus T_\phi, \psi) + \phi(E^-).
\end{aligned}$$

■

Using Theorem 6.5.20, we may now obtain a good characterization of $r(G,T,\phi)$ for general weights:

6.5.22. COROLLARY. *For every graph G, every even subset T of $V(G)$, and every integral weighting ϕ of the lines of G,*

$$r(G,T,\phi) = \nu(G, T \oplus T_\phi, 2\psi)/2 + \phi(E^-).$$

■

By specializing, the solution of the weighted T-join problem also yields a solution of the weighted matching problem. We do not go into the details here, since the weighted matching, perfect matching and T-join problems are best understood in the context of linear programming, and this approach will be the subject of the next chapter.

We do point out that if we have a polynomial-time algorithm to find a minimum weight T-join in a graph with non-negative weights, then the construction in the proof above yields a polynomial-time algorithm to find a minimum weight T-join in a graph with *arbitrary* weights. Such an algorithm will follow from the results in Chapter 9.

6.6. Optimum Paths, Cycles, Joins and Cuts

In this section we survey a number of further graph theoretical results which are in one way or the other related to matching theory. Many of these are, in fact, directly reducible to a matching (or T-join) problem. In other cases such reductions are much more complicated and in still other cases the problems are only analogous to matching problems.

Let G be a graph and suppose $u, v \in V(G)$. The problem of finding a shortest $u - v$ path in G is easily solved by the breadth-first search algorithm (Dijkstra (1959); see also Box 2A and Christofides (1975)). This simple algorithm also extends to the case when each line has a non-negative *length* assigned. It also applies to digraphs.

We may also formulate a minimax theorem for the minimum length of a $u - v$ path. Let D be a digraph and let us assign a non-negative "length" $\phi(e)$ to each line $e \in E(D)$. A **potential** on the weighted digraph (D, ϕ) is a mapping $\pi : V(D) \to \Re$ such that

$$\pi(y) - \pi(x) \leq \phi(xy)$$

for each line $xy \in E(D)$.

6.6.1. THEOREM. *Let (D, ϕ) be a weighted digraph in which the weights are non-negative and assume $u, v \in V(D)$. Assume also that there is a directed $u - v$ path in D. Then the minimum length of a directed $u - v$ path in (D, ϕ) is equal to the maximum of $\pi(v) - \pi(u)$, where π is a potential on (D, ϕ).*

PROOF. I. Let $P = (u = v_0, v_1, \ldots, v_k = v)$ be any directed $u - v$ path and π, any potential. Then by the definition of a potential,

$$\phi(E(P)) = \sum_{i=1}^{k} \phi(v_{i-1}v_i) \geq \sum_{i=1}^{k} (\pi(v_i) - \pi(v_{i-1})) = \pi(v) - \pi(u).$$

II. It remains to be seen that there exists a potential π which gives equality here. Such a potential can be defined as follows:

$$\pi(x) = \begin{cases} \text{the minimum length of a } u-x \text{ dipath} \\ \qquad \text{if such exists, and} \\ \infty, \text{otherwise.} \end{cases}$$

Then it is easy to check that π is indeed a potential and hence the theorem is proved. ■

The problem of finding a $u-v$ path of *maximum* length contains the Hamilton cycle problem and is, therefore, NP-complete. Since the problem of finding a maximum length $u-v$ path may also be viewed as the problem of finding a shortest $u-v$ path in a graph where each line has length -1, it follows that the problem of finding a shortest $u-v$ path in a graph, where the lines are given arbitrary integral weights, is also NP-complete. By way of contrast, we have just seen that this problem with *non-negative* weights is easily solvable! In fact we may relax the condition that all weights are non-negative and still obtain a tractable problem as follows.

Let G be a graph [digraph] and assume that every line is assigned an integral "length" $\phi(e)$. We say that the weighted graph [digraph] (G, ϕ) is **conservative** if every cycle [directed cycle] has non-negative total length. (The name comes from the observation that in a digraph of this type we cannot "gain" by going around any directed cycle. In other words, "energy is conserved".)

In the case of directed graphs, Theorem 6.6.1 remains valid also for conservative weightings, along with its proof. But when is a weighting conservative? And do the potentials in the statement of the theorem always exist? Fortunately, these questions answer each other, as the following result shows. The proof is left to the reader.

6.6.2. THEOREM. *Let (D, ϕ) be any weighted digraph. Then there exists a potential on (D, ϕ) if and only if (D, ϕ) contains no cycle with negative weight.* ■

Let us also show an alternative treatment of the shortest dipath problem to point out the connection with matchings.

Let (D, ϕ) be a conservative weighted directed graph and suppose $u, v \in V(D)$. Construct a bipartite graph G as follows. For each point $x \in V(D)$, take two new points x', x'' and for each directed line $(x, y) \in E(D)$, connect x' to y'' by a line. Furthermore, connect x' to x'' by a line for each $x \in V(D)$. Let the lines $x'x''$ have weight 0 and the lines $x'y''((x,y) \in E(D))$ have weight $\phi((x,y))$.

6.6.3. LEMMA. *The minimum length of a directed $u - v$ path in D is equal to the minimum weight of a perfect matching in $G - u' - v''$.*

PROOF. I. Let P be a directed $u - v$ path in D of minimum length. Then the lines $x'y''((x,y) \in E(P))$ and lines $x'x''(x \notin V(P))$ form a perfect matching in $G - u' - v''$ the weight of which is equal to the length of P.

II. Let F be a minimum weight perfect matching in $G - u' - v''$ and form the set $\overline{F} = \{(x,y) \in E(D) \mid x'y'' \in F\}$. Then $\overline{F}$ is the point-disjoint union of a directed $u - v$ path P and some directed cycles.

Moreover, the total length of $\overline{F}$ is equal to the weight of F. Since each directed cycle has non-negative length by hypothesis, the minimality of F implies that $\overline{F}$ contains only cycles with weight 0 and so the length of P is equal to the weight of F. This concludes the proof. ∎

Lemma 6.6.3 reduces the problem of finding a minimum length directed $u - v$ path in a conservative weighted digraph to the problem of finding a minimum weight perfect matching in an associated weighted bigraph. We leave it to the reader to derive the minimax formula of Theorem 6.6.1 from this observation.

Now let us turn to the case of undirected graphs. In this case the problem of finding a shortest path in a *conservative* weighted graph is substantially more difficult than the same problem for *non-negative* weights. It is possible to define a "distance" as the minimum length of paths joining two given points, and then introduce an analogue of the "potential". This treatment, initiated by Frank, Sebő and Tardos (1983) and developed by Sebő (1984a, 1984b) leads to a new proof and substantial strengthening of the Chinese Postman Theorem, and also to an extension of the Gallai-Edmonds Structure Theorem to T-joins. Here we shall follow a different road and show how the problem of finding a shortest path in a conservative weighted graph can be reduced to the Chinese Postman Problem. By the results of Chapter 9, this will also yield a polynomial-time algorithm for finding shortest paths in conservative weighted graphs.

Let (G, ϕ) be a conservative weighted graph and suppose $u, v \in V(G)$. Then a minimum length $\{u,v\}$-join is the line-disjoint union of a $u - v$ path and cycles. Since the graph is conservative, the removal of these cycles does not increase the length of the $\{u,v\}$-join and hence it follows that among the minimum $\{u,v\}$-joins there is a $u-v$ path. So it suffices to determine the minimum length of a $\{u,v\}$-join. Applying Theorem 6.5.21, we obtain the following minimax result:

6.6.4. THEOREM. *Let (G,ϕ) be a conservative weighted graph and $u,v \in V(G)$. Let $d_\phi(u,v)$ denote the minimum length of all $u-v$ paths. Let T be the set of those points of G which are incident with an odd number of negative lines. Then*

$$d_\phi(u,v) = \nu(G, T \oplus \{u,v\}, 2|\phi|)/2 + \phi(E^-).$$

■

A polynomial-time algorithm to find a minimum length $u-v$ path (i.e., a minimum length $\{u,v\}$-join) in a conservative weighted graph will follow from the results of Chapter 9 as remarked before.

But which graphs and digraphs are conservative? It is interesting (although perhaps not too surprising) that the methods just used above provide a good characterization of these graphs in the undirected case.

6.6.5. THEOREM. *Let (G,ϕ) be a weighted graph. Then (G,ϕ) is conservative if and only if there exists a collection of (not necessarily distinct) cuts such that each of them contains exactly one negative line, every non-negative line e is contained in at most $2\phi(e)$ of the cuts, while every negative line e is contained in exactly $-2\phi(e)$ of them.*

PROOF. Clearly (G,ϕ) is conservative if and only if every Eulerian subgraph (or $\emptyset$-join) in G has non-negative weight (i.e., if and only if $r(G,\emptyset,\phi) \geq 0$). By Corollary 6.5.21, this happens if and only if $\nu(G,T_\phi,2\psi) + 2\phi(E^-) \geq 0$, where $\psi = |\phi|$; that is, if and only if there exists a 2ψ-packing of $-2\phi(E^-) = 2\psi(E^-)$ T_ϕ-cuts. Hence the "if" part of the assertion is obvious.

To see the converse, let $\mathcal{L}$ be a 2ψ-packing of $\psi(E^-)$ T_ϕ-cuts. By Lemma 6.5.7, each T_ϕ-cut contains an odd number of negative lines. Hence

$$|\mathcal{L}| \leq \sum_{J\in\mathcal{L}} |J \cap E^-| \leq \sum_{e\in E^-} 2\psi(e) = 2\psi(E^-) = |\mathcal{L}|.$$

Since equality holds here, each cut in $\mathcal{L}$ must contain exactly one negative line and each negative line e must be contained in exactly $\psi(e)$ cuts of $\mathcal{L}$. So $\mathcal{L}$ has the properties claimed. ■

We may wonder if instead of $2|\phi|$-packings, the conservatism of the weighted graph (G,ϕ) could be characterized by the existence of an appropriate $|\phi|$-packing. The weighting of K_4 shown in Figure 6.6.1 is a counterexample. This graph is conservative; on the other hand, it does not contain two line-disjoint cuts at all, and one cut does not suffice to prove its conservatism.

6.6.6. EXERCISE. Assume that the negative lines of (G, ϕ) form a spanning tree T of G. Prove that (G, ϕ) is conservative if and only if

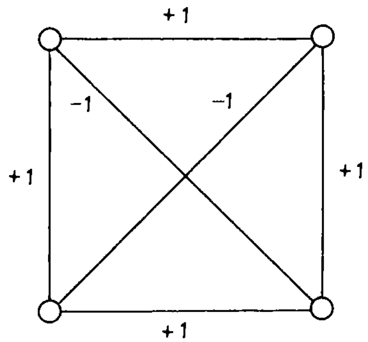


FIGURE 6.6.1.

every fundamental cycle with respect to T (i.e., every cycle containing exactly one line not in T) has non-negative length.

6.6.7. EXERCISE. Let (G, ϕ) be a conservative graph. Find a good characterization for the minimum length of cycles in G. (Korach (1982)).

We have derived the characterization of conservative weighted graphs (Theorem 6.6.5) from Theorem 6.5.10. These two results are, however, "equivalent" in a sense. More exactly, let J be any T-join and define a weighting $w_J : E(G) \to \{-1, 1\}$ by $w_J(e) = -1$ if and only if $e \in J$. Then the following lemma can be proved easily.

6.6.8. LEMMA. (*Guan* (1962).) *Let J be any T-join in a graph G and let (G, w_J) be as defined above. Then J is a minimum T-join if and only if the weighted graph (G, w_J) is conservative.* ∎

Using this lemma, the non-trivial part of Theorem 6.5.10 can be derived by considering a minimum T-join J and applying Theorem 6.6.5 to the weighted graph (G, w_j).

But how can we find a shortest *odd* path? For the case of non-negative weights, this problem was solved by Edmonds (unpublished) who reduced it to a matching problem as follows. For the sake of simplicity, we shall only consider the case in which each line has weight 1.

Let G' be a second copy of G and construct a graph G'' by considering the disjoint union of G and G', and then connecting every point $v \in V(G)$ to the corresponding point $v' \in V(G')$ by a new line. Let these new lines have weight 0, all the other lines, weight 1. Then the following analogue of Lemma 6.6.3 can be proved by the same argument.

6.6.9. LEMMA. *The minimum length of an odd $u-v$ path in G is equal to the minimum weight of a perfect matching in $G''-u-v$.* ■

Lemma 6.6.9 reduces the problem of finding a minimum length odd $u-v$ path to the problem of finding a minimum weight perfect matching in G''. Thus Theorem 6.5.20 specialized to G'', yields a minimax formula for the minimum length of an odd path. Since this minimax formula is not very appealing, we will not go into details.

6.6.10. EXERCISE. Show that the problem of finding the minimum length of an *even* $u-v$ path in a graph with non-negative weights is also reducible to a matching problem.

6.6.11. EXERCISE. Show that the problem of finding a minimum odd (respectively, even) *cycle* in a graph with non-negative weights can be reduced to a shortest path problem (respectively, to a weighted matching problem).

One may also ask for the shortest odd $u-v$ path in a *conservative* graph. We do not know if this problem is still reducible to the matching problem or tractable by any other means.

There are a number of different notions of "join" for which the problem of finding a minimum join can be solved and these considerations lead to a number of interesting graph theoretical results.

First, let us consider minimal sets of lines which meet not only all T-cuts (as T-joins do), but *all* cuts. These sets are clearly spanning trees. Since all spanning trees have the same number of lines, only the weighted problem is interesting. The problem of finding a spanning tree with minimum weight was solved by Borůvka (1926a, 1926b) and Kruskal (1956). (For a thorough historical account of this problem, see Graham and Hell (1982).) We obtain the solution by repeatedly choosing the line with least weight which does not form a cycle when taken together with the lines already chosen. This algorithm is called *greedy* (see Box 1C), and the fact that it always yields an optimum spanning tree is very important. It turns out that the reason this procedure works is the fact that the spanning trees of a graph form the bases of a matroid. We could formulate a minimax theorem as well, but it would not be as transparent as the above greedy algorithm.

This is a case when the algorithm provides the best description of the optimum value. (A somewhat similar situation was encountered in the problem of finding a shortest $u-v$ path. There too the algorithm was easier to state than the corresponding minimax theorem.)

We have two rather more difficult problems if we consider directed graphs. The first of these may be viewed as an oriented analogue of the Chinese Postman Problem.

The police of Nashpest, in trying to cope with increasing traffic density, made all streets one-way. However, they were not careful enough and it turned out that there were points in the city from which other points could not be reached at all! So the police had to allow 2-way traffic on some streets to resume. Of course they wanted to minimize the number of such streets. How can we find the minimum number of streets such that when they are made two-way, every point in the city can be reached from every other point?

The problem is then to find, in a given digraph D, a set of lines of minimum cardinality such that if we add an "anti-parallel" line to each we obtain a strongly connected digraph. It is clear that this property is equivalent to saying that the set meets every directed cut in D. A set with this property is called a **join**, and the minimum size of a join will be called the **join number** of G.

The join number is strongly related to another important invariant of digraphs, the **feedback number**. The latter is defined as the minimum number of lines whose deletion destroys every directed cycle in the graph. (Note that the undirected version of the feedback number is the cyclomatic number.) If the digraph D describes a "flow chart" of something (perhaps a computer program, an investment, or a chemical reaction among several compounds), the analysis is often easier if the digraph D does not contain any directed cycles, and it can be reduced to this case by analyzing what happens at such directed cycles. This suggests that to find the feedback number is an important first step in analyzing such flow charts. Let us hasten to point out that the problem of computing the feedback number is NP-complete (Karp, (1972, 1975)) and we shall not go into the details of results concerning it. But there is an important connection between the feedback number and the join number. Let D be a planar digraph and D^*, its dual. (The orientation of D defines an orientation of D^* in a natural way: let, say, each line of D^* cross the corresponding line of D from the left to the right.) Then it is easy to see that the feedback number of D is equal to the join number of D^*. Thus the solution of the join number problem which follows will also yield a solution of the feedback number problem for planar graphs.

It turns out that the situation with joins in digraphs is simpler than with T-joins in graphs; the following theorem provides a simple minimax relation for the join number. This relation was conjectured by Robertson (unpublished) and Younger (1969), and proved by Lucchesi and Younger (1978).

6.6.12. THEOREM. *The join number of any digraph D is equal to the maximum number of line-disjoint directed cuts in D.* ∎

Theorem 6.6.12 can be proved by copying the proof of Theorem 6.5.10, with very little modification. In fact, at several points the arguments become much simpler, as might be expected considering the simpler formulation of the theorem. For the details the reader is referred to Lucchesi and Younger (1978) and to Lovász (1976a).

By planar duality as mentioned above, we obtain:

6.6.13. COROLLARY. *The feedback number of any planar digraph D is equal to the maximum number of line-disjoint directed cycles in D.* ∎

A further related problem arises if we consider a weighted digraph (D, ω) and a specified point $r \in V(D)$ called a **root**. For every non-empty subset $S \subseteq V(D) - \{r\}$, the set of lines entering S will be called a **cut rooted at** r or, more concisely, an r-**cut**. Observe that the minimal sets of lines covering all r-cuts are precisely the spanning arborescences rooted at r, that is, the spanning trees of D which are oriented away from the root r. These arborescences will be called r-**branchings**. We are interested in the minimum weight of an r-branching. (Note that since all branchings have the same cardinality, the unweighted problem is trivial. But — to date, at least — the weighted problem is not, and this indicates that the weighted problem very likely cannot be reduced to the unweighted case in any straightforward way.) In the discussion to follow we shall assume that D contains at least one r-branching; that is, every point can be reached from the root on a directed path. The following theorem was proved implicitly by Edmonds (1967b) and explicitly by Fulkerson (1974).

6.6.14. THEOREM. *For every digraph D, root r and weighting ω of the lines by non-negative integers, the minimum weight of an r-branching is equal to the maximum size of an ω-packing of r-cuts.* ∎

We now turn our attention to the problem of finding minimum cuts. This problem is in many respects dual to the minimum join (path, cycle) problem. Let us start by recalling that the problem of finding a minimum uv-cut in a graph (or digraph) is solved by the Max-Flow-Min-Cut Theorem 2.1.4 and by the various flow algorithms discussed in Section 2.2. By applying these results to every pair of points u, v, the minimum cut problem for the whole graph can be solved. An elegant relation between minimum cuts for all pairs $\{u, v\}$ is provided by the Flow-equivalent Tree Theorem of Gomory and Hu (Theorem 2.3.2), which will also play some role in the present discussion.

Let G be a graph and T, an even subset of $V(G)$. What is the minimum size of a T-cut of G? The following result of Padberg and Rao (1982) answers this question.

6.6.15. THEOREM. *Let G be a graph and T an even subset of $V(G)$. Let F be a cut-equivalent tree for G. Then the minimum T-cut among the cuts determined by F is a minimum T-cut of G.*

PROOF. (It will be clear from this proof that there is always *at least one* T-cut among the cuts determined by F.) Let C be a minimum T-cut in G with shores S_1 and S_2. It suffices to show that F contains a line e which connects S_1 to S_2 and which determines a T-cut C_1 of G. For then C_1 is minimal among all cuts separating the endpoints of e. Since C is such a cut, it follows that $|C_1| \leq |C|$. But since C is a minimum T-cut, we see that C_1 too is a minimum T-cut.

Let H be the set of those lines of F which determine a T-cut of G, that is, for which both components of $F - e$ contain an odd number of points of T. It is immediate that these lines form a T-join in F. Since (S_1, S_2) determines a T-cut of F, it follows that H contains at least one line which connects S_1 to S_2. ∎

So the minimum T-cut problem is linked to matching theory via flow theory. A more intrinsic and interesting connection between minimum cuts and maximum matchings will be discussed in Chapter 9, where it will be shown that from an algorithmic point of view these two problems are equivalent.

6.6.16. EXERCISE. Let (G, ϕ) be a weighted graph with ϕ non-negative. Show that the problem of finding a T-cut with minimum weight can be reduced to the problem of finding a minimum cardinality T-cut.

6.6.17. EXERCISE. Let D be a digraph. Show that the problem of finding a minimum directed cut in D can be reduced to the problem of finding a minimum cut.

Now let us consider the case of general (not necessarily non-negative) weights. If all weights are -1, then the problem of finding a minimum weight cut is equivalent to the problem of finding a maximum cut, or equivalently, to the problem of finding the minimum number of lines whose removal destroys all odd cycles. This problem is known to be NP-complete (Karp, (1972, 1975)). There is, however, an important special case for which it can be solved, namely the case in which the graph is planar (Hadlock (1975), Orlova and Dorfman (1972)). In fact, in the case of *planar* graphs the minimum weight cut problem can be solved for general (positive and negative) weights.

Let G be a planar graph and ϕ a weighting of its lines by arbitrary integers. Let G^* be the planar dual of G. Then ϕ may also be viewed as a weighting of the lines of G^*. A (not necessarily proper) cut in G corresponds to an Eulerian subgraph ($\emptyset$-join) in G^* and vice versa. Thus we are interested in the minimum weight of a $\emptyset$-join in G^*. Applying Theorem 6.5.20, we obtain the following results.

6.6.18. THEOREM. *Let (G, ϕ) be a weighted planar graph and let m denote the maximum size of a $2|\phi|$-packing of cycles in G such that each of the cycles contains an odd number of negative lines. Then the minimum weight of a (not necessarily proper) cut of G is equal to $m/2 + \phi(E^-)$.* ■

By this trick, the problem of finding a minimum weight cut in a planar graph (even if negative weights are allowed) can be reduced to finding a minimum $\emptyset$-join in its dual. This, in turn, can be reduced to a matching problem and solved in polynomial time by the methods of Chapter 9. We also obtain the following corollary. (Note the nice analogy with the feedback number!)

6.6.19. COROLLARY. *Let G be a planar graph. Then the minimum number of lines which destroy all odd cycles is equal to half the maximum size of a 2-packing of odd cycles.* ■

The complete graph K_4 serves to show that, instead of considering a 2-packing of odd cycles and then dividing by 2, it would not suffice to consider a 1-packing of odd cycles. Also, if we consider larger complete graphs, we see that the condition of planarity cannot be dropped. In K_5 we need 4 lines to destroy all odd cycles. On the other hand, every odd cycle has at least 3 lines and hence every k-packing of odd cycles consists of no more than $10k/3 < 4k$ cycles.

At this point, just as in the case of joins, it is natural to study the problems of finding minimum (and minimum weight) cuts, directed cuts and r-cuts. Each of these problems can be solved by the flow algorithm, applied to various pairs of points. We shall bring this section to a close by presenting a minimax theorem concerning r-cuts which is truly elegant. This result is due to Edmonds (1970, 1973).

6.6.20. THEOREM. *Let G be a digraph with root r. Then the minimum size of an r-cut is equal to the maximum number of line-disjoint r-branchings.* ■

(A weighted version of the above result is easily derived by considering multiple lines.)

6.6.21. EXERCISE. Let D be a k-line-connected digraph, and suppose $a_1, \ldots, a_k, b_1, \ldots, b_k \in V(G)$. Then there are k line-disjoint directed paths $P_1, \ldots, P_k$ in D such that P_i connects a_i to b_i. Show that no similar assertion is true for undirected graphs. (Shiloach (1979)).

6.6.22. EXERCISE. Formulate theorems dual (in the planar sense) to Theorems 6.6.14 and 6.6.20.

BOX 6B. Packing Paths, Cycles, Joins and Cuts

Theorem 6.5.10, the Lucchesi-Younger Theorem 6.6.12 and the Edmonds-Fulkerson Theorem 6.6.14 are all examples in which the problem of optimum (fractional, respectively integral) packing of cuts can be characterized by a minimax formula. There are many results in graph theory which are similarly related. We mention the results of Seymour (1978) on packing 2-commodity cuts, and of Lins (1981) on packing "coronas". In these results on packing cuts, a key fact is that some version of the uncrossing procedure can be carried out.

Along with various types of cuts, packings of various "joins" constitute an interesting, but in many respects more difficult, set of problems. The packing problem for T-joins with $|T| = 2$ is just Menger's Theorem, which states that the maximum number of line-disjoint T-joins is equal to the minimum size of a T-cut. This relation does not generalize for $|T| > 2$. If $T = V(G)$, then a cubic graph has 3 line-disjoint $V(G)$-joins if and only if it is 3-line-colorable, which is in turn an NP-complete problem (Holyer (1981)).

It was conjectured by Edmonds and Giles (1977) that the maximum number of line-disjoint joins in a digraph is equal to the minimum size of a directed cut, but this was recently disproved by Schrijver (1980a). On the other hand, this relation is known to hold for a large class of digraphs (Schrijver (1982)).

The Disjoint Branching Theorem of Edmonds (1973) may also be viewed as a minimax result for packing branchings. The maximum number of line-disjoint spanning trees of a graph was characterized by Tutte (1961) and Nash-Williams (1961, 1964). If we study packing problems for various other notions of "joins", we find theorems on 2-commodity flows (Hu (1963), Rothschild and Whinston (1966a, 1966b), Seymour (1979c), Okamura and Seymour (1981)), on graph connectivity and path packing (Gallai (1961), Lovász (1970d, 1976b), and Mader (1978a, 1978b, 1979a, 1979b)) and on and on. Closely related to these are cycle packing problems (Gallai (1957/58), Seymour (1979b) and Lins (1981)). But at this point an enormous amount of graph theory is sneaking in, so we quickly close Pandora's box!

7

Matching and Linear Programming

7.0. Introduction

During the last 20 years, linear programming has become one of the most widely used and most powerful tools in combinatorics. One of the early successes in the application of linear programming to combinatorial problems was the treatment of flows and bipartite matching, using the concept of "total unimodularity" (Ford and Fulkerson (1958), Hoffman and Kruskal (1956)). These results show that the Max-Flow Min-Cut Theorem, as well as König's Theorem, may be viewed as sharpened (integral) versions of the Duality Theorem of Linear Programming, for certain special, highly structured cases. In a celebrated paper, Edmonds (1965b) extended the effective range of linear programming methods to non-bipartite matching problems, by determining the defining inequalities for the convex hull of matchings. This work has served as a prototype for treating many other combinatorial optimization problems, and a number of these have been solved completely or partially in this sense (branching, matroid intersection, vertex packing, travelling salesman, and others).

In this chapter we first discuss solving the bipartite matching problem as a consequence of the Duality Theorem. In considering two natural ways of generalizing this problem, we will be led to the notion and study of "fractional" matchings on the one hand, and to Edmonds' Theorem about the matching polytope on the other. Observing that the latter may be viewed as a characterization of the "fractional chromatic index", we shall make a brief detour to line colorations, and prove the famous theorem of Vizing. The descriptions of the matching polytope and of the fractional matching polytope have many other nice applications and we will treat some of these.

In the last section we set out to determine the dimension of the perfect matching polytope. This problem turns out to be rather difficult and the solution makes use of the structure theory developed in Chapter 5.

BOX 7A. Cones, Polytopes and Polyhedra, and other Preliminaries from Linear Programming

Let $v_1, \ldots, v_m$ be vectors in $\Re^n$. Vector $v = \lambda_1 v_1 + \cdots + \lambda_m v_m$, $(\lambda_i \in \Re)$ is called a **linear combination** of $v_1, \ldots, v_m$. An **affine combination** is a linear combination with $\lambda_1 + \cdots + \lambda_m = 1$; a **conical combination** is a linear combination in which each $\lambda_i \geq 0$; and a **convex combination** is an affine combination in which each $\lambda_i \geq 0$. The **linear (affine, conical, convex) hull** of $\{v_1, \ldots, v_m\}$ is the set of all linear (affine, conical, convex) combinations of $v_1, \ldots, v_m$. A set of vectors is called **affinely independent** if no one of them is an affine combination of the others. Note that the column vectors $v_1, \ldots, v_n$ are affinely independent if and only if the augmented column vectors $\binom{1}{v_1}, \ldots, \binom{1}{v_n}$ are linearly independent.

The convex hull of a finite set of vectors is called a (**convex polyhedral**) **cone**. A cone is **pointed** if $\mathbf{0}$ is a vertex, that is, if there is a hyperplane which intersects the cone in the unique point $\mathbf{0}$. This is equivalent to saying that $\mathbf{0}$ is not a non-trivial non-negative linear combination of the defining vectors.

A theorem due to Weyl (1935) says that every convex polyhedral cone is the intersection of a finite number of halfspaces which have $\mathbf{0}$ on their boundary. Algebraically, this means that the cone is the solution set of a system of homogeneous linear inequalities. Conversely, every system of homogeneous linear inequalities has a convex cone as its solution set.

The reader should consider himself duly warned that the number of vectors needed to generate a polyhedral cone may be exponentially large in the dimension and in the number of linear inequalities defining the cone. As an example consider the cone of all non-negative $n \times n$ matrices with equal row and column sums. This cone in $\Re^{n^2}$ is defined by n^2 inequalities and $2(n-1)$ equations, but we need $n!$ matrices (the permutation matrices). (See the Birkhoff–von Neumann Theorem 1.4.13.) It may also happen that the number of linear inequalities needed to define a cone is exponentially large in the dimension and in the number of generating vectors.

The above two definitions of convex cones (as non-negative combinations of defining vectors and as the solution set of a system of homogeneous defining inequalities) together provide a *good characterization* of convex cones. More exactly, if we are given both a list of vectors generating the cone and a list of linear inequalities defining the same cone, then to prove that a vector belongs to the convex cone, it suffices to represent it as a non-negative combination of defining vectors. To prove that it does not belong to the convex cone, it suffices to find a defining inequality which is violated by this vector.

One frequently meets cones given by some description other than defining vectors or defining inequalities. (See, for example, Exercise 7.3.5.) In this case it may happen that both the number of defining vectors *and* the number of defining inequalities are exponentially large in terms of the "size" of this description!

So to have a good characterization of a cone given by some description of length n, we need

(a) a way to exhibit that a defining vector belongs to the cone in time polynomial in n, and

(b) a way to exhibit that a defining inequality is valid for every vector in the cone in time polynomial in n.

Usually we have (a) or (b) trivially and then we try to achieve the other. If the cone is given by a list of defining vectors, then (a) is the trivial task of checking whether a given vector is in the list; (b) is also trivially solved by substituting the defining vectors into the given inequality. The above mentioned good characterization of polyhedral cones can be stated in the following more algebraic form. This result is fundamental in linear programming; for a proof see Chvátal (1983).

7A.1. LEMMA. (*The Farkas Lemma*). *If $a_1, \ldots, a_m, b \in \Re^n$, then either b is a non-negative linear combination of $a_1, \ldots, a_m$, or there exists a vector $u \in \Re^n$ such that $a_i \cdot u \leq 0$, but $b \cdot u > 0$.* ■

Returning to good characterizations of convex cones, a further minor, but subtle, point has to be clarified. Suppose that we have a procedure for (a). How can we exhibit that an arbitrary vector x belongs to the cone? We represent x as a non-negative combination of the defining vectors, and then prove that these defining vectors are indeed in the cone. But the number of defining vectors may be very large, and so just to write down this representation may take an exponentially long time! Fortunately, there is a useful fact (called Carathéodory's Theorem (1907)) which says that every vector in the cone can be represented as a non-negative linear combination of at most n defining vectors, where n is the dimension of the space.

Given a finite number of points (vectors) in $\Re^n$, their convex hull is called a **polytope**. A classical result of Minkowski and Weyl states that every polytope is the intersection of a finite number of halfspaces. Algebraically, this says that every polytope can be defined as the set of points satisfying a system of linear inequalities. Knowing both a set of defining points and a system of linear inequalities provides a *good characterization* of the polytope: if we want to exhibit that a certain vector belongs to the polytope, all we have to do is to give a representation of it as a convex combination of defining points; if we want to exhibit that a certain vector does not belong to the polytope, all we have to do is to find one of the defining inequalities which is violated by the vector. To make this precise, the same observations have to be made as were made above for cones.

For every polytope there exists a unique minimal set of defining points, the set of its **vertices**. A vertex can also be characterized as a point in the polytope which is not contained in a segment connecting two other points. Another characterization of a vertex is as a point for which there exists some valid inequality such that this is the only point yielding equality.

The situation is somewhat more complicated for defining inequalities. An **interior point** is a point satisfying all defining inequalities with strict inequality. A polytope is called **full dimensional** if it has an interior point.

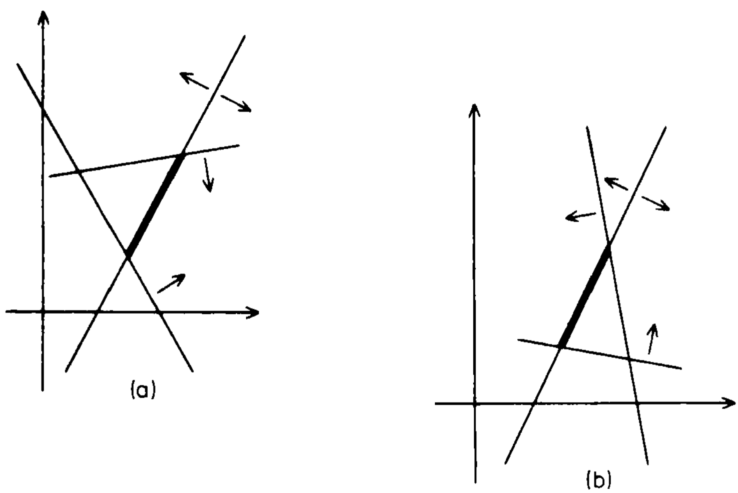


FIGURE 7A.1. Two different descriptions of a polytope

In this case, there is a *unique* minimal set of defining inequalities (up to multiplication by a positive number). Inequalities belonging to this unique minimal system are called **essential**. The set of points of a polytope which satisfy an essential inequality with equality is called a **facet**. There is a one-to-one correspondence between facets and essential inequalities (up to a positive multiplier).

Essential inequalities are characterized by the fact that there are n *affinely* independent vertices which give equality in them. Equivalently, an inequality is essential for a full dimensional polytope if it is valid and is not a non-negative combination of other valid inequalities. Yet another characterization says that a valid inequality "$*$" is a facet if and only if every valid inequality which holds with equality whenever $*$ does, is a positive multiple of $*$.

Vertices and facets are special cases of the notion of a *face*. A **face** is the set of points giving equality in some valid inequality.

If the polytope is not full dimensional then there exist linear equations which are satisfied by all points of the polytope. Adding a multiple of one of these equations to any of the defining inequalities does not change the polytope obtained thereby. So in general there is no unique minimal set of inequalities defining the polytope, and we shall be content with finding a collection of inequalities which is "reasonable" in some sense. (See Figure 7A.1.)

The solution set of a system of linear inequalities is a **polyhedron**. Every polytope is a polyhedron, but not conversely. In fact, a polyhedron is a polytope if and only if it is bounded. If we want to describe a polyhedron — which is not a polytope — it does not suffice to list its vertices. Motzkin (1936) proved that we can represent any polyhedron as the sum of a polytope and a convex polyhedral cone in the following sense: we list some *defining points* and some other vectors called *defining rays*, and then form all points which arise as a convex combination of the defining points plus a non-negative combination of the rays. (See Figure 7A.2 for an example.)

Once a list of defining points and defining rays is found, we have an efficient way to exhibit that a point *is* in the polyhedron: we just represent it as a non-negative combination of defining points and rays, where the coefficients of defining points add up to 1. Of course we have an efficient way to exhibit that a point is *not* in the polyhedron by exhibiting a violated defining inequality. For a more precise description of this idea of good polyhedral characterization, remarks similar to those for cones given prior to Lemma 7A.1 apply.

The **dimension** of a polyhedron P, $\dim P$, is defined as the dimension of the affine subspace spanned by P. Note that the maximum number of affinely independent points is equal to $1 + \dim P$.

THE FUNDAMENTAL PROBLEM OF LINEAR PROGRAMMING is to find the minimum or maximum of a linear *objective function* over a polyhedron, that is, subject to *linear constraints*. There are many standard forms in which linear programming problems can be written. For combinatorial applications, one common form is

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c}\cdot\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \geq \mathbf{0} \\ & A\boldsymbol{x} \leq \boldsymbol{b}, \end{array} \tag{7A.1}$$

where A is an $n \times m$ matrix, $\boldsymbol{b} \in \Re^n$, $\boldsymbol{c} \in \Re^m$, and $\mathbf{0}$ stands for the zero vector of dimension m. We define the **dual program** of (7A.1) (and denote it by 7A.2) as follows:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{b}\cdot\boldsymbol{y} \\ \text{subject to} & \boldsymbol{y} \geq \mathbf{0} \\ & A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}. \end{array} \tag{7A.2}$$

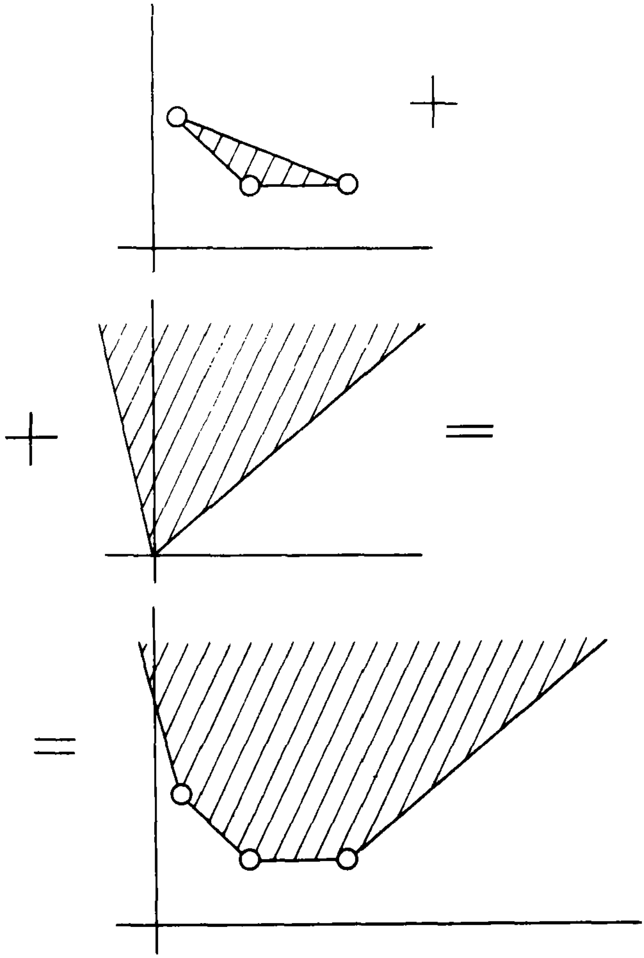


FIGURE 7A.2. A polyhedron as the sum of a polytope and a cone

The relation between these two programs is described by the following fundamental theorem. It appears implicitly first in a working paper written by von Neumann (1947), and later as a full-blown theorem with proof in Gale, Kuhn and Tucker (1950).

7A.2. THEOREM. (*The Duality Theorem of Linear Programming*). *If either one of the programs* (7A.1) *or* (7A.2) *has a solution and a finite optimum, then so does the other and the two optima are equal.* ■

This theorem does not only transform one linear programming problem to another, but it also yields a good characterization of the optimum value of the program in the following sense. Assume that we want to prove that the optimum value of program (7A.1) is a. Then we can show that it is not less than a by exhibiting a feasible solution of (7A.1) for which the value of the objective function is a. To show that the optimum value is not larger than this, we can use a feasible solution of the dual program (7A.2) for which the value of the dual objective is a.

It is very useful to know that if (7A.1) has a feasible solution and a finite optimum, and if the corresponding polyhedron has at least one vertex, then this optimum is attained at a vertex of the polyhedron. Hence it follows, among other things, that if the matrix A and the vector b are rational, then among the optimum solutions of (7A.1) at least one has rational entries.

As general references on linear programming we suggest Dantzig (1963) and Chvátal (1983).

BOX 7B. Linear Programming Algorithms

In the troubled days of the late 1930's and early 1940's, Kantorovich first formulated the linear programming problem in a truly mathematical setting. (See Kantorovich (1939, 1942) and Kantorovich and Gavurin (1949).) Unfortunately this work lay unnoticed for many years. Also the crucial idea of a formal solution procedure — that is, an *algorithm* — was not yet formulated.

In the years just after World War II, Dantzig and von Neumann independently discovered the new discipline. Moreover, a giant step forward was taken when Dantzig (1951) produced the first — and still the best — algorithm for linear programming, the now-famous *Simplex Method*. Linear programming has since become one of the most widely and successfully applied branches of mathematics. The reason for this success is not only because linear programs are at the core of a wide variety of mathematical models in economics, industry and science, but also because now we have the Simplex Method to solve them.

From the geometric point of view, the Simplex Method is very easy to understand. To solve a linear program we must find a maximum, say, of a linear objective function over a polyhedron P (given as the solution set of a system of linear inequalities). Without loss of generality we may visualize this objective function as "elevation", that is, we want to find the highest point of the polyhedron P. This can be achieved — quite naturally — as follows. Start at any vertex v of P. Look for an edge connecting v to a vertex v' which is higher than v, and use it to travel to v'. Repeating this, we sooner or later reach a vertex from which all edges go down. It is easy to show then that this vertex is the highest point of P.

Technically speaking, the above description of the Simplex Method is complete only if the linear program is non-degenerate, that is, only n

of the constraints are satisfied with equality at the same vertex, where n is the dimension of P. In this case, only n edges are incident with any vertex and so it is easy to decide for them all whether they go up or down. However, it is easy to reduce the general case to the case of simple polytopes — say by "perturbing" the defining inequalities just a bit. (In practice, however, there are better ways to do this). So in this informal discussion we shall confine ourselves to such non-degenerate linear programs.

Not only is it easy to explain the geometric content of the Simplex Method, but it is easy to implement it as a sequence of simple manipulations on the columns of the matrix describing P. Such a description of the Simplex Method can be found in any textbook on linear programming, and we shall not go into details. Let us remark at least, however, that these manipulations of the columns can be described in a completely combinatorial manner, and such a combinatorial abstraction of linear programming has been one of the sources of the theory of oriented matroids (Bland (1977), Bland and Las Vergnas (1978), Folkman and Lawrence (1978), Fukuda (1982) and Mandel (1982)).

Now let us consider the running time of the Simplex Method. We start by making a preliminary observation. Suppose that we are at a vertex of P from which more than one edge is going "up". Which of these edges should we follow? Several simple rules-of-thumb can be applied to select this edge: steepest ascent, largest increase in the objective function value, etc. Such a rule is called a **pivoting rule**. Whatever pivoting rule we choose, the Simplex Method will terminate in a finite number of steps with an optimum solution, but the running time does depend on the pivoting rule chosen. (This finite termination is automatically valid only if P is simple, but we will stick to this case).

Now in practice the Simplex Method seems to run extremely fast with any of several simple pivoting rules; it can solve linear programs up to several thousand variables and constraints on a reasonably fast computer. However, for each of these pivoting rules we can construct diabolical programs on which the Simplex Method takes *exponentially* long to run! So while the average performance of the Simplex Method is very good, its worst-case performance is very bad (see Klee-Minty (1972), Zadeh (1973), Jeroslow (1973)). For theoretical results on the *average* performance of the Simplex Method, see Borgwardt (1982a, 1982b), Smale (1982, 1983), Haimovich (1983) and Adler, Megiddo and Todd (1984). Thus from a theoretical point of view, the existence of the Simplex Method has not solved the question of algorithmic complexity of linear programming: can linear programs be solved in polynomial time? This problem is of great *theoretical* importance since the optimum value of a program is a well-characterized function of the input by the Duality Theorem. In fact, for some time, the optimum value of a linear program was the Number One

Candidate for a well-characterized function which could not be computed in polynomial time.

However, Hačijan (1979) showed that a method of non-linear optimization, due to Šor (1970, 1977) and to Judin and Nemirovskiĭ (1976), but unfortunately not widely known, could be adapted to solve linear programs in polynomial time. The method has become known as the *Ellipsoid Method*. The popular press picked up this result and — unfortunately — advertised it as an important breakthrough *in practice*, which it was never claimed to be! The Simplex Method was — and is — satisfactory for all practical purposes. But the Ellipsoid Method did represent a theoretical breakthrough and some of its most important implications are in the field of combinatorial optimization which render it important for the purposes of this book.

Even though the geometry of the Ellipsoid Method is more complicated than that of the Simplex Method, it is still fairly easy to explain. Let us start with a very simple version. Suppose that we want to decide if a system of strict linear inequalities

$$\boldsymbol{a}_i \cdot \boldsymbol{x} < b_i \qquad (i = 1, \ldots, m) \tag{7B.1}$$

has a solution, where $\boldsymbol{a}_i \in Z^n$, $b_i \in Z$. Let T denote the maximum absolute value of entries of $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ and $b_1, \ldots, b_m$. First we can prove that if system (7B.1) has a solution, then it has a solution with $|x_j| \leq (nT)^n$. Thus, we may add the inequalities

$$-(nT)^n < x_j < (nT)^n \qquad (j = 1, \ldots, n) \tag{7B.2}$$

to the system without changing the answer to our question. Then, if P denotes the solution set of (7B.1) together with (7B.2), then P is contained in the ball E_0 of radius $R = n^{n+1}T^n$ and center $\mathbf{0}$.

For later reference let us remark that if P is non-empty then it has a positive volume; in fact, in this case we have the following lower bound on its volume:

$$\operatorname{vol} P \geq (nT)^{-(n+1)^2}.$$

The Ellipsoid Method consists of constructing a sequence $E_1, E_2, \ldots$ of ellipsoids, each including P and shrinking in volume. Suppose that E_k has been constructed. Let $\boldsymbol{x}^k$ be its center. Test to see if $\boldsymbol{x}^k$ satisfies all the inequalities (7B.1). If the answer is yes, then we can stop, for $\boldsymbol{x}^k$ is a solution. Suppose that $\boldsymbol{x}^k$ violates one of the inequalities in (7B.1), for example $\boldsymbol{a}_i \cdot \boldsymbol{x}^k \geq b_i$, for some $1 \leq i \leq m$. Then consider the intersection

$$E_k' = E_k \cap \{\boldsymbol{x} \mid \boldsymbol{a}_i \cdot \boldsymbol{x} \geq b_i\}$$

and include this in an ellipsoid with minimum volume. Let this new ellipsoid be E_{k+1}. By this construction, each E_k will include P.

We can show that

$$\operatorname{vol} E_{k+1} \leq e^{-1/(2(n+1))} \operatorname{vol} E_k,$$

from which it follows that

$$\operatorname{vol} E_k \le e^{-k/(2(n+1))} \operatorname{vol} E_0 \le e^{-k/(2(n+1))} (nT)^{(n+1)^2}.$$

So the volumes of the ellipsoids E_k tend to 0 as a geometric progression. Since $E_k \supseteq P$, we also have

$$\operatorname{vol} E_k \ge \operatorname{vol} P,$$

which implies that if P is non-empty, then the procedure must terminate

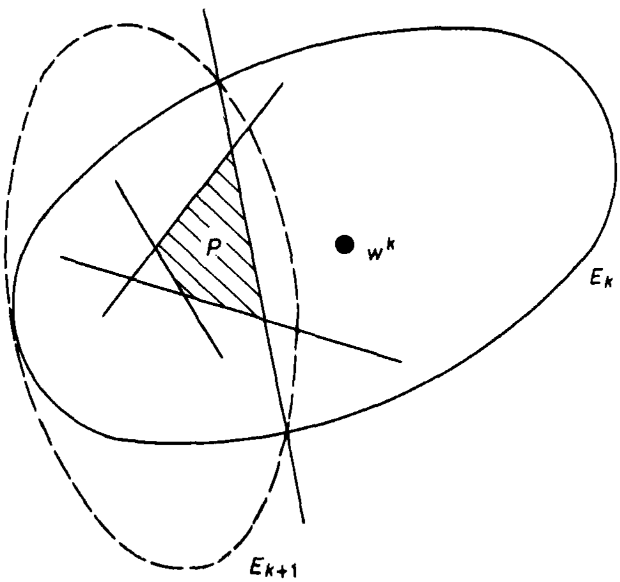


FIGURE 7B.1. A step in the Ellipsoid Method

in a finite number of steps. More precisely, suppose that E_k is constructed, and that P is non-empty. Then

$$(nT)^{-(n+1)^2} \le \operatorname{vol} P \le \operatorname{vol} E_k \le e^{-k/(2(n+1))} (nT)^{(n+1)^2}.$$

Hence

$$k \le 4(n+1)^3 (\log n + \log T).$$

Thus if the sequence $E_0, E_1, E_2, \ldots,$ does not terminate after $2(n+1)^3 \lceil \log n + \log T \rceil$ steps, we can stop and be certain that (7B.1) has no solution at all. Since the input length of the problem (i.e., the number of binary digits necessary to write down all the coefficients in the problem) is at least $\log T$ and also at least n, this means that only a polynomial number of ellipsoids need be constructed.

To obtain a polynomial-time algorithm for linear programming from this, we must add a fairly large number of details, which we shall not develop here. Let us mention a few, however. How do we describe an ellipsoid? How do we determine the ellipsoid with minimum volume including E'_k? What do we do if rounding errors occur? How large can the numbers grow during these iterations? How do we handle the case of non-strict inequalities? What do we do with the objective function? None of these represents unresolvable difficulties, but the details are quite tedious, and therefore we refer the reader to the papers cited above, as well as to Gács and Lovász (1981), Bland, Goldfarb and Todd (1981), and Grötschel, Lovász and Schrijver (1981).

From a practical point of view, the Ellipsoid Method has proved inferior to the Simplex Method because of the numerical problems accompanying it and also because its running time, though *always* polynomial, is usually quite unattractive. But from the point of view of theoretical applications to combinatorial optimization problems, it has two very nice features. First, the bound on the number of iterations

$$4(n+1)^3(\log T + \log n)$$

does not depend on the number of constraints; and second, one does not need to list all the constraints in advance. All we need is a way to identify a constraint which is violated by the center x_k of the current ellipsoid. These features can be exploited to prove far-reaching reduction principles between combinatorial optimization problems (see Grötschel, Lovász and Schrijver (1981)). Let us formulate here one version of such "equivalence principles", which takes relatively little preparation. We shall apply this result in Chapter 9.

Let $\mathcal{K}$ be a class of polytopes. Suppose that each member P of $\mathcal{K}$ has some "description" or "name", where we assume that the dimension of the linear space which contains P is bounded by a polynomial in the length of the "name" of P. Also assume that each polytope in $\mathcal{K}$ has rational vertices and that the numbers of binary digits in numerators and denominators of entries of vertices of P are bounded by a polynomial in the length of the name of P. For the purposes of this box, let us call such a class $\mathcal{K}$ "well-behaved".

7B.1. THEOREM. *Let $\mathcal{K}$ be a well-behaved class of polytopes, and suppose that there is a polynomial algorithm which, given the "name" of any $P \in \mathcal{K}$, $P \subseteq \Re^n$ and any $x \in \Re^n$, decides if $x \in P$ and if it is not, finds a hyperplane separating x from P. Then we can find the maximum of any linear objective function $c \cdot x$ $(c \in \mathcal{Q}^n)$ over $x \in P$ in polynomial time.* ■

A weaker version of this theorem (for the case when the polytope is full dimensional) is proved in Grötschel, Lovász and Schrijver (1981). The present version is proved in Grötschel, Lovász and Schrijver (1984c). Extensions to the non-full dimensional case, but with restrictions on the separation subroutine, were proved earlier by Karp and Papadimitriou (1980, 1982) and Padberg and Rao (1981).

Quite recently, Karmarkar (1984) has designed a new polynomial-time algorithm to solve linear programs. This algorithm is based upon ideas from projective geometry and is too involved to be discussed here. However, it is important to remark that the first tests show this algorithm to be comparable with (and indeed even faster than) the Simplex Method. On the other hand, in its present form, it does not seem to yield the combinatorial applications mentioned above.

7.1. Linear Programming and Matching in Bigraphs

Let G be a bigraph with bipartition $V(G)=(U,W)$. We shall work in the following two linear spaces: $\Re^{E(G)}$, the set of all vectors whose entries are indexed by the lines of G, and $\Re^{V(G)}$, the set of all vectors whose entries are indexed by the points of G. Every subset $F \subseteq E(G)$ can be described by its **incidence vector**, the $|E(G)|$-tuple $\boldsymbol{q}^F = (\alpha_e : e \in E(G)) \in \Re^{E(G)}$, where

$$\alpha_e = \begin{cases} 1, & \text{if } e \in F, \\ 0, & \text{otherwise.} \end{cases}$$

Similarly we can assign an incidence vector to every subset $S \subseteq V(G)$.

Let us adopt one more general convention. If we have an arbitrary vector $\boldsymbol{a} = (a_e : e \in E(G)) \in \Re^{E(G)}$ and if $F \subseteq E(G)$, then we set

$$\boldsymbol{a}(F) = \boldsymbol{q}^F \cdot \boldsymbol{a} = \sum_{e \in F} a_e.$$

It is trivial that a 0–1 vector $\boldsymbol{x}$ in $\Re^{E(G)}$ is the incidence vector of a matching in G if and only if

$$\boldsymbol{x}(\nabla(v)) \leq 1, \tag{7.1.1}$$

for every point $v \in V(G)$.

In fact, if $\boldsymbol{x}$ is the incidence vector of any set $F \subseteq E(G)$, then the left hand side of (7.1.1) just counts the number of lines in F incident with point v. Also, the cardinality of F may be expressed very simply in terms of its incidence vector $\boldsymbol{x} = \boldsymbol{q}^F$ by

$$|F| = \boldsymbol{x}(E(G)) = \boldsymbol{x} \cdot \mathbf{1}, \tag{7.1.2}$$

where $\mathbf{1} = \boldsymbol{q}^{E(G)}$ is the vector with all entries equal to 1. So the problem of finding a maximum matching in G is equivalent to finding a 0–1 vector $\boldsymbol{x}$ which maximizes (7.1.2) subject to and $|V(G)|$ constraints of type (7.1.1). If the assumption that $\boldsymbol{x}$ is a 0–1 vector were dropped, we would have a *linear programming problem*, which could be solved easily by the powerful methods of linear programming. However, x must be a 0–1 vector, and this condition cannot be expressed by linear inequalities. Let us try to do as much as we can, however, and write

$$0 \leq x_e \leq 1 \qquad (e \in E(G)).$$

Here the upper bounds may as well be omitted, since they are implied by inequalities (7.1.1). So we are left with the $|E(G)|$ inequalities

$$x_e \geq 0 \qquad (e \in E(G)). \tag{7.1.3}$$

We often write (7.1.1) – (7.1.3) in the more compact form

$$\text{maximize} \quad \mathbf{1} \cdot \boldsymbol{x} \tag{7.1.4}$$

$$\text{subject to} \qquad \boldsymbol{x} \geq \mathbf{0} \tag{7.1.5}$$

$$A\boldsymbol{x} \leq \mathbf{1},$$

where $A = (a_{ve})$ is the **point-line incidence matrix** of G, that is,

$$a_{ve} = \begin{cases} 1, & \text{if } v \text{ is an endpoint of } e, \\ 0, & \text{otherwise.} \end{cases}$$

The reader may wonder at this point why we want to formulate (7.1.4) – (7.1.5), since it is not equivalent to the original problem of finding a maximum matching. In particular, (7.1.3) does not imply that $\boldsymbol{x}$ is a 0–1 vector, and hence an optimum solution of (7.1.4) – (7.1.5) may not have any combinatorial meaning as a set of lines. However, in the case of bipartite graphs we can easily show that *among the optimum solutions of* (7.1.4) – (7.1.5) *there will be one, which is a* 0–1 *vector*. In fact, the solutions of (7.1.5) form a polytope $M(G)$, and among those points of this polytope which maximize the linear objective function $\mathbf{1} \cdot \boldsymbol{x}$, at least one will be a vertex. So if we show that the vertices of the polytope $M(G)$ are 0–1 vectors we are done. In view of (7.1.3), it suffices to show that they are integral. To this end, recall that a vertex of a polytope in $\Re^q$ can be obtained as the unique intersection point of the hyperplanes

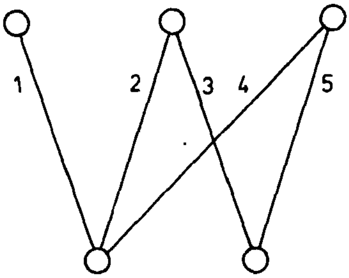


FIGURE 7.1.1.

of q of its facets. In our case this means that we have to select q linearly independent rows of the matrix $\binom{A}{I}$, turn the corresponding inequalities in (7.1.5) into equalities, and solve the resulting system of linear equations. Eliminating those variables which are immediately given as 0's, we are left with a system

$$A_1 \boldsymbol{x}' = \mathbf{1}, \tag{7.1.6}$$

where A_1 is a square non-singular submatrix of A.

As an example, take the graph in Figure 7.1.1. Then the system of linear inequalities (7.1.5) is the following:

$$\begin{array}{lllllcl}
x_1 & & & & & \le 1 & \qquad x_1 \ge 0 \\
 & x_2 & +x_3 & & & \le 1 & \qquad x_2 \ge 0 \\
 & & x_3 & & +x_5 & \le 1 & \qquad x_3 \ge 0 \\
x_1 & +x_2 & & +x_4 & & \le 1 & \qquad x_4 \ge 0 \\
 & & & x_4 & +x_5 & \le 1 & \qquad x_5 \ge 0.
\end{array} \tag{7.1.7}$$

Any vertex of the polytope described by (7.1.7) will satisfy 5 of the inequalities with linearly independent left hand sides with equality. (Of course, it also must satisfy the rest as well.) Suppose, for example that this vertex gives equality in the first four inequalities of the group of inequalities on the left hand side of (7.1.7) and in the last inequality on the right. Then, substituting $x_5 = 0$, we obtain from the first four equations

$$\begin{array}{llllll} x_1 & & & & = 1 \\ & x_2 + x_3 & & & = 1 \\ & & x_3 & & = 1 \\ x_1 + x_2 & & & + x_4 & = 1. \end{array}$$

The matrix of this system

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}$$

has determinant 1. Now if we solve (7.1.6) by Cramer's Rule, the entries of the solution will be rational numbers whose denominators are $\det A_1$. Thus the fact that the solution is integral follows immediately from the following lemma:

7.1.1. LEMMA. *If G is bipartite, then every square submatrix of A has determinant 0 or ± 1.*

PROOF. Consider any square $k \times k$ submatrix Q of A. We use induction on k; the case $k = 1$ being trivial.

If Q has a column consisting only of zeros, then $\det Q = 0$. If Q has a column containing exactly one 1, then we can expand $\det Q$ about this column and proceed by induction. So we may assume that every column of Q contains two 1's.

The rows of Q correspond to the points of G and so they can be partitioned into two classes according to the two color classes of G. Let, say, the first t rows correspond to one color class of G, and the last $k-t$ rows to the other color class. Then every column of Q has one 1 in the first t rows and one 1 in the last $k-t$ rows. But this means that the sum of the first t row vectors is equal to the sum of the last $k-t$ row vectors (both sums being equal to $(1, \ldots, 1)$). So the rows of Q are linearly dependent, and hence $\det Q = 0$. ∎

It is worth-while to notice that we have, in fact, proved the following:

7.1.2. THEOREM. *Let G be a bipartite graph. Then the vertices of the polytope (7.1.5) are 0–1 vectors. In fact, they are exactly the incidence vectors of matchings.* ∎

The problem of finding minimum point covers is quite analogous. A 0–1 vector $\boldsymbol{y}$ is the incidence vector of a point cover if and only if it satisfies

$$y_u + y_v \geq 1,$$

for every $uv \in E(G)$.

If we replace the condition that $\boldsymbol{y}$ is a 0–1 vector by the constraint that $\boldsymbol{y}$ is non-negative, we obtain similarly as before, the following linear programming problem:

$$\text{minimize} \quad \mathbf{1} \cdot \boldsymbol{y} \tag{7.1.8}$$

$$\begin{aligned} \text{subject to} \quad & \boldsymbol{y} \geq \mathbf{0} \\ & A^{\mathrm{T}} \boldsymbol{y} \geq \mathbf{1}. \end{aligned} \tag{7.1.9}$$

Using Lemma 7.1.1 similarly as before, we obtain that the vertices of the polyhedron (7.1.9) are integral. It is easy to see that those integral solutions of (7.1.9) which are not 0–1, or which are 0–1, but are not the incidence vectors of minimal point covers, are not vertices. Thus we obtain the following result.

7.1.3. THEOREM. *If G is a bipartite graph, then the vertices of the polyhedron* (7.1.9) *are exactly the incidence vectors of* (*inclusionwise*) *minimal point covers.* ■

Note that (7.1.9) does not describe a bounded body, and so we cannot conclude that it is the convex hull of incidence vectors of point covers. Actually, it is the sum of this convex hull with the non-negative orthant.

König's Minimax Theorem is an immediate consequence of these results. Just note that Theorem 7.1.2 implies that

$$\nu(G) = \max\{\mathbf{1} \cdot \boldsymbol{x} \mid \boldsymbol{x} \in \Re^{E(G)},\ \boldsymbol{x} \geq \mathbf{0},\ A\boldsymbol{x} \leq \mathbf{1}\},$$

and Theorem 7.1.3 implies

$$\tau(G) = \min\{\mathbf{1} \cdot \boldsymbol{y} \mid \boldsymbol{y} \in \Re^{V(G)},\ \boldsymbol{y} \geq \mathbf{0},\ A^{\mathrm{T}} \boldsymbol{y} \geq \mathbf{1}\}.$$

But the two programs on the right hand sides are dual to each other, and since their solvability is obvious, we get $\nu(G) = \tau(G)$ by the Duality Theorem.

But Theorems 7.1.2 and 7.1.3 say more than König's Theorem. Their most essential consequence is that they imply minimax formulas for the *weighted matching problem* and the *weighted point cover problem.* First, let a non-negative integer weight w_e be assigned to every line e, and write $\boldsymbol{w} = (w_e : e \in E(G))$. We would like to find a matching in the bipartite graph G having maximum weight. This may be formulated as the problem of maximizing the linear objective function $\boldsymbol{w} \cdot \boldsymbol{x}$ over all 0–1 vectors $\boldsymbol{x}$ that are incidence vectors of matchings. By Theorem 7.1.2, this maximum is equal to

$$\max\{\boldsymbol{w} \cdot \boldsymbol{x} | \, \boldsymbol{x} \in \Re^{E(G)}, \, \boldsymbol{x} \geq \mathbf{0}, \, A\boldsymbol{x} \leq \mathbf{1}\}. \tag{7.1.10}$$

By the Duality Theorem, this is equal to

$$\min\{\mathbf{1} \cdot \boldsymbol{y} | \, \boldsymbol{y} \in \Re^{V(G)}, \, \boldsymbol{y} \geq \mathbf{0}, \, A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{w}\}. \tag{7.1.11}$$

Using Lemma 7.1.1 it is easy to prove that the vertices of the polyhedron $\{\boldsymbol{y} \in \Re^{V(G)} \mid \boldsymbol{y} \geq \mathbf{0}, \, A^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{w}\}$ are integral. (It does not follow that they are 0–1 vectors!) We may interpret an integer vector $\boldsymbol{y}$ solving (7.1.11) as a collection of points of G, where point v occurs y_v times in the collection. So we have proved the following result (Egerváry (1931)):

7.1.4. THEOREM. *Let G be a bipartite graph and let a non-negative integral weight w_e be assigned to every line e of G. Then the maximum weight of a matching in G is equal to the minimum number of points in a collection from which every line e meets at least w_e points.* ■

Let a non-negative integer weight b_v be assigned to every point v. A b-**matching** is an assignment of non-negative integral weights to the lines such that the sum of weights of lines incident with any given point v is at most b_v. An argument similar to that above also proves the following:

7.1.5. THEOREM. *Let G be a bipartite graph. Let a non-negative integral weight w_e be assigned to every line e, and also let a non-negative integral weight b_v be assigned to every point v. Then the maximum weight of a t-matching in G is equal to the minimum b-weight of a collection of points from which every line e contains at least w_e elements.* ■

7.1.6. EXERCISE. Show that Theorem 7.1.5 can easily be deduced by applying Theorem 7.1.4 to the bipartite graph obtained from G by replacing each point v by b_v points and connecting all the b_u copies of u to all the b_v copies of v if and only if u and v are adjacent in G. (No similar simple trick is known for obtaining Theorem 7.1.4 from König's Theorem.)

7.1.7. EXERCISE. Give a proof of the Max-Flow Min-Cut Theorem and the Flow Integrality Theorem using the same kind of argument as above.

7.1.8. EXERCISE. Use Theorem 7.1.2 to derive the Birkhoff–von Neumann Theorem (Corollary 1.4.13).

7.1.9. EXERCISE. Formulate the line coloration problem for bipartite graphs as an integer linear programming problem. Derive König's Line Coloring Theorem from this formulation.

7.1.10. EXERCISE. Prove that every bipartite graph with maximum degree $\leq k$ has a k-line coloration in which the number of lines of each color is $\lfloor q/k \rfloor$ or $\lceil q/k \rceil$. (For reference see de Werra (1972).)

BOX 7C. The Hoffman-Kruskal Theorem and Other Conditions of Integrality

The proof of König's Theorem given above may be adapted to prove other minimax theorems, including the Max-Flow Min-Cut Theorem (see Exercise 7.1.7). The crucial point in these proofs is to show that among the optimal solutions of an appropriate linear program, there is at least one whose entries are integers. A sufficient condition for this is that the vertices of the polytope (or polyhedron) of feasible solutions are integral points. There are some general criteria which guarantee this. Probably the most important one to date was abstracted from the above proof of König's Theorem by Hoffman and Kruskal (1956). A matrix A is called **totally unimodular** if every square submatrix of A has determinant 0 or ± 1.

7C.1. THEOREM. *Let A be a totally unimodular $n \times m$ matrix. Then for all integral vectors $\boldsymbol{b}_1, \boldsymbol{b}_2 \in Z_+^n$ and $\boldsymbol{d}_1, \boldsymbol{d}_2 \in Z^m$, the polyhedron*

$$\{\boldsymbol{x} \in \Re^m \mid \boldsymbol{b}_1 \leq A\boldsymbol{x} \leq \boldsymbol{b}_2,\ \boldsymbol{d}_1 \leq \boldsymbol{x} \leq \boldsymbol{d}_2\}$$

has all integral vertices. Conversely, if this polyhedron has all integral vertices for every choice of integral vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{d}_1$ and $\boldsymbol{d}_2$, then matrix A is totally unimodular. ■

Another simple but very useful result on the integrality of the vertices of a polyhedron was proved by Hoffman (1974). For the extension to the unbounded case, see Edmonds and Giles (1977).

7C.2. THEOREM. *A polytope has all integral vertices if and only if for every linear objective function with integral coefficients which has a maximum over the polyhedron, this maximum is integral.* ■

The way this theorem is usually applied is to show that for any linear objective function with integral coefficients, and for an appropriate description of the polyhedron by linear inequalities having integral coefficients, the dual problem has at least one integral optimal solution. This automatically implies that the optimum value of the objective function (which is common for the primal and dual problems) is an integer. (A system of linear inequalities such that for any linear objective function with integral coefficients the dual program has an integral solution is called **totally dual integral**.) To derive, for example, König's Theorem in this way, we only have to show that if $G = (U, W)$ is a bipartite graph and w is any non-negative integral weighting of the lines, then among the optimum solutions of (7.1.11) at least one is integer-valued. The same property of (7.1.10) will then follow by the theorem above.

Proving that (7.1.11) has integral vertices can be done directly, without using total unimodularity. In fact, let $\boldsymbol{x}_0 = (x_{01}, \dots, x_{0p})$ be a non-integral point of (7.1.11). Let A and B denote the sets of points of U and W, respectively, which correspond to non-integral entries in $\boldsymbol{x}_0$, and let $\boldsymbol{a}$ and $\boldsymbol{b}$ be their incidence vectors. Then for a small enough $\epsilon > 0$, both vectors

$$\boldsymbol{x}_0 + \epsilon(\boldsymbol{a} - \boldsymbol{b}), \qquad \boldsymbol{x}_0 - \epsilon(\boldsymbol{a} - \boldsymbol{b})$$

satisfy (7.1.11), and since $\boldsymbol{x}_0$ is on the segment connecting them, it follows that $\boldsymbol{x}_0$ cannot be a vertex.

7.2. Matchings and Fractional Matchings

If we consider non-bipartite graphs, Theorem 7.1.2 does not remain valid. For example, if G is a triangle then the point $(1/2, 1/2, 1/2)$ is a fractional vertex of the polytope described by (7.1.5). It is still true that the incidence vector of every matching belongs to the polytope (7.1.5), but (7.1.5) is in general not the convex hull of these matchings.

So in the case of general (not necessarily bipartite) graphs two more polytopes may be introduced. One is the **matching polytope**, denoted $M(G)$, which is the convex hull of incidence vectors of all matchings in the graph G. The other is the polytope of solutions of (7.1.5), which we call the **fractional matching polytope** and which will be denoted by $FM(G)$. As remarked above, $M(G) \subseteq FM(G)$.

We shall also consider the convex hull of all *perfect* matchings which we call the **perfect matching polytope** of G and denote by $PM(G)$.

Obviously, $PM(G) \subseteq M(G)$; in fact, $PM(G)$ is the face of $M(G)$ determined by the valid inequality $\boldsymbol{x}(E(G)) \leq p/2$.

7.2.1. EXERCISE. Prove that $M(G) = FM(G)$ if and only if G is bipartite.

Similarly, Theorem 7.1.3 fails to hold for non-bipartite graphs, and this motivates the introduction of two other polyhedra. One is the convex hull of all point covers, called the **point cover polyhedron**, and the other is the solution set of (7.1.9), which we shall call the **fractional point cover polyhedron**. They will be denoted by $PC(G)$ and $FPC(G)$, respectively.

The main objective of this chapter is to study the five polyhedra $M(G)$, $FM(G)$, $PM(G)$, $PC(G)$, and $FPC(G)$. Let us remark at the outset that these five objects are not equally important or equally difficult. For us the most important will be the matching polytope. The study of the facets of this polytope will be the central theme in our treatment, and this will have the most important applications for us. Note that since $M(G)$ is given by its vertices, the key problem is to find its description in terms of inequalities. In the case of the polyhedra $FM(G)$ and $FPC(G)$, which are described to us by inequalities, the main problem will be to find the vertices. This will be relatively easy. (It is, in fact, essentially equivalent to the bipartite matching problem.) Finally, the polyhedron $PC(G)$ turns out to be a real villain; it has very complicated facets which have not been — and probably never will be — described in full! We shall only survey the main directions of attack on $PC(G)$ in Chapter 12.

7.3. The Matching Polytope

Recall that the matching polytope $M(G)$ of a graph G is defined as the convex hull of incidence vectors of matchings in G. In this section we prove an important theorem which characterizes the facets of $M(G)$. The more important half of this result, which says that every facet of $M(G)$ is of the form given below, is a celebrated result of Edmonds (1965b). The converse, that every inequality given below does indeed define a facet, was proved by Pulleyblank (1973) (see also Pulleyblank and Edmonds (1974)). The proof given here is due to Lovász (1979a).

7.3.1. THEOREM. *The facets of $M(G)$ are given by the following inequalities:*

(i) $\quad \boldsymbol{x} \geq \mathbf{0}$

(ii) $\quad \boldsymbol{x}(\nabla(v)) \leq 1 \quad$ (*where v is a non-isolated point such that if v has only one neighbor u, then $\{uv\}$ is a connected component of G, and if v has exactly two neighbors, then they are not adjacent*)

(iii) $\quad \boldsymbol{x}(E(S)) \leq (|S|-1)/2 \quad$ (*where S spans a 2-connected factor-critical subgraph*).

The somewhat lengthy condition in (ii) serves only to rule out some degenerate cases. It is equivalent to saying that no line non-incident with v contains all neighbors of v. The interesting inequalities are those in (iii).

The inequalities (ii) are valid inequalities for the matching polytope for every point v. Similarly, the inequalities (iii) are valid for every odd set of points S. The restrictions on v and S are only to rule out those inequalities which are superflous in as much as they follow from the other inequalities.

PROOF. First we show that every facet of $M(G)$ is determined by an inequality of the given type. Let

$$\sum_{e \in E(G)} a_e x_e \leq b \tag{7.3.1}$$

be an inequality which defines a facet of $M(G)$. Every matching in G which satisfies (7.3.1) with equality will be called **an extremal matching**. The main thrust of our argument will be that if we find an inequality "$*$" such that every extremal matching satisfies it with equality, then this inequality may differ from (7.3.1) only by a positive scalar factor, since the polytope is full dimensional. In fact, the assumption that (7.3.1) is a facet implies that there are q extremal matchings which span the hyperplane

$$\sum_{e \in E(G)} a_e x_e = b \tag{7.3.2}$$

and so if we find a hyperplane which is incident with every extremal matching, we know that it must be identical to hyperplane (7.3.2). Thus inequality "$*$" must be either a non-negative multiple of inequality (7.3.1) or the inequality obtained from (7.3.1) by reversing the inequality sign. This latter possibility can be trivially ruled out by noticing that $M(G)$ is full dimensional and so it cannot be contained in the hyperplane (7.3.2).

Case 1. Assume first that there exists an $e \in E(G)$ with $a_e < 0$. Then no extremal matching can contain the line e, for then the deletion of e from the extremal matching would yield a matching violating inequality (7.3.1). Thus the incidence vector of every extremal matching satisfies $x_e = 0$. As remarked above, this implies that (7.3.1) is equivalent to one of the inequalities (i).

Case 2. Now assume that $a_e \geq 0$ for all lines e and also that there is a point v which is covered by every extremal matching. Then every extremal matching $\boldsymbol{x}$ satisfies

$$\boldsymbol{x}(\nabla(v)) = 1.$$

If v satisfies the condition in (ii) then this implies that inequality (7.3.1) is equivalent to one of the inequalities (ii). If v does not satisfy the condition in (ii), then there is a line e non-incident with v such that no matching covering v contains e. Thus every extremal matching has $x_e = 0$ and we conclude as in Case 1.

Case 3. So assume that $a_e \geq 0$ for each line $e \in E(G)$ and also that for every point of the graph, there exists an extremal matching which misses it. Let G_0 be the subgraph of G formed by those lines e with $a_e > 0$.

Claim 1. G_0 is connected. Assume to the contrary that G_0 is disconnected. Then it is the union of two disjoint non-empty subgraphs G_1 and G_2. Let

$$g_e^t = \begin{cases} a_e, & \text{if } e \in E(G_t) \\ 0, & \text{otherwise,} \end{cases}$$

and let

$$\boldsymbol{g}^t = (g_e^t : e \in E(G)) \in \Re^{E(G)} \quad (t = 1, 2).$$

Then trivially

$$\boldsymbol{g}^1 + \boldsymbol{g}^2 = \boldsymbol{a}. \tag{7.3.3}$$

Let $\boldsymbol{x}^i$ be the incidence vector of a matching maximizing the objective $\boldsymbol{g}^i \cdot \boldsymbol{x}$, and let h^i be the value of this maximum. Then

$$h^1 + h^2 = b. \tag{7.3.4}$$

In fact, if $\boldsymbol{x}^0$ is the incidence vector of an extremal matching, then

$$h^1 + h^2 \geq \boldsymbol{g}^1 \cdot \boldsymbol{x}^0 + \boldsymbol{g}^2 \cdot \boldsymbol{x}^0 = \boldsymbol{a} \cdot \boldsymbol{x}^0 = b.$$

On the other hand, we may assume that $\boldsymbol{x}^i$ is non-zero only on the lines of G_i and hence $\boldsymbol{x}^1+\boldsymbol{x}^2$ is a matching, and so

$$h^1+h^2=\boldsymbol{g}^1\cdot\boldsymbol{x}^1+\boldsymbol{g}^2\cdot\boldsymbol{x}^2=(\boldsymbol{g}^1+\boldsymbol{g}^2)\cdot(\boldsymbol{x}^1+\boldsymbol{x}^2)=\boldsymbol{a}\cdot(\boldsymbol{x}^1+\boldsymbol{x}^2)\le b.$$

Now equations (7.3.3) and (7.3.4) show that inequality (7.3.1) is the sum of two other valid inequalities, which is impossible since it is a facet.

Claim 2. No extremal matching misses two or more points of G_0. Assume to the contrary that an extremal matching M_1 misses two points u, v of G_0. We may also assume that u and v have minimum distance in G_0 among all pairs of points simultaneously missed by some extremal matching. Trivially, this distance is at least 2. Let w be any point on any shortest path connecting u to v. (Such a path exists by Claim 1). By the minimality property of the distance between u and v, M_1 must cover w. Let M_2 be an extremal matching which misses w. Then $M_1\cup M_2$ contains an alternating path A which starts from w. We must have

$$\boldsymbol{a}(M_1\cap A)=\boldsymbol{a}(M_2\cap A) \tag{7.3.5}$$

for if we "alternate" on A, one of the resulting matchings will violate inequality (7.3.1). But even if equation (7.3.5) holds, if we alternate on A we get from M_1 an extremal matching which misses w and one of u and v, which contradicts the choice of u and v.

Claim 3. G_0 is factor-critical. For let us consider an extremal matching M which misses a point v of G_0. We may delete any line of M which is not in G_0 and still have an extremal matching, so assume that $M\subseteq E(G_0)$. By Claim 2, matching M cannot miss any other point of G_0, so it is a perfect matching of G_0-v.

Claim 4. Every extremal matching contains exactly $(|S|-1)/2$ lines spanned by the set $S=V(G_0)$. For let M be any extremal matching. Delete all lines not in G_0 from M. This leaves us with another extremal matching M'. By Claim 2, matching M' misses at most one point of S. So M contains at least $(|S|-1)/2$ lines spanned by S. On the other hand, S is odd by Claim 3, and so M cannot contain more than $(|S|-1)/2$ lines spanned by S.

Now Claim 4 says that if $\boldsymbol{x}$ is the incidence vector of any extremal matching, then

$$\boldsymbol{x}(E(S))=\frac{|S|-1}{2}$$

and hence by the remark at the beginning of this proof, inequality (7.3.1) must be equivalent to the inequality

$$\boldsymbol{x}(E(S)) \le \frac{|S|-1}{2}. \tag{7.3.6}$$

It also follows that G_0 is the subgraph spanned by S, (by the definition of G_0). So to show that (7.3.1) is equivalent to an inequality in list (iii), we only have to prove one more property.

Claim 5. G_0 is 2-connected. Assume indirectly that G_0 is the union of G_1 and G_2, where G_1 and G_2 have more than one point and the two graphs have exactly one point in common. Let $S_i = V(G_i)$. Since G_0 is factor-critical, S_1 and S_2 must be odd, and hence the following inequalities hold for every matching $\boldsymbol{x}$:

$$\boldsymbol{x}(E(S_i)) \le \frac{|S_i|-1}{2} \quad (i = 1, 2).$$

Adding these we get inequality (7.3.6). But this contradicts the assumption that (7.3.1) (or equivalently (7.3.6)) defines a facet.

Thus we have proved that every facet of the matching polytope is of the form (i), (ii) or (iii). Now we are going to prove that these inequalities do indeed define facets.

(i). Consider the inequality $x_h \ge 0$ $(h \in E(G))$. The incidence vector of every one-element matching $\{e\}$ $(e \ne h)$, as well as the empty matching, satisfy this with equality. These are q affinely independent points in $\Re^E$, which implies that $x_h \ge 0$ defines a facet.

(ii). Consider the inequality

$$\boldsymbol{x}(\nabla(v)) \le 1, \tag{7.3.7}$$

where v is a point as described in the theorem. Construct the following q matchings: for every line e incident with v, take $\{e\}$; for every line e non-incident with v, pick a line f incident with v which is disjoint from e and take the matching $\{e, f\}$. These matchings all give equality in (7.3.7) and are affinely independent. This implies that inequality (7.3.7) defines a facet.

(iii). We could follow an argument similar to those in the previous two cases by exhibiting q affinely independent matchings, all giving equality in a given inequality

$$\boldsymbol{x}(E(S)) \le \frac{|S|-1}{2}, \tag{7.3.8}$$

where S spans a 2-connected factor-critical subgraph G_1. This construction would be considerably more complicated than the previous two. It could be based on the Ear Structure Theorem for 2-connected factor-critical graphs (Theorem 5.5.2). Essentially this approach was taken by Pulleyblank (1973). Instead we give another proof which uses more of the techniques of linear programming.

Suppose indirectly that (7.3.8) does not define a facet. Then there is a facet

$$\sum_{e\in E(G)} a_e x_e \le b, \tag{7.3.9}$$

such that every matching which gives equality in (7.3.8) also gives equality in (7.3.9). (We know that (7.3.9) is equivalent to one of the inequalities in the list (i) – (iii), but we can make little use of this fact at this point.)

Claim 1. If $e \notin E(S)$ then $a_e = 0$. In fact e has at most one endpoint in S and hence, using the assumption that G_1 is factor-critical, we can find a near-perfect matching M of G_1 such that $M + e$ is a matching. Now both M and $M + e$ give equality in (7.3.8) and hence they both must give equality in (7.3.9). But this clearly implies that $a_e = 0$.

Claim 2. If $e, f \in E(S)$, then $a_e = a_f$. Assume, to the contrary, that this is not the case. Since G_1 is connected, we may as well assume that e and f are incident with a common point v. Split v into two points v' and v'', where v' is incident with those lines h which used to be adjacent to v and which have $a_h = a_e$, and v'' is incident with the remaining lines originally incident with v. Using Tutte's Theorem, it follows easily that the graph G_2 obtained from G_1 in this way has a perfect matching (Exercise 3.1.11). Let M be a perfect matching of G_2, and let e_1 and e_2 be the two lines of M adjacent to v' and v'', respectively. Then both $M - e_1$ and $M - e_2$ correspond to near-perfect matchings in G_1, and so they both give equality in (7.3.8). But then they both have to give equality in (7.3.9). But this is not the case since the left hand sides of (7.3.9), upon substitution of $M - e_1$ and $M - e_2$, will differ by $a_{e_1} - a_{e_2} \neq 0$.

By these two claims we see that, possibly after multiplication by a positive constant, inequality (7.3.9) has the same left hand side as (7.3.8). Since the right hand sides must then be equal as well (just substitute any near-perfect matching of G_1), we see that (7.3.8) and (7.3.9) are equivalent. But this contradicts the assumption that (7.3.9) defines a facet while (7.3.8) does not. ■

7.3.2. EXERCISE. Use the above theorem to prove that if G is a 2-connected factor-critical graph, then G contains q linearly independent near-perfect matchings.

At this point let us formulate without proof another important property of the system of inequalities (i)–(iii) given in Theorem 7.3.1, due to Cunningham and Marsh (1978).

7.3.3. THEOREM. *Let G be a graph and $\mathbf{c} \in Z^{E(G)}$ any integral vector. Then there exists an optimal dual integral solution for the program:*

$$\begin{array}{lrl}
\text{maximize} & \mathbf{c} \cdot \mathbf{x} & \\
\text{subject to} & \mathbf{x} \geq \mathbf{0} & \\
& \mathbf{x}(\nabla(v)) \leq 1 & (v \in V(G)) \\
& \mathbf{x}(E(S)) \leq (|S|-1)/2 & (S \subseteq V(G),\ |S|\ \mathit{odd}).
\end{array}$$

■

Define the **perfect matching polytope** $PM(G)$ of the graph G as the convex hull of incidence vectors of perfect matchings in G. Then the perfect matching polytope is a face of the matching polytope, since it is the intersection of the matching polytope with the hyperplane

$$\mathbf{1} \cdot \mathbf{x} = p/2. \tag{7.3.10}$$

So a description of $PM(G)$ in terms of inequalities is trivial; all we have to add to the list of Theorem 7.3.1 is equation (7.3.10). But, unlike the matching polytope, the perfect matching polytope is not full dimensional, and hence its description by a minimal set of inequalities is not unique. A very appealing alternate description can be obtained as follows, however. Assume that G has an even number of points. (Otherwise $PM(G) = \emptyset$.) Recall from Chapter 6 that an **odd cut** (or **$V(G)$-cut**) of the graph G is the set of lines connecting S to $V(G)-S$, where $S \subseteq V(G)$ and $V(G)-S$ both have odd cardinality. The sets S and $V(G)-S$ will be called the **shores** of the cut. An odd cut is **trivial**, if one of its shores is a singleton; that is, if the cut is the star of a point.

7.3.4. THEOREM. *The perfect matching polytope $PM(G)$ may be described by the following constraints:*

$$\begin{array}{lll}
\text{(i}'\text{)} & \mathbf{x} \geq \mathbf{0} & \\
\text{(ii}'\text{)} & \mathbf{x}(C) = 1 & (\mathit{where}\ C\ \mathit{is\ a\ trivial\ odd\ cut}) \\
\text{(iii}'\text{)} & \mathbf{x}(C) \geq 1 & (\mathit{where}\ C\ \mathit{is\ a\ non\text{-}trivial\ odd\ cut}).
\end{array}$$

PROOF. We already know that (i), (ii) and (iii) in Theorem 7.3.1, together with equation (7.3.10), describe $PM(G)$. Since the inequalities (i′), (ii′) and (iii′) are trivially valid for every point in $PM(G)$, it suffices to show that they imply every inequality in (i), (ii), (iii) and (7.3.10). Now (i) is the same as (i′) and inequality (ii) is weaker than (ii′). Inequality (iii) for a given subset $S \subseteq V(G)$ follows by adding (ii′) for the star of every point in S, subtracting (iii′) for the odd cut determined by S, and then dividing by 2. Finally, equation (7.3.10) follows by adding (ii′) for every point and then dividing by 2. ■

7.3.5. EXERCISE. Let G be a graph with an even number of points and suppose $v_0 \in V(G)$. (a) Prove that the cone generated by the incidence vectors of the perfect matchings of G can be defined by the equations and inequalities:

$$\begin{aligned} \boldsymbol{x} &\geq 0 \\ \boldsymbol{x}(\nabla(v)) &= \boldsymbol{x}(\nabla(v_0)) && (\text{for } v \in V(G) - v_0) \\ \boldsymbol{x}(C) &\geq \boldsymbol{x}(\nabla(v_0)) && (\text{for } C \text{ a non-trivial odd cut}). \end{aligned}$$

(b) Prove that if G is complete, then none of these inequalities can be omitted.

As pointed out earlier, the set of inequalities given in Theorem 7.3.4 is not uniquely determined. It is not minimal either. To determine a minimal subset of (i′) – (iii′) which suffices to describe $PM(G)$ is a more difficult problem. We shall return to this problem in Section 7.6.

We have discussed T-joins in the previous chapter as generalizations of perfect matchings. The following theorem shows that they are a very natural generalization indeed.

Let G be a graph and T an even subset of $V(G)$. We define the **T-join polyhedron** $TJ(G)$ as the sum of the convex hull of incidence vectors of T-joins with the non-negative orthant; that is, a vector belongs to $TJ(G)$ if and only if it majorizes some convex combination of T-joins. The reason we do not simply take the convex hull of T-joins is that since we are only interested in minimal T-joins this difference does not matter, and the formulas will turn out to be simpler.

7.3.6. THEOREM. *The T-join polyhedron of G is described by the following inequalities:*

$$\begin{aligned} &(\text{i}'') && \boldsymbol{x} \geq 0 \\ &(\text{ii}'') && \boldsymbol{x}(C) \geq 1 && (\textit{for every } T\textit{-cut } C). \end{aligned}$$

Note that the halfspace defined by $\sum x_e \geq p/2$ contains this polyhedron, and the intersection of the T-join polyhedron with the hyperplane $\sum x_e = p/2$ is just the perfect matching polytope. From this observation we see that Theorem 7.3.4 could be easily deduced from this theorem. We shall show here, however, that Theorem 7.3.6 can be deduced from the "elementary" minimax result on T-joins (Theorem 6.5.10). This contrasts considerably with the case of the matching polytope, where the combinatorial minimax theorem (Berge's Formula 3.1.14) does not suffice to prove the characterization of facets.

PROOF. For the time being, let P denote the polytope described by (i″) – (ii″). It is trivial that $TJ(G) \subseteq P$. So we must prove that if $\boldsymbol{a} \cdot \boldsymbol{x} \geq b$ is any valid inequality for $TJ(G)$, then it is also valid for P. It is clear that we may assume that $\boldsymbol{a}$ is an integral vector and b is an integer.

Let $\boldsymbol{a} = (a_e : e \in E(G))^T$. From the fact that $TJ(G)$ contains all vectors majorizing a T-join and $\boldsymbol{a} \cdot \boldsymbol{x} \geq b$ is a valid inequality for all these vectors, it follows that $a_e \geq 0$.

Let us construct a graph G' by subdividing every line e of G by $a_e - 1$ new points. (If $a_e = 0$ then we contract the line.) Let T' be the set of points in G' which arises from T by these subdivisions and contractions. (Recall that a point is in set T' if and only if precisely one of its two parents belonged to T.) We will show:

$$r(G', T') \geq b. \tag{7.3.11}$$

In fact, let J' be a minimum T-join in G'. Trivially, J' contains all or none of the lines of G' coming from a line of G. Thus J' corresponds in a natural way to a T-join J of G. Let $\boldsymbol{x}_1$ be the incidence vector of J. Then $\boldsymbol{x}_1$ is a point in $TJ(G)$ and so

$$b \leq \boldsymbol{a} \cdot \boldsymbol{x}_1 = \boldsymbol{a}(J) = |J'| = r(G', T').$$

This proves (7.3.11).

By Theorem 6.5.10, $\nu_2(G', T') \geq 2b$ and so G' contains a 2-packing of $2b$ T'-cuts. These T'-cuts yield in G a list $\{C_1, \ldots, C_{2b}\}$ of T-cuts such that every line e is contained in at most $2a_e$ of them. Hence for any point $\boldsymbol{x} \in P$,

$$2\sum_{i=1}^{q} a_e x_e \geq \sum_{j=1}^{2b}\Big(\sum_{e \in C_j} x_e\Big) \geq 2b,$$

that is, $\boldsymbol{a} \cdot \boldsymbol{x} \geq b$ is a valid inequality of the polyhedron P as well. ∎

7.3.7. EXERCISE. Derive Theorem 7.3.4 from Theorem 7.3.6. Then derive the theorem characterizing the matching polytope (Theorem 7.3.1) from Theorem 7.3.4.

7.3.8. EXERCISE. Prove that the dual of the linear program "minimize $\mathbf{c} \cdot \mathbf{x}$ subject to (i″) and (ii″) of Theorem 7.3.6," in the case where $\mathbf{c}$ is integral, has a half-integral optimal dual solution. (N.B.: This is not so easy!)

7.3.9. EXERCISE. Deduce the T-join Polyhedron Theorem 7.3.6 using the perfect matching polytope Theorem 7.3.4. (Hint: let H be the complete graph on T, and consider the distance between x and y in G as the weight of the line xy of H. Look for a perfect matching in H with minimum weight. The union of minimum paths of G, connecting those pairs of points of T which are matched by this minimum weight perfect matching in H, will be a minimum T-join in G.)

7.3.10. EXERCISE. (Chvátal (1975).) Two matchings are adjacent as vertices of the matching polytope if and only if one arises from the other by switching on an alternating cycle or on an alternating path.

BOX 7D. Cutting Planes

Edmonds' Theorem on the matching polytope suggests the following general question. Let us suppose that we want to find the convex hull of a set S of integral vectors in $\Re^n$. Quite often it is easy to find a system of inequalities whose *integral* solutions are exactly the vectors in S. For example, when S is the set of incidence vectors of matchings, the inequalities (7.1.5) defining the fractional matching polytope have only the matchings as integral solutions. In other words, we are given a polytope P and we would like to determine the convex hull of integral points inside P.

The following general procedure to obtain this convex hull was found by Chvátal (1973a). It is closely related to Gomory's Algorithm for solving integer linear programs (1958, 1963). In fact, the result could be proved by analyzing Gomory's Algorithm.

Let

$$\mathbf{a}_i \cdot \mathbf{x} \leq b_i \quad (i = 1, \ldots, m) \tag{7.3.12}$$

be a system of inequalities. Let $\mathbf{c} \cdot \mathbf{x} \leq d$ be an inequality which follows from the given system for every real vector $\mathbf{x}$. (By the Farkas Lemma, this happens if and only if there exist non-negative reals y_i such that $\mathbf{c} = \sum y_i \mathbf{a}_i$ and $d = \sum y_i b_i$.) Assume, moreover, that $\mathbf{c}$ is an integral vector. Then the inequality

$$c \cdot x \leq \lfloor d \rfloor \tag{7.3.13}$$

is called a **cut** derived from the original system. Inequality (7.3.13) does not follow from the inequalities (7.3.12) for *every* vector x, but if x is an *integral* vector satisfying (7.3.12), then it also satisfies (7.3.13). Geometrically, (7.3.13) describes a halfspace which does not contain P,

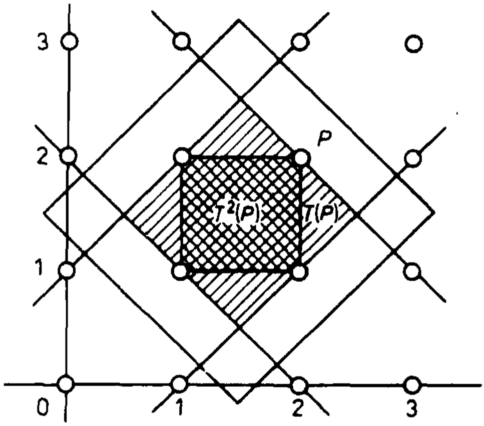


FIGURE 7D.1.

but the portion of P which is cut off by (7.3.13) certainly does not contain any integral points.

We call the set

$$T(P) = \{x \mid c \cdot x \leq d \text{ for every cut derived from (7.3.12)}\}$$

the **Gomory-Chvátal truncation** of P. It is clear that $T(P)$ is a convex set. In fact, it is a polyhedron.

We may consider $T(T(P)) = T^2(P)$, which is the set of vectors satisfying all cuts derived from valid inequalities for $T(P)$, $T^3(P) = T(T(T(P)))$, etc. Figure 7D.1 shows an example in which $P \neq T(P) \neq T^2(P)$.

7D.1. THEOREM. *Assume that P is bounded. Then there exists an integer $k \geq 0$ such that $T^k(P)$ is the convex hull of integral points inside P.* ■

Schrijver (1980b) showed that if the inequalities defining P have rational coefficients, then the boundedness of P need not be assumed.

The difficulty in applying the above theorem is that it may take an enormous number of cuts to obtain a facet of the convex hull of integral points in P. In this respect, Theorem 7.3.1 is interesting because it tells us that if we start with the fractional matching polytope (which has a very simple and natural description by inequalities), then generating cuts just *once* gives us the facets of the matching polytope; that is, $T(FM(G)) = M(G)$. In fact, let S be any odd subset of the points of G. Sum the inequalities

$$x(\nabla(v)) \leq 1$$

for every $v \in S$, and subtract from this sum all inequalities

$$x_e \geq 0$$

for lines e connecting S to $V(G) - S$. After division by 2 we get

$$x(E(S)) \leq \frac{|S|}{2}.$$

If we take the integer part on the left hand side, we get the inequality in list (iii) corresponding to the set S. So by Theorem 7.3.1, all facets of $M(G)$ occur among the cuts generated in one step. (In fact, this result of Edmonds motivated Chvátal in formulating the results discussed in this box.)

7.4. Chromatic Index

Recall that a **line coloration** of a graph G is a coloration of the lines such that any two adjacent lines have different colors. The **chromatic index** $\chi_e(G)$ of G is the minimum number of colors in a line coloration.

Ever since Tait (1878-80b, 1880) proved that the Four Color Theorem is equivalent to the statement that every 2-connected planar cubic (3-regular) graph has chromatic index 3, graph theorists have been very much interested in the study of the chromatic index. So how can we compute this parameter? We have already discussed the theorem of König (1916a, 1916b), which asserts that the chromatic index of a bipartite graph is equal to its maximum degree. On the negative side, Holyer (1981) proved that the problem of finding the chromatic index of a 3-regular non-bipartite graph is NP-complete.

For non-bipartite graphs, the most important result on chromatic index is the theorem of Vizing (1964, 1965).

7.4.1. THEOREM. (*Vizing's Theorem*). *The chromatic index of a simple graph is at most one larger than the maximum degree.*

The proof given here is due to Ehrenfeucht, Faber and Kierstead (1982). We start with a lemma which will enable us to augment partial line colorations.

7.4.2. LEMMA. *Assume that G is a simple graph with maximum degree k. Let $e_1, \ldots, e_r$ be some of the lines of G adjacent to the same point v, and suppose $G - e_1 - \cdots - e_r$ is k-colored so that for every e_i there is a color missing at both endpoints, and for all but one of the lines e_i there are at least two such colors. Then G is k-line colorable.*

PROOF. We use induction on r. For $r = 1$ the assertion is trivial. Assume that $r \geq 2$. Let F_i be the set of colors from $1, \ldots, k$ which are missing at both endpoints of e_i. We may assume that the indices are chosen so that $|F_1| \geq 1$ and $|F_i| \geq 2$ for $i \geq 2$. Choose $F'_i \subseteq F_i$ with $|F'_1| = 1$ and $|F'_i| = 2$ for $i = 2, \ldots, r$.

Case 1. There exists a color α which occurs in exactly one set F'_i. Then color line e_i with color α. No color of F'_j appears this way at the endpoints of e_j ($j \neq i$), and so the k-line colorability of G follows by the induction hypothesis.

Case 2. No color occurs in exactly one set F'_i. Then there is a color α which is missing at v, but does not belong to any of the sets F'_i. In fact, there are at most $k-r$ lines incident with v which have been colored, so there are at least r colors missing at v. If each of these occurs in one of the sets F'_i, then by assumption each of them occurs in two sets. But $\sum_{i=1}^{r} |F'_i| = 2r - 1$, so this is impossible.

Let $e_1 = uv$, $F'_1 = \{\beta\}$ and consider the alternating $\alpha - \beta$ path P starting from u. If we interchange colors α and β along this path, we get another valid coloration of $G - e_1 - \cdots - e_r$. Upon coloring e_1 with color α, we get a valid coloration of $G - e_2 - \cdots - e_r$. Furthermore, the set of colors occurring has changed only at three points: v, where α too now occurs; u, where both α and β now occur; and the second endpoint of the path P. Thus all but at most one of the lines e_i ($2 \leq i \leq r$) still miss both colors in F'_i at both of their endpoints, and the exceptional line (which must then be incident with the second endpoint of P) still misses one of the elements of F'_i. Thus the k-line colorability of G follows, again by the induction hypothesis. ∎

PROOF (of Theorem 7.4.1). Let G be a simple graph with maximum degree $k - 1$. We show by induction on $|V(G)|$ that G is k-line colorable.

Let $v \in V(G)$ and let $e_1, \ldots, e_r$ $(r \leq k-1)$ be the lines of G incident with v. Consider a k-line coloration of $G-v$. All former neighbors of v then have degree at most $k-2$ in $G-v$, and hence at least two colors are missing at each of them. So we may apply Lemma 7.4.2 to see that G is k-line colorable. ∎

The above proof of Vizing's Theorem provides a polynomial algorithm for a $(k+1)$-line coloration of a graph with maximum degree k.

7.4.3. EXERCISE. Prove that if G is a simple graph with maximum degree k and if every cycle in G contains a point with degree at most $k-1$, then G is k-line colorable.

Let us point out that it has been conjectured that it is sufficient in the above exercise to require only that every *odd* cycle contain a point of degree at most $k-1$.

7.4.4. EXERCISE. Let G be a simple graph with maximum degree $k-1$ and M a maximal matching in G. Then G has a k-line coloration in which M is one of the color classes.

7.4.5. EXERCISE. Show by an example that Vizing's Theorem, in the form given above, does not hold for non-simple graphs.

Vizing's Theorem admits an extension to non-simple graphs. In a somewhat different direction on the other hand, interesting — although rather restrictive — conditions have been discovered under which the chromatic index is equal to the maximum degree. We will not go into the details here, but refer the reader to the monograph of Fiorini and Wilson (1977) and to Gupta (1974).

What concerns us here is a linear programming approach to the chromatic index problem and its relationship with the matching problem. Observe that a line coloration of a graph G may be defined as a partition of its line set into matchings. In the spirit of the present chapter, we may rephrase this as follows. Let us introduce a variable y_M for every matching M, and as before, let $\boldsymbol{f}_M$ denote the incidence vector of the matching M. Then a line coloration may be viewed as a 0–1 solution to the system of equations

$$\sum_M y_M \boldsymbol{f}_M = \mathbf{1}. \tag{7.4.1}$$

Moreover, the number of colors is counted by the sum

$$\sum_M y_M. \tag{7.4.2}$$

So the chromatic index is the minimum of (7.4.2) taken over all 0–1 vectors satisfying (7.4.1).

Similarly as before, we may consider the "linear relaxation" of this problem. That is, we may want to determine the minimum of (7.4.2) subject to (7.4.1) and to the constraint

$$y_M \geq 0. \tag{7.4.3}$$

Let this minimum be called the **fractional chromatic index** of G and denote it by $\chi_e^*(G)$.

Obviously, $\chi_e^*(G) \leq \chi_e(G)$. We are going to show that a minimax formula for $\chi_e^*(G)$ can be derived from Edmonds' description of the matching polytope. This application of the matching polytope was hinted at by Edmonds and worked out by Stahl (1979) and Seymour (1979a).

7.4.6. THEOREM. *For every graph G,*

$$\chi_e^*(G) = \max\left\{\Delta(G), \max\left\{\frac{2|E(S)|}{|S|-1} \mid S \subseteq V(G), |S| \text{ odd}\right\}\right\}.$$

PROOF. Let us consider any solution of (7.4.1) and (7.4.3), and let $w = \sum_M y_M$. Set $y'_M = (y_M)/w$. Then we have

$$\sum_M y'_M = 1, \qquad \sum_M y'_M \boldsymbol{f}_M = (1/w)\mathbf{1},$$

that is, the vector $(1/w)\mathbf{1}$ is a convex combination of matchings. So what we want to determine is the largest value of w for which

$$(1/w)\mathbf{1} \in M(G).$$

By Theorem 7.3.1, this is equivalent to saying that

$$\begin{array}{lll} \text{(i)} & 1/w \geq 0 & \\ \text{(ii)} & (1/w)\deg(v) \leq 1 & (v \in V(G)) \\ \text{(iii)} & (1/w)|E(S)| \leq (|S|-1)/2 & (S \subseteq V(G),\ |S| \text{ odd}). \end{array}$$

It is clear that the largest value of w satisfying these inequalities is the one given in the statement of the theorem. ■

The most interesting special case arises when the graph is r-regular. If it is also simple, then Vizing's Theorem implies that the chromatic index is either r or $r+1$. Furthermore, if the chromatic index is r, then in every optimum line coloration every color forms a perfect matching. Hence such a line coloration is only possible if the number of points is even. An even stronger necessary condition which can be derived is that for every subset S with odd cardinality, at least r lines connect S to $V(G)-S$. These conditions are not sufficient, however, as the Petersen graph shows. They *are*, however, sufficient for *something*! The next result is also due to Seymour (1979a).

7.4.7. COROLLARY. *Let G be an r-regular graph such that for every subset $S \subseteq V(G)$ having odd cardinality, at least r lines connect S to $V(G)-S$. Then $\chi_e^*(G)=r$.* ■

The verification that the formula in Theorem 7.4.6 gives r under these conditions is left to the reader.

Let us remark that if G has an even number of points, then $(r-1)$-line-connectivity of G implies that the condition of Corollary 7.4.7 is satisfied. Another formulation of this result is the corollary below, which follows upon noticing that (7.4.1)–(7.4.3) have a *rational* optimum solution, and then multiplying by the least common denominator of the y_M's.

7.4.8. COROLLARY. *Under the conditions of Corollary 7.4.7, there exists an integer $t>0$ such that if each line of G is replaced by t parallel lines, then the resulting graph G' has chromatic index tr.* ■

The least value of t in Corollary 7.4.8 is not known. A conjecture of Fulkerson states that if $r=3$, then $t=2$ is always sufficient. (The question as to which 3-regular graphs admit $t=1$ is equivalent to asking which 3-regular graphs have chromatic index 3, which is an NP-complete problem). The best result in the direction of Fulkerson's Conjecture is due to Seymour (1979a), who proved that if $r=3$ then t can be chosen to be any even integer, provided it is large enough.

How much information on the chromatic index did we obtain by determining the fractional chromatic index? The answer to this question is not known. It is conjectured by Seymour (1979a) that $\chi_e(G) \le \chi_e^*(G)+1$ holds for every graph G. This conjecture would imply most of the generalizations of Vizing's Theorem referred to above.

Another "relaxation" of the chromatic index problem was proposed by Seymour (1979a). Let G be an r-regular graph. Then the fact that it has chromatic index r may also be formulated as follows: the system of linear equations

$$\sum y_P \boldsymbol{f}_P = \mathbf{1} \tag{7.4.4}$$

(where P ranges over all perfect matchings of G) has a 0–1 solution. We may relax this and ask if this system has an *integral* solution. So the integral solvability of equation (7.4.4) is a necessary condition for the r-line colorability of G. The Petersen graph does not even satisfy this necessary condition; that is, if G is the Petersen graph then (7.4.4) has no integral solution. On the other hand, Seymour proved that if G is a 3-regular 3-line-connected graph which contains no subdivision of the Petersen graph then (7.4.4) does have an integral solution.

One further relaxation of the chromatic index problem is to ask for a solution to equation (7.4.4) over $GF(2)$. This problem can be phrased as follows. Is there a list of perfect matchings of G containing every line an odd number of times? Again, no such list exists for the Petersen graph.

Unfortunately, very little is known about these two problems, although there is some hope that a satisfactory theory for them can be developed (cf. Box 7E).

BOX 7E. Good Characterizations other than the Farkas Lemma

The problem of finding a solution to equation (7.4.4) is potentially easier then the chromatic index problem. The reason for this is the following theorem, which is analogous to the Farkas Lemma.

7E.1. THEOREM. *Let A be an $n\times m$ matrix and $\boldsymbol{b}$ an n-dimensional vector. Then $A\boldsymbol{x} = \boldsymbol{b}$ has an integral solution if and only if there does not exist a vector $\boldsymbol{y}$ such that $A^{\mathrm{T}}\boldsymbol{y}$ is an integer vector, but $\boldsymbol{b}\cdot\boldsymbol{y}$ is not an integer.* ■

An analogous, but much more trivial, condition holds for solvability over $GF(2)$:

7E.2. THEOREM. *Let A be an $n \times m$* 0–1 matrix and suppose $\boldsymbol{b}$ is an n-dimensional 0–1 vector. Then $A\boldsymbol{x} = \boldsymbol{b}$ has a solution over $GF(2)$ if and only if there does not exist an m-dimensional vector $\boldsymbol{y}$ over $GF(2)$ such that $A^{\mathrm{T}}\boldsymbol{y} = \mathbf{0}$, but $\boldsymbol{b}^{\mathrm{T}}\cdot\boldsymbol{y} = 1$. ■

Of course this last theorem is just the well-known condition on the solvability of a system of linear equations over a field. Let us remark that no similar condition for the 0–1 solvability or non-negative integral solvability of a system of linear equations is known.

Now these two theorems enable us to formulate the two relaxations of the chromatic index problem mentioned earlier as follows:

Equation (7.4.4) has an integral solution if and only if there does not exist a weighting of the lines such that the weight of every perfect matching is an integer, but the total weight of lines is not.

Equation (7.4.4) has solution over $GF(2)$ if and only if there is no subgraph which meets every perfect matching in an even number of lines, but has itself an odd number of lines.

These conditions are, however, not *good* characterizations, since to check (or prove) that for a given weighting of the lines every perfect matching has integral weight is an unsolved problem (and similarly for the second condition). We remark, however, that the complexity of these problems is not known and it may well turn out that they can be solved polynomially!

7.5. Fractional Matching Polytopes and Cover Polyhedra

Recall that the fractional matching polytope $FM(G)$ of a graph G was defined as the solution set of the system of linear inequalities:

$$\boldsymbol{x} \geq \mathbf{0} \tag{7.5.1}$$

$$\boldsymbol{x}(\nabla(v)) \leq 1 \quad (\text{for each } v \in V(G)). \tag{7.5.2}$$

Since this polytope is defined in terms of inequalities, in order to give a good characterization of it we must describe its vertices. The following theorem is due to Balinski (1965).

7.5.1. THEOREM. *Let $\boldsymbol{w}$ be a basic 2-matching of G. Then $\boldsymbol{w}/2$ is a vertex of $FM(G)$, and every vertex of $FM(G)$ is of this form.*

PROOF. The point $\boldsymbol{w}/2$ is the unique point of $FM(G)$ which maximizes the linear objective function having coefficients $c_e = w_e$ if $w_e > 0$, and $c_e = -1$ if $w_e = 0$. Hence it is a vertex.

Conversely, let $\boldsymbol{y} = (y_e : e \in E(G))$ be a vertex of $FM(G)$. First we show that $\boldsymbol{y}$ has half-integral entries. Let $2t$ be a common denominator of the entries of $\boldsymbol{y}$. Define an auxiliary graph G' by replacing each line e of G by $2ty_e$ parallel copies. The fact that $\boldsymbol{y}$ satisfies inequality (7.5.2) implies that G' has maximum degree at most $2t$. Embed G' in a $2t$-regular graph G''. Then by Petersen's Theorem 6.2.3, G'' is the union of t 2-factors. Hence G' is the union of t subgraphs, each having maximum degree at most 2. This implies that G has t 2-matchings $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_t$ such that $2t\boldsymbol{y} = \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_t$. Let $\boldsymbol{x}_i = \boldsymbol{w}_i/2$. Then

$$\boldsymbol{y} = \frac{\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_t}{t}$$

and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t \in FM(G)$. Since by hypothesis $\boldsymbol{y}$ is a vertex of $FM(G)$, this implies that $\boldsymbol{y} = \boldsymbol{x}_1 = \cdots = \boldsymbol{x}_t$. Thus $\boldsymbol{y}$ is indeed half-integral and so $2\boldsymbol{y}$ is a 2-matching. The fact that this 2-matching is basic follows easily by noticing that non-basic 2-matchings can be represented as convex combinations of basic ones. ■

There is a very similar theorem which describes the fractional point cover polyhedron $FPC(G)$. Recall that this polyhedron is defined as the solution set of the linear inequalities

$$\boldsymbol{x} \geq \mathbf{0}$$
$$x_u + x_v \geq 1 \quad \text{(for each } uv \in E(G)).$$

Again, recall that our purpose is to describe the vertices of these two polyhedra. The following theorem has been attributed to Lorentzen (1966) and, independently, to Balinski and Spielberg (1969). Proofs may be found in Trotter (1973) and in Nemhauser and Trotter (1974).

7.5.2. THEOREM. *Let ϕ be a basic 2-cover of the graph G. Then $\phi/2$ is a vertex of the polyhedron $FPC(G)$. Every vertex of $FPC(G)$ is of this form.* ■

7.5.3. EXERCISE. Deduce the main theorem about 2-matchings — that is, Theorem 6.1.4 — from Theorems 7.5.1 and 7.5.2.

7.5.4. EXERCISE. Find a minimal system of linear inequalities that defines the cone generated by the rows of the incidence matrix of a graph G.

7.6. The Dimension of the Perfect Matching Polytope

In this section we study the dimension of the perfect matching polytope $PM(G)$ of a graph G. As remarked earlier, this problem is also related to the problem of determining the facets of $PM(G)$, and we shall outline how the solution to the problem of finding the dimension of the perfect matching polytope can be used to help characterize its facets. Throughout this chapter we shall assume that G is a connected 1-extendable graph.

Determining the dimension of a polytope is equivalent to obtaining a linear basis for all those equations which hold true for all points in the polytope. These always include the following p equations:

$$\boldsymbol{x}(\nabla(v)) = 1 \qquad \text{(for every } v \in V(G)). \tag{7.6.1}$$

Our first task is to determine the dimension of the solution set of (7.6.1).

7.6.1. LEMMA. *Let G be a connected graph. Then the dimension of the solution space of system* (7.6.1) *is $q-p+2$, if G is bipartite and is $q-p+1$, if G is non-bipartite.*

PROOF. The assertion is equivalent to saying that the matrix composed of the rows consisting of the left hand sides of (7.6.1), that is, the point-line incidence matrix of G, has rank $p-1$ if the graph is bipartite and rank p if the graph is non-bipartite. To see this, note that the point-line incidence matrix has rank $\leq p$ (since it has p rows), and in the bipartite case there is a linear relation among the rows: the sum of rows corresponding to one color class is equal to the sum of rows corresponding to the other. Thus the rank of the point-line incidence matrix in the bipartite case is at most $p-1$.

To show that the rank is not less than the value claimed, consider any linear relation between the rows of the point-line incidence matrix. This is best visualized as an assignment of real weights to the points of G, not all 0, such that the sum of weights of the two endpoints of any line is 0. Let a point v have weight $w \neq 0$. Then all its neighbors have weight $-w$. Continuing, we see that every point has weight w or $-w$ and that every line connects a point with weight w to a point with weight $-w$. Hence G is bipartite and the linear relation is a scalar multiple of the one already discovered. This proves the lemma. ∎

It might appear at this point that our goal of determining the dimension of $PM(G)$ is achieved, since the other constraints in the description of the perfect matching polytope (Theorem 7.3.4) are inequalities. However, this is not so! Inequalities may imply further equations which are independent of (7.6.1), and in fact, they do in general! So the values given in Lemma 7.6.1 are only *upper bounds* for the dimension of $PM(G)$. In the bipartite case, however, the bound turns out to be the exact value. Although this will follow from a later discussion, we give a simple direct proof of this fact here based on the ear decomposition theorem for elementary bipartite graphs (Theorem 4.1.6).

7.6.2. THEOREM. *The dimension of the perfect matching polytope of an elementary bipartite graph G is $q-p+1$.*

PROOF. As remarked above, it suffices to show that $\dim PM(G) \geq q-p+1$, that is, that G contains $q-p+2$ affinely independent perfect matchings. Let

$$G = P_0 + P_1 + \cdots + P_k$$

be an ear decomposition of G, where $P_0 = \{e_0\}$, $k = q-p+1$ and P_{i+1} is an odd path which has precisely its endpoints in common with $P_0 + P_1 + \cdots + P_i$ $(i = 0, \dots, k-1)$. Let e_i be one of the endlines of P_i and let M_i be the matching consisting of those lines of P_i which are at an odd distance from its endpoints (thus $e_i \notin M_i$). Let M_i' be a perfect matching of $P_0+P_1+\cdots+P_i$ containing e_i. Consider the following perfect matchings of G:

$$M_i' \cup M_{i+1} \cup \cdots \cup M_k \qquad (i = 0, \dots, k).$$

These matchings are affinely — even linearly — independent, since if we restrict their incidence vectors to the entries corresponding only to the lines $e_0, e_1, \dots, e_k$, we obtain a triangular matrix with 1's on the diagonal. Thus G contains $k+1 = q-p+2$ affinely independent perfect matchings. ∎

Note that we can also derive Theorem 7.6.2 from Theorem 7.3.4 by noticing that there are no factor-critical subgraphs since the graph is bipartite and no non-negativity constraint holds with equality since the graph is elementary.

7.6.3. EXERCISE. Show that for perfect matchings, "affinely independent" and "linearly independent" are equivalent.

We now turn our attention to the case of non-bipartite graphs. This general case was first solved by Naddef (1982).

It is easy to obtain a good characterization of $\dim PM(G)$. To show that $\dim PM(G) \geq k$, it suffices to exhibit $k+1$ affinely independent perfect matchings. To show that $\dim PM(G) \leq k$, it suffices to exhibit $q-k$ linearly independent equations which are each satisfied by every perfect matching. To exhibit that a given linear equation $\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{x} = \alpha$ is satisfied by all perfect matchings, it suffices to show that the inequalities $\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{x} \leq \alpha$ and $\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{x} \geq \alpha$ are satisfied by all perfect matchings. This, in turn, can be shown using Theorem 7.3.4 and linear programming duality.

Naddef's main contribution is a much more transparent form of this dual solution. To state his characterization of the dimension of $PM(G)$, we need some definitions.

Let G be a graph on an even number of points and C an odd cut of G. We say that C is **tight** if every perfect matching of G intersects C in exactly one line (of course, every perfect matching contains *at least one* line of C). (See Figure 7.6.1.)

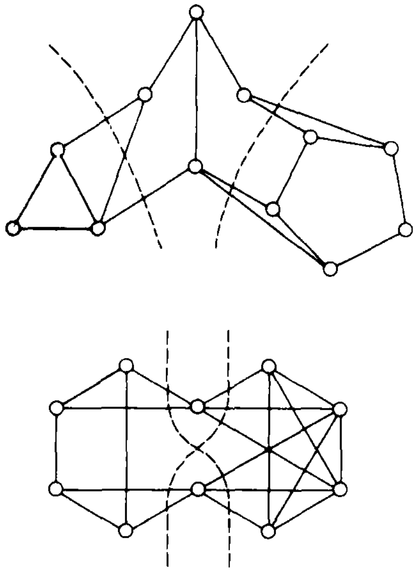


FIGURE 7.6.1. Tight odd cuts in 1-extendable graphs

Let us remark immediately that the problem as to whether a given odd cut is tight is a special case of the weighted matching problem for G. Just take weights 1 for the lines in C and weights 0 for the other lines, and look for a matching with maximum weight. We may also reduce this problem to a polynomial number of unweighted matching problems by noting that an odd cut C is tight if and only if no pair of lines of C extends to a perfect matching of G; that is, if and only if $G-V(e_1)-V(e_2)$ does not contain a perfect matching for every pair e_1, e_2 of disjoint lines of C.

Let $\mathcal{F}$ be a set of non-crossing non-trivial odd cuts. Let $C \in \mathcal{F}$ and let $S_1(C)$, $S_2(C)$ be the two shores of C. Contract $S_1(C)$ to a single point.

Also contract every $S(C')$ such that $C' \in \mathcal{F} - \{C\}$ and $S(C') \subseteq S_2(C)$. The graph obtained this way from G will be called an **$\mathcal{F}$-contraction** of G. We say that the family $\mathcal{F}$ has the **odd cycle property** if no $\mathcal{F}$-contraction of G is bipartite. Note that this property implies, in particular, that $\mathcal{F}$ consists of distinct cuts. Also note that if we add all trivial cuts to the family, this property is not affected.

7.6.4. THEOREM. *Let G be a 1-extendable graph. Then*

$$\dim PM(G) = q - \max\{|\mathcal{F}| \mid \mathcal{F} \text{ is a non-crossing family of tight odd cuts with the odd cycle property}\}. \qquad \blacksquare$$

Naddef also derives the following result by similar methods.

7.6.5. THEOREM. *Each inclusionwise maximal non-crossing family of tight odd cuts with the odd cycle property has the same cardinality.* ■

We shall not prove Theorem 7.6.4 in this book, but rather the following more recent result, which makes use of the structure theory of Chapter 5. This characterization of the dimension of the perfect matching polytope was found by Edmonds, Lovász and Pulleyblank (1982).

Recall that in Chapter 5 we described how to "paste together" every 1-extendable graph from bicritical and elementary bipartite graphs, and showed that bicritical graphs are "built up" from 3-connected bicritical graphs. These 3-connected bicritical graphs will be called the **bricks** of G.

7.6.6. THEOREM. *For every 1-extendable graph G,*

$$\dim PM(G) = q - p + 1 - k,$$

where k is the number of bricks of G.

From this result, Theorem 7.6.4 can be easily deduced. We leave this to the reader as an exercise.

The most difficult case in the proof of Theorem 7.6.6 occurs when G is itself a brick. In this case the theorem asserts that the dimension of $PM(G)$ is equal to the upper bound provided by Lemma 7.6.1. (Another such case is that of bipartite graphs, which was settled in Theorem 7.6.2.) To settle the case for bricks, we need several lemmas.

Let $\mathcal{F}$ be any family of subsets of $E(G)$ where we allow multiple membership. For $e \in E(G)$, we denote by $\mathcal{F}_e$ the family of members of $\mathcal{F}$ containing e. We say that a family $\mathcal{F}'$ **majorizes** the family $\mathcal{F}$, if

$|\mathcal{F}'_e| \geq |\mathcal{F}_e|$ for every $e \in E(G)$. Here in $|\mathcal{F}|$ each element of $\mathcal{F}$ is counted as often as it appears in $\mathcal{F}$.

7.6.7. LEMMA. *Let G be a graph with a perfect matching and let C be an odd cut of G. Then C is tight if and only if there exists a family $\mathcal{F}$ of odd cuts containing C and another family $\mathcal{F}'$ of trivial cuts such that $|\mathcal{F}| = |\mathcal{F}'|$ and $\mathcal{F}'$ majorizes $\mathcal{F}$.*

PROOF. Assume first that there exist families $\mathcal{F}$ and $\mathcal{F}'$ with the given properties. Let M be any perfect matching. Then by the hypothesis that $\mathcal{F}'$ majorizes $\mathcal{F}$, we have

$$\sum_{A\in\mathcal{F}'} |A\cap M| = \sum_{e\in M} |\mathcal{F}'_e| \geq \sum_{e\in M} |\mathcal{F}_e| = \sum_{A\in\mathcal{F}} |A\cap M|$$

Here the left hand side is $|\mathcal{F}'|$, while the right hand side is at least $|\mathcal{F}|$. Since $|\mathcal{F}| = |\mathcal{F}'|$, it follows that we must have equality throughout and, in particular, $|C\cap M| = 1$. Since this holds for every perfect matching M, it follows that C is tight.

Conversely, assume that C is tight. Then the maximum of $\boldsymbol{x}(C)$ over all $\boldsymbol{x} \in PM(G)$ is 1. Using the description of $PM(G)$ given in Theorem 7.3.4, we can write this as a linear program:

$$\begin{array}{lrl}
\text{maximize} & \boldsymbol{x}(C) & \\
\text{subject to} & \boldsymbol{x} \geq \mathbf{0} & \\
 & \boldsymbol{x}\nabla(v)) = 1 & \quad\text{(for every point } v\text{)} \\
 & \boldsymbol{x}(A) \geq 1 & \quad\text{(for every odd cut } A\text{)}.
\end{array}$$

We know that the maximum of this objective function is 1. Consider an optimum dual solution: we shall have a variable y_v for every point v and a variable z_A for every odd cut A, and the following constraints:

$$z_A \geq 0,$$
$$\sum_{v\in V(e)} y_v - \sum_{e\in A} z_A \geq \begin{cases} 1, & \text{if } e \in C, \\ 0, & \text{if } e \in E(G) - C. \end{cases}$$

By the Duality Theorem of Linear Programming, the dual objective function has value 1 for an optimum dual solution:

$$\sum_v y_v - \sum_A z_A = 1.$$

If we let $z'_A = z_A + 1$ when $A = C$ and $z'_A = z_A$ otherwise, then

$$z'_A \geq 0, \quad z'_C > 0, \tag{7.6.2}$$

$$\sum_{v \in V(e)} y_v \geq \sum_{e \in A} z'_A \quad \text{(for all } e\text{)}, \tag{7.6.3}$$

and

$$\sum_v y_v = \sum_A z'_A. \tag{7.6.4}$$

We may assume in addition that $y_v \geq 0$ for every point v, since adding the same number to y_v and $z_{\nabla(v)}$ does not influence the validity of (7.6.2), (7.6.3) and (7.6.4). By similar reasoning, we may assume that every y_v and z_A is an integer, since we may multiply by their common denominator without interfering with (7.6.2), (7.6.3) or (7.6.4).

If we let the family $\mathcal{F}$ consist of z_A copies of A for every odd cut A, and let the family $\mathcal{F}'$ consist of y_v copies of $\nabla(v)$ for every point v, then $\mathcal{F}'$ majorizes $\mathcal{F}$ by (7.6.3), $|\mathcal{F}'| = |\mathcal{F}|$ by (7.6.4) and $C \in \mathcal{F}$ by (7.6.2). ∎

Our next step is to prove the theorem for 3-connected bicritical graphs. In this case the theorem asserts that the dimension of $PM(G)$ is equal to the upper bound provided by Lemma 7.6.1. So this case of the theorem will follow if we can show that none of the inequalities (i') and (iii') describing $PM(G)$ in Theorem 7.3.4 holds with equality for all perfect matchings. This is evident for the inequalities in (i'), since every bicritical graph is 1-extendable. Thus it suffices to show the following.

7.6.8. LEMMA. *A 3-connected bicritical graph has no non-trivial tight cut.*

PROOF. By Lemma 7.6.7, it suffices to prove the following: if G is a 3-connected bicritical graph, $\mathcal{F}$ a family of odd cuts, $\mathcal{H}$ a family of trivial cuts, $|\mathcal{F}| = |\mathcal{H}|$ and $\mathcal{H}$ majorizes $\mathcal{F}$, then $\mathcal{F}$ also consists of trivial cuts. Suppose, to the contrary, that there exist such families $\mathcal{F}$ and $\mathcal{H}$, and $\mathcal{F}$ contains at least one non-trivial cut.

Let M be any perfect matching in G. Then

$$|\mathcal{H}| = \sum_{e \in M} \mathcal{H}(e) \geq \sum_{e \in M} \mathcal{F}(e) = \sum_{C \in \mathcal{F}} |M \cap C| \geq |\mathcal{F}|.$$

Thus we must have equality throughout. In particular, it follows that

$$\mathcal{F}(e) = \mathcal{H}(e) \tag{7.6.5}$$

for every line e, and also that all cuts in $\mathcal{F}$ are tight. (This of course is also a consequence of Lemma 7.6.6).

Next we show that we can assume that $\mathcal{F}$ consists of non-crossing cuts. This follows by the same "uncrossing" procedure as used in the proof of Lemma 6.5.13, but at one point some additional work is needed. Assume that C_1 and C_2 are two crossing cuts in $\mathcal{F}$. Choose a shore S_i of C_i such that $S_1 \cap S_2$ is odd. Then the cuts $C' = \nabla(S_1 \cap S_2)$ and $C'' \doteq \nabla(S_1 \cup S_2)$ are odd. Further, if $\mathcal{F}' = \mathcal{F} - \{C_1, C_2\} \cup \{C', C''\}$, then $\mathcal{H}$ majorizes $\mathcal{F}'$ and $|\mathcal{F}'| = |\mathcal{H}|$.

The only new feature is that we have to verify that $\mathcal{F}'$ still contains at least one non-trivial cut. We claim that at least one of C' and C'' is non-trivial. For if both of them are trivial then $S_1 \cap S_2 = \{u\}$ and $S_1 \cup S_2 = V(G) - v$, for some $u, v \in V(G)$. Since G is 3-connected, the pair $\{u, v\}$ does not separate G and so there must be a line x_1x_2 connecting $S_1 - u$ to $S_2 - u$. Then $\mathcal{F}'(x_1x_2) < \mathcal{F}(x_1x_2)$, which contradicts the fact that, by equation (7.6.5),

$$\mathcal{F}'(x_1x_2) = \mathcal{H}(x_1x_2) = \mathcal{F}(x_1x_2).$$

This contradiction proves that at least one of C' and C'' is non-trivial and so we can repeat the uncrossing procedure until we get two families $\mathcal{F}$, $\mathcal{H}$ of tight cuts such that $\mathcal{F}$ is non-crossing and contains at least one non-trivial cut, $\mathcal{H}$ consists of trivial cuts, and $\mathcal{F}(e) = \mathcal{H}(e)$ for every line e. (Incidently, this is the only point in the entire proof of the lemma where 3-connectivity of G is used.)

As a next trivial step, we note that we may delete those trivial cuts occurring in both $\mathcal{F}$ and $\mathcal{H}$ from both families without violating any of these properties, so we may assume that $\mathcal{F} \cap \mathcal{H} = \emptyset$.

Let $C \in \mathcal{F}$, and $e = xy \in C$. Then $\mathcal{F}(e) > 0$, so $\mathcal{H}(e) > 0$; that is, at least one of x and y is the center of a trivial cut in $\mathcal{H}$. Let, say, x be the center of a trivial cut in $\mathcal{H}$ and let S be the shore of C containing x. We may assume that C, e and x are chosen so that $|S|$ is as small as possible. Since the trivial cut with center x occurs in $\mathcal{H}$, it does not occur in $\mathcal{F}$, and so $|S| > 1$.

Let $S_1, \ldots, S_k$ be the maximal proper subsets of S which are shores of cuts in $\mathcal{F}$. Since $\mathcal{F}$ is non-crossing, $S_1, \ldots, S_k$ are mutually disjoint, and by the minimality of $|S|$, none of them contains x. Let $X = S - S_1 - \cdots - S_k$ and let Y be the set of those points of X which are centers of trivial cuts in $\mathcal{H}$. Thus $x \in Y$.

Let us now observe the following facts.

(a) No line connects S_i to S_j $(1 \le i < j \le k)$. In fact, at least one of the two endpoints of such a line would have to be the center of a trivial cut in $\mathcal{H}$ and then we could replace S by S_i or S_j, which would contradict the minimality of $|S|$.

(b) No line connects $X - Y$ to any S_i for the same reasons as given in (a).

(c) No line connects Y to $X - Y$ or two points in Y. In fact, such a line f would imply $\mathcal{H}(f) > 0$ and $\mathcal{F}(f) = 0$, contradicting (7.6.5).

(d) $|Y| = k + 1$. In fact, consider a perfect matching M containing the line e. Since each S_i is a shore of a tight cut, exactly one line of M connects S_i to $V(G) - S_i$. By (a), this line cannot end in any S_j $(j \neq i)$ or in $X - Y$. Furthermore, it cannot end in $V(G) - S$, since C is tight and it already contains the line $e \in M$. So exactly one line of M connects S_i to Y. By (c), the points in Y cannot be incident with any other line in M, except that x is incident with e. Hence the assertion follows.

(e) No line connects S_i to $V(G) - S, (1 \le i \le k)$. For suppose there is such a line f. Consider a perfect matching M containing f. The same argument as in (d) implies that $|Y| = k - 1$, a contradiction.

(f) Thus, $G - Y$ has at least $|Y| - 1$ odd components. Hence, by parity, $G - Y$ must have at least $|Y|$ odd components. This contradicts the hypothesis that G is bicritical, unless $|Y| = 1$. Thus by (d), $k = 0$; that is, S has no proper subsets which are shores of cuts in $\mathcal{F}$. So $X = S$ and $Y = \{x\}$.

(g) Since G is connected, there must be a line f connecting $S - x$ to $V(G) - (S - x)$. (Note that $S - x \neq \emptyset$ since $|S| > 1$. Let M be a perfect matching containing f. Let g be the line in M containing x. Obviously, both f and g belong to C. But this is a contradiction since C is tight. This completes the proof of Lemma 7.6.8. ∎

The last lemma we need for the proof will enable us to use induction. For every set $S \subseteq V(G)$ denote by $G \times S$ the graph obtained by contracting S to a single point.

7.6.9. LEMMA. *Let G be a 1-extendable graph, C, a tight cut in G and let S_1, S_2 be the two shores of C. Then*

$$\dim PM(G \times S_1) + \dim PM(G \times S_2) = \dim PM(G) + |C| - 1.$$

Before proving this lemma, let us remark that the hypothesis that C is tight implies that the image of a perfect matching of G under the contraction of S_i is a perfect matching of $G \times S_i$. Hence both $G \times S_1$ and $G \times S_2$ are 1-extendable.

PROOF. Let $\mathcal{M}_i$ be a maximum family of linearly (or affinely) independent perfect matchings in $G \times S_i$. For $e \in C$, let $\mathcal{M}_i^e$ denote the set of perfect matchings in $\mathcal{M}_i$ containing e, and select one $F_i^e \in \mathcal{M}_i^e$ for every i and e. (Since $G \times S_i$ is 1-extendable, this is possible). Consider the matchings $F_1 \cup F_2^e$ (where $F_1 \in \mathcal{M}_1^e$) and $F_1^e \cup F_2$ (where $F_2 \in \mathcal{M}_2^e$). The total number of such matchings is

$$\sum_{e \in C} |\mathcal{M}_1^e| + \sum_{e \in C} |\mathcal{M}_2^e| - |C| = |\mathcal{M}_1| + |\mathcal{M}_2| - |C|.$$

(Since the matchings $F_1^e \cup F_2^e$ were counted in both the first and the second sum, we had to subtract the number $|C|$ of such matchings.) Furthermore, these matchings are linearly independent. In fact, assume for each $F_1 \in \mathcal{M}_1$ and for each $F_2 \in \mathcal{M}_2$ that $\lambda(F_1)$ and $\lambda(F_2)$ are real numbers such that

$$\sum_{e \in C} \sum_{F_1 \in \mathcal{M}_1^e} \lambda(F_1) \boldsymbol{q}^{F_1 \cup F_2^e} + \sum_{e \in C} \sum_{\substack{F_2 \in \mathcal{M}_2^e \\ F_2 \neq F_2^e}} \lambda(F_2) \boldsymbol{q}^{F_1^e \cup F_2} = 0.$$

Restricting this relation to $E(G \times S_1)$, we obtain

$$\sum_{F_1 \in \mathcal{M}_1} \lambda(F_1) \boldsymbol{q}^{F_1} + \sum_{e \in C} \Bigg(\sum_{\substack{F_2 \in \mathcal{M}_2^e \\ F_2 \neq F_2^e}} \lambda(F_2) \Bigg) \boldsymbol{q}^{F_1^e} = 0. \tag{7.6.6}$$

Since the perfect matchings in $\mathcal{M}_1$ are linearly independent, it follows that $\lambda(F_1) = 0$, except possibly for the matchings $F_1 = F_1^e$ $(e \in C)$. Similarly, $\lambda(F_2) = 0$, except possibly for the matchings $F_2 = F_2^e$. But then in (7.6.6) the second term is 0 and hence $\lambda(F_1) = 0$ for every $F_1 \in \mathcal{M}_1$ and similarly $\lambda(F_2) = 0$ for every $F_2 \in \mathcal{M}_2$.

Thus we have $|\mathcal{M}_1| + |\mathcal{M}_2| - |C| = \dim PM(G \times S_1) + \dim PM(G \times S_2) + 2 - |C|$ linearly independent perfect matchings in G and hence

$$\dim PM(G) \geq \dim PM(G \times S_1) + \dim PM(G \times S_2) + 1 - |C|.$$

The reverse inequality follows by quite similar arguments. Hence equality holds and the proof is complete. ■

PROOF (of Theorem 7.6.6). We use induction on $|V(G)|$. If G is 3-connected and bicritical then the assertion follows from Lemma 7.6.8. Suppose that G is bicritical, but not 3-connected, and let $\{u, v\}$ be any separating set of G. Let G_1 and G_2 be two line-disjoint subgraphs of G such that $G_1 \cup G_2 = G$, $V(G_1) \cap V(G_2) = \{u, v\}$ and $|V(G_i)| > 2$. Let $S_1 = V(G_1) - u$ and $S_2 = V(G_2) - v$. Then $G \times S_i$ is bicritical. In fact, up to the multiplicity of the line uv, the graphs $G \times S_i$ are the same as the graphs obtained by applying the procedure for decomposing bicritical graphs into bricks, using the specific cutset $\{u, v\}$. Hence it follows that if $G \times S_i$ has k_i bricks then G has $k_1 + k_2$ such bricks.

The cut C determined by S_1 and S_2 is tight, as has been remarked before. Hence we may apply Lemma 7.6.9 and obtain

$$\begin{aligned}\dim PM(G) &= \dim PM(G \times S_1) + \dim PM(G \times S_2) + 1 - |C| \\ &= |E(G \times S_1)| - |V(G \times S_1)| + 1 - k_1 \\ &\qquad + |E(G \times S_2)| - |V(G \times S_2)| + 1 - k_2 - |C| \\ &= |E(G)| - |V(G)| + 1 - k,\end{aligned}$$

and we are done.

The case when G is not bicritical can be handled similarly. Suppose $S \in \mathcal{P}(G)$ and $|S| \geq 2$. Consider the connected components of $G - S$. If each of these is a singleton, then the assertion of the theorem follows from Theorem 7.6.2. So suppose that $G - S$ has a component H with more than one point. Let $S_1 = V(H)$ and $S_2 = V(G) - V(H)$. Then the cut C determined by S_1 and S_2 is tight. Hence $G \times S_1$ and $G \times S_2$ are 1-extendable. It is also clear by looking at the Brick Decomposition Procedure that if k_i denotes the number of bricks of $G \times S_i$, then the number of bricks of G is $k_1 + k_2$. Hence the theorem follows by the same computation as in the bicritical case. ∎

REMARK. Theorem 7.6.6 can be used to characterize the inequalities (i′) and (ii′) in Theorem 7.3.4 which correspond to facets of $PM(G)$. In fact, an inequality $\boldsymbol{a} \cdot \boldsymbol{x} \leq b$ defines a facet of a polytope P if and only if it is valid for all points in P and moreover,

$$\dim(P \cap \{\boldsymbol{x} \mid \boldsymbol{a} \cdot \boldsymbol{x} = b\}) = \dim P - 1.$$

(Cf. also Box 7F below.)

It is a fortunate circumstance that in the case of perfect matching polytopes of 1-extendable graphs, the polytopes $P \cap \{\boldsymbol{x} \mid \boldsymbol{a} \cdot \boldsymbol{x} = b\}$ are also perfect matching polytopes, or are at least closely related to such polytopes. In fact, if $\boldsymbol{a}\cdot\boldsymbol{x} \le b$ is one of the inequalities (i′), say $x_e \ge 0$, then $PM(G) \cap \{\boldsymbol{x} \mid x_e = 0\} = PM(G-e)$. If $\boldsymbol{a}\cdot\boldsymbol{x} \le b$ is one of the inequalities in (iii′), say $\boldsymbol{x}(C) \ge 1$ where C is an odd cut with shores S_1 and S_2, then $\dim(PM(G)\cap\{\boldsymbol{x} \mid \boldsymbol{x}(C) = 1\}) = \dim PM(G\times S_1) + \dim PM(G\times S_2) - |C| - 1$. This follows by an argument similar to the proof of Lemma 7.6.9. Thus in any case, to determine whether a defining inequality corresponds to a facet of $PM(G)$ is reduced to computing the dimension of the perfect matching polytopes of G and either $G-e$ or $G\times S_1$ and $G\times S_2$, which can be done easily by the results of this section. The explicit formulation of the results is somewhat awkward and the reader is referred to the paper of Edmonds, Lovász and Pulleyblank (1982).

7.6.10. COROLLARY. *If G is bicritical, then* $\dim PM(G) \ge p/2$. *Consequently, G contains at least $p/2+1$ perfect matchings.*

PROOF. If G is a brick then by Theorem 7.6.6, $\dim PM(G) = q - p \ge (3/2)p - p = p/2$, since the degree of each point is at least 3.

If G is not a brick, then the assertion follows by induction on the number of bricks of G by the argument in the proof of Theorem 7.6.6.■

BOX 7F. The Dimension of a "Good" Polytope

In Section 7.6 we studied the dimension of the perfect matching polytope, and derived the rather elegant formula found in Theorem 7.6.6. Had we been concerned only with the question of a polynomial time algorithm to compute this number, it would have been enough to know that we can optimize any linear objective function in polynomial time over $PM(G)$. In fact, we have the following general procedure to determine the dimension of a polytope with rational vertices, provided we have a subroutine to maximize (or minimize) an arbitrary linear objective function with rational coefficients over said polytope (Edmonds, Lovász and Pulleyblank (1982)).

Let P be any polytope with rational vertices. During the algorithm, we shall form a set Z of affinely independent vertices of P and a set A of linearly independent linear functions on $\Re^n$, such that each function in A is constant on P. We start with $A = Z = \emptyset$. Suppose we have formed sets A and Z. Clearly,

$$|Z| - 1 \le \dim P \le n - |A|,$$

so if $|A|+|Z| = n+1$, then we are done, for $\dim P$ is determined. Assume that $|A|+|Z| \le n$. Then there exists a linear function $\boldsymbol{a}$ on $\Re^n$, such that $\boldsymbol{a}$ is linearly independent of the functions in A and $\boldsymbol{a}$ is constant on Z. Compute the maximum and minimum of $\boldsymbol{a}\cdot\boldsymbol{x}$ for $\boldsymbol{x} \in P$. If these two are equal, then P is contained in the hyperplane $a \cdot x = \max\{a \cdot y \mid y \in P\}$, so we may add a to A. If the maximum is attained at some point $z_1 \in P$ and the minimum is attained at $z_2 \in P$ and if $a \cdot z_1 \neq a \cdot z_2$, then at least one of z_1 and z_2 is not in the affine hull of Z (since $a \cdot z$ is constant for $z \in Z$ and hence also for z in the affine hull of Z). Thus we may add at least one of z_1 and z_2 to Z. In both cases, we were able to increase $|A| + |Z|$, and thus the algorithm will terminate after at most $n+1$ iterations.

Suppose that P is a polytope with rational vertices and that we have an algorithm to maximize any linear objective function with rational coefficients over P. Suppose, in addition, that we know explicitly a natural number T, such that the numerator and denominator of any entry of any vertex of P are bounded by T. As we have seen above, we can determine the dimension of P in polynomial time. Using this fact, we can also check if a given inequality is a facet of P. In fact, let $a \cdot x \leq b$ be an inequality valid for all $x \in P$. (Here $a \in \mathcal{Q}^n$ and $b \in \mathcal{Q}$). First we show that we can also maximize any linear objective function $c \cdot x$ ($c \in \mathcal{Q}^n$) over the polytope

$$P_1 = P \cap \{x \in \Re^n \mid a \cdot x = b\}.$$

In fact, we claim that it suffices to maximize the linear objective function $(c + Na) \cdot x$ over P, where N is a very large number, because of the following equation which is valid if N is large enough:

$$\max\{c \cdot x \mid x \in P_1\} = \max\{(c + Na) \cdot x \mid x \in P\} \\ - N \max\{a \cdot x \mid x \in P\}.$$

Of course, we must estimate how large N has to be in terms of T, n, $|a|$ and $|c|$ and then show that the numbers occurring are not too large. But, with some effort, this can all be carried out; for details the reader is referred to Edmonds, Lovász and Pulleyblank (1982).

Having an algorithm to maximize $c \cdot x$ over an arbitrary face of P enables us (by the algorithm sketched at the beginning of this box) to decide whether an inequality $a \cdot x \leq b$ determines a facet of P. In fact, it suffices to check whether or not

$$\max\{a \cdot x \mid x \in P\} = b,$$

and if the answer is yes, to compute the dimension of the polytope

$$P_1 = P \cap \{x \mid a \cdot x = b\}.$$

Thus the facts that both the dimension of the perfect matching polytope of a graph can be computed, and the facets of this polytope recognized, in polynomial time, follow quite generally from the fact (to be shown in Chapter 9) that every linear objective function can be maximized over $PM(G)$ in polynomial time. Of course, the combinatorial solution given above gives more insight and faster algorithms. There are several other combinatorically defined polyhedra over which linear objective functions can be optimized in polynomial time (branching polytopes, matroid intersection polytopes, vertex packing polytopes of perfect graphs etc.). For combinatorial characterizations of dimensions of such polyhedra see Giles (1978), Fonlupt and Zemirline (1983) and Frank and Tardos (1984).

This Page Intentionally Left Blank

8

Determinants and Matchings

8.0. Introduction

Let G be a simple bipartite graph with bipartition (U,W) where $U=\{u_1,\ldots,u_n\}$ and $W=\{w_1,\ldots,w_n\}$. Form the **biadjacency matrix** $A=(a_{ij})$, where

$$a_{ij}=\begin{cases}1, & \text{if } u_iw_j\in E(G),\\ 0, & \text{otherwise.}\end{cases}$$

The starting point of our discussion is the observation that every non-zero expansion term of $\det A$ corresponds to a perfect matching in G and vice versa. *So if G has no perfect matching then* $\det A=0$.

Unfortunately, the converse is not true! Expansion terms which correspond to different perfect matchings may cancel each other. For example, if $G=K_{n,n}$ is a complete bipartite graph, then A is a matrix with entries all 1 and hence $\det A=0$.

To overcome this difficulty, several methods have been proposed:

(a) modify the definition of the determinant so that all expansion terms have the same sign;
(b) replace the 1's in A by indeterminates (or equivalently by algebraically independent transcendentals) so that no cancellation occurs;
(c) replace some of the 1's in A by -1 so that every expansion term in $\det A$ becomes the same (i.e., all are $+1$ or all are -1);
(d) replace some of the 1's in A by -1 at random and take the average of the squares of the resulting determinants;
(e) work over the finite field $GF(2)$; then the signs of the expansion terms do not matter.

Whatever information can be gained from (e) is summarized in Exercises 8.0.1 and 8.0.2 below.

We shall discuss (a) in Section 8.1. The other approaches can be extended to non-bipartite graphs and will be discussed in Sections 8.2–8.4. Most of the results presented will concern the number of perfect matchings in various classes of graphs. Let us point out that since the enumeration problem for perfect matchings in general graphs (even in bipartite graphs) is NP-hard, it makes sense not only to seek good upper and lower bounds, but also to seek special classes for which the problem can be solved exactly.

A further relationship between determinants and matchings is the following. Let G' be a forest and A its adjacency matrix. Consider the characteristic polynomial $\det(xI - A)$. It is easy to verify that the coefficient of x^{p-2k} in this polynomial is (up to its sign) the number of k-element matchings in G'. While this nice relation does not remain valid for general graphs, we can introduce the corresponding generating function for the numbers of matchings of different sizes in G. Many deep properties of these polynomials have been discovered by Heilman and Lieb and Godsil and Gutman and we shall survey these in Section 8.5. Many of these results were motivated by problems in quantum chemistry.

Complementing certain previous results, some further lower bounds on the number of perfect matchings in various graphs will be discussed in Section 8.6. Some exercises deal with the number of matchings of various sizes in certain special classes of graphs. These results really fall under the heading of classical (enumerative) combinatorics and we shall only touch upon them.

The last section of this chapter sketches applications of matching theory to a problem in chemistry — topological resonance energy — and to a topic in solid state physics — the Ising model for ferromagnetism. This classical application of the theory of matchings motivated Kasteleyn, a physicist, to develop the Pfaffian method for the enumeration of perfect matchings in planar graphs. But, perhaps quite surprisingly, other areas of matching theory, namely matching algorithms and matching polytopes, have proved very useful in this physical application as well.

We conclude this introduction with a brief discussion of two related enumeration problems. Throughout the chapter, our main concern will be the enumeration of perfect matchings. But what about the enumeration of *all* matchings? Furthermore, in Chapters 6 and 7 we found that many problems on matchings are just as interesting, and sometimes even more easily handled, when generalized to T-joins. So what about enumerating T-joins? The problem of enumerating *all* matchings can be reduced to the problem of enumerating *perfect* matchings. (See Exercise 8.0.3.) The problem of enumerating T-joins is trivial using Exercise 8.0.5 below, but there are related questions which are not at all trivial. In fact, quite a few questions in connection with these two problems have never been studied at all! Some later exercises will also show how certain key results on enumerating perfect matchings carry over to all matchings and to T-joins.

We first present two results on the *parity* of the number of perfect matchings. (See Little (1972).)

8.0.1. EXERCISE. A graph G has an even number of perfect matchings if and only if there is a set $S \subseteq V(G)$, $S \neq \emptyset$, such that every point in $V(G)$ is adjacent to an even number of points in S.

8.0.2. EXERCISE. Every Eulerian graph has an even number of perfect matchings.

8.0.3. EXERCISE. Show that every graph G can be embedded in a graph G' with at most $3|V(G)|$ points so that, for every matching M of G, there exists a unique perfect matching of G' such that $F \cap E(G) = M$. (Thus the number of all matchings in G is equal to the number of perfect matchings in G'.) If G is outerplanar then G' can be chosen to be planar. (A graph is **outerplanar** if can be drawn in the plane in such a way that no lines intersect and all points of the graph lie on the boundary of the infinite face.)

8.0.4. EXERCISE. Let G be a graph with $T \subseteq V(G)$, $|T|$ odd. Show that one can construct a graph G' with at most $4q$ points, such that G can be obtained from G' by contracting some set U of lines, and every T-join of G is the image of a unique matching of G'. (Thus the number of T-joins of G is equal to the number of perfect matchings of G'.) If G is planar then G' can also be chosen to be planar.

8.0.5. EXERCISE. Let G be a connected graph with $T \subseteq V(G)$, $|T|$ odd. Prove that the number of T-joins of G is 2^{q-p+1}.

8.1. Permanents

Let $A = (a_{ij})$ be an $n \times n$ matrix. Define the **permanent** of A by

$$\operatorname{per} A = \sum a_{1\pi(1)} a_{2\pi(2)} \cdots a_{n\pi(n)},$$

where the summation extends over all permutations π of $\{1, \ldots, n\}$. Note that the only difference between this and the definition of the determinant is that here all terms are taken with the same sign.

It follows by our remarks above that if A is the biadjacency matrix of a simple bipartite graph, then each non-zero term in the definition of per A corresponds to a perfect matching of G and vice versa. So per A counts the number of perfect matchings in G. Modifying the definition of A appropriately, this observation extends to non-simple bipartite graphs.

Let G be a (not necessarily simple) bipartite graph with bipartition (U, W), where $U = \{u_1, \ldots, u_n\}$, and $W = \{w_1, \ldots, w_n\}$. Let a_{ij} denote the number of lines connecting u_i to v_j and let $A = (a_{ij})$. As before, $\Phi(G)$ denotes the number of perfect matchings in G. The following result is then immediate.

$$\operatorname{per} A = \Phi(G). \tag{8.1.1}$$

The main problem with this approach is that the permanent, similar as it may seem to the determinant, is incomparably more difficult to handle.

Recall that the value of a determinant is "well-behaved" under certain row (and column) operations, like adding and subtracting rows from other rows and multiplying a row or column by a scalar. Under ordinary matrix multiplication, we have the nice formula: $\det(AB) = \det A \det B$. Also, convenient schemes for evaluating determinants like expanding by a row or column are well-known. The reader may find it instructive to work out which of these apply to permanents! For a comprehensive treatment of permanents the reader is directed to the book by Minc (1978).

But the difficulty lies not only in the necessity of developing a new calculus of permanents. If we are given an $n \times n$ determinant whose entries are small integers, we can use the methods of any textbook to evaluate it in about $O(n^3)$ or $O(n^4)$ steps (depending on how we count arithmetic operations). But if we want to evaluate a permanent, we have a problem which is at least as hard as any NP-complete problem (L. Valiant (1979a, 1979b)). So a nice calculus of permanents analogous to that of determinants probably does not exist!

But there are two inequalities involving permanents which are very useful in the study of matching theory. The first of these was conjectured by van der Waerden in 1926, but only proved by Falikman (1981) and Egoryčev (1980, 1981) half a century later. We quote this result without proof. Recall that a matrix is called **doubly stochastic** if it is non-negative and all its row sums and column sums are 1.

8.1.1. THEOREM. *Let A be a doubly stochastic $n \times n$ matrix. Then*

$$\operatorname{per} A \geq \frac{n!}{n^n},$$

where equality holds if and only if every entry of A is $1/n$. ∎

Another fundamental inequality concerning permanents was conjectured by Minc and proved by Brègman (1973). For a simpler proof also see Schrijver (1978).

8.1.2. THEOREM. *Let A be a* 0–1 matrix with row sums $r_1, \ldots, r_n$. Then

$$\operatorname{per} A \leq (r_1!)^{1/r_1} \cdots (r_n!)^{1/r_n}.$$

■

As applications of these inequalities, we derive the following estimates on the number of perfect matchings in regular bipartite graphs.

8.1.3. THEOREM. *Suppose G is a k-regular bipartite graph having $2n$ points. Then*

$$\Phi(G) \geq n!\left(\frac{k}{n}\right)^n.$$

PROOF. Let A be the biadjacency matrix of G. Then

$$\Phi(G) = \operatorname{per} A = k^n \operatorname{per}\left(\frac{1}{k}A\right) \geq k^n \frac{n!}{n^n}$$

by Theorem 8.1.1. ■

8.1.4. THEOREM. *Let G be a k-regular simple bipartite graph on $2n$ points. Then*

$$\Phi(G) \leq (k!)^{n/k}.$$

PROOF. If A is the biadjacency matrix of G, then

$$\Phi(G) = \operatorname{per} A \leq \left((k!)^{1/k}\right)^n$$

by Theorem 8.1.2. ■

Two remarks are in order. First, if G is not assumed to be simple, then of course no better upper bound can be given on $\Phi(G)$ than the trivial bound k^n. This follows by considering the graph consisting of n components, each having 2 points and k parallel lines. The upper bound in Theorem 8.1.4 is sharp whenever $k \mid n$; this is seen by considering the union of n/k copies of $K_{k,k}$.

Second, if we approximate the factorials in the above bounds using Stirling's formula:

$$n! = n^n e^{-n} \sqrt{2\pi n}\, r(n),$$

(where $r(n) \to 1$ as $n \to \infty$), we obtain the following bounds on $\Phi(G)$ for a simple k-regular graph G on $2n$ points:

$$\sqrt{2\pi n}\left(\frac{k}{e}\right)^n r(n) \le \Phi(G) \le (2\pi k)^{n/k}\left(\frac{k}{e}\right)^n r(k)^{n/k}.$$

Hence it follows that

$$\lim \frac{1}{k}\sqrt[n]{\Phi(G)} = \frac{1}{e}$$

if $k, n \to \infty$. This shows that the number of perfect matchings in a k-regular simple bipartite graph depends remarkably little on the structure of G.

While the upper bound on $\Phi(G)$ given by Theorem 8.1.4 is sharp for infinitely many values of n if k is fixed, the lower bound given by Theorem 8.1.3 is attained only if n is a divisor of k. Let $\Phi(n,k)$ denote the minimum number of perfect matchings a k-regular bipartite graph on $2n$ points can have. Schrijver and W.G. Valiant (1980) conjectured that the following limit holds for fixed k:

8.1.5. CONJECTURE.

$$\lim_{n\to\infty} \sqrt[n]{\Phi(n,k)} = \frac{(k-1)^{k-1}}{k^{k-2}}.$$

This conjecture is trivially valid for $k \le 2$. Schrijver and W.G. Valiant (1980) proved that the relation $\le$ holds for every $k \ge 1$. On the other hand, Voorhoeve (1979) proved that the conjecture is valid for $k = 3$, the first non-trivial case. Their results are precisely stated in the next two theorems.

8.1.6. THEOREM. *Suppose $k, n \ge 1$. Then there exists a k-regular bipartite graph on $2n$ points having at most $k^{2n}/\binom{kn}{n}$ perfect matchings.*

PROOF. Let us start with a graph H consisting of kn independent lines connecting two sets A and B each of cardinality kn. Let P_1 and P_2 be partitions of A and B, respectively, into n subsets of cardinality k. If we identify those points belonging to the same class in these partitions, we get a k-regular bipartite (multi-)graph $H(P_1, P_2)$ on $2n$ points. We shall evaluate the sum

$$\sum_{P_1,P_2} \Phi(H(P_1,P_2)).$$

Let M be any set of n lines of H. The number of pairs of partitions (P_1, P_2), such that $H(P_1, P_2)$ contains M as a perfect matching, is

$$\left(\frac{((k-1)n)!}{((k-1)!)^n}\right)^2.$$

(This is the number of partitions of the $(k-1)n$ remaining points of A and of B, respectively, into n labelled classes of size $k-1$). The set M can be chosen $\binom{kn}{n}$ ways. This yields that

$$\sum_{P_1,P_2} \Phi(H(P_1,P_2)) = \binom{kn}{n}\left(\frac{((k-1)n)!}{((k-1)!)^n}\right)^2.$$

The *number of terms* on the left hand side is the square of the number of partitions of kn points into n *un*labelled classes of size k, which is

$$\left(\frac{(kn)!}{n!(k!)^n}\right)^2.$$

Thus there must be a term on the left hand side which is at most

$$\binom{kn}{n}\left(\frac{((k-1)n)!}{((k-1)!)^n}\right)^2 \bigg/ \left(\frac{(kn)!}{n!(k!)^n}\right)^2 = k^{2n} \bigg/ \binom{kn}{n}. \qquad \blacksquare$$

REMARK. The above proof may also be interpreted as showing that the average number of perfect matchings in a k-regular bipartite graph on $2n$ points is $k^{2n} / \binom{kn}{n}$. If we restrict ourselves to *simple* bipartite graphs, then we can sharpen this average somewhat. O'Neil (1969) proved that the averge number of perfect matchings in a k-regular simple bipartite graph with $2n$ points is $k^{2n} / \binom{kn}{n}\exp(-1/2+o(1))$. The error term here has been improved by several authors. (See McKay (1982).) The expectation and variance of the number of perfect matchings in simple regular non-bipartite graphs have been determined by Bollobás and McKay (1982).

8.1.7. THEOREM. *Let G be a cubic bipartite graph on $2n$ points. Then $\Phi(G) \geq \left(\frac{4}{3}\right)^n$.*

Note that Theorem 8.1.3 gives a lower bound of (roughly) $\left(\frac{3}{e}\right)^n < \left(\frac{4}{3}\right)^n$. The proof which follows is rather simple and it may be expected that a similar argument would yield a proof of Conjecture 8.1.5 for general k. Such an extension of this proof, however, has resisted all our efforts to date.

PROOF (of Theorem 8.1.7). We prove the following slightly more general assertion by induction on n: if G is a bipartite graph on $2n$ points such that each color class of G contains 1 point of degree 2 and $n-1$ points of degree 3, then G contains at least $2(4/3)^{n-1}$ perfect matchings. To see that this assertion does indeed imply the assertion in the theorem, let G be a cubic bipartite graph on $2n$ points and let e_1, e_2, e_3 be three lines of G incident with the same point. Then if the assertion formulated above is true, each of the graphs $G-e_i$ contains at least $2(4/3)^{n-1}$ perfect matchings. Since every perfect matching of G is counted in exactly two graphs $G-e_i$, this implies that G contains at least $3(4/3)^{n-1} > (4/3)^n$ perfect matchings.

Let u be a point of G of degree 2, and let uv_1 and uv_2 be the two lines incident with u.

Case 1. $v_1 = v_2$. Then the graph $G-u-v_1$ is either cubic or has exactly one point of degree 2 in each color class. So by the induction hypothesis, $G-u-v_1$ contains at least $2(4/3)^{n-2}$ perfect matchings. Each of these can be augmented by either uv_1 or uv_2. Thus we have found $2 \cdot 2 \cdot (4/3)^{n-2} > 2 \cdot (4/3)^n$ perfect matchings.

Case 2. $v_1 \neq v_2$. Delete u from G and identify v_1 and v_2 to obtain a new bipartite graph G'. Observe that $\Phi(G) = \Phi(G')$, since the perfect matchings of G and G' correspond to each other in a natural way.

Case 2a. Assume that one of v_1 and v_2 has degree 2 in G. Then G' is cubic, and so by the induction hypothesis G' has at least $3 \cdot (4/3)^{n-2} > 2 \cdot (4/3)^{n-1}$ perfect matchings.

Case 2b. Assume that v_1 and v_2 have degree 3 in G. Then $v_1 = v_2$ has degree 4 in G'. Let v_1w_i $(i = 1, \ldots, 4)$ be the four lines of G' incident with v_1. The bigraph $G' - v_1w_i$ has one point of degree 2 in each color class, and all other points are of degree 3. So by the induction hypothesis, each graph $G' - v_1w_i$ has at least $2 \cdot (4/3)^{n-2}$ perfect matchings. Adding this for $i = 1, 2, 3, 4$, we count every perfect matching of G' three times. Thus G' has at least $4 \cdot 2 \cdot (4/3)^{n-2}/3 = 2 \cdot (4/3)^{n-1}$ perfect matchings.■

The results in this section show that the number of perfect matchings in a k-regular bipartite graph on $2n$ points grows exponentially with n for every $k \geq 3$. It is natural to conjecture that similar lower bounds exist for non-bipartite graphs as well, but a proof seems to be very difficult. We must be careful though, since the regularity of a non-bipartite graph does not imply the existence of even *one* perfect matching. A natural assumption to make at this point, therefore, is that the graph be elementary without forbidden lines, that is, 1-extendable.

8.1.8. CONJECTURE. *For $k \geq 3$ there exist constants $c_1(k) > 1$ and $c_2(k) > 0$ such that every k-regular elementary graph on $2n$ points, without forbidden lines, contains at least $c_2(k) \cdot c_1(k)^n$ perfect matchings. Furthermore $c_1(k) \to \infty$ as $k \to \infty$.*

We conclude the section with another conjecture of Schrijver (1983b). He has proved this conjecture for infinitely many values of k, including all powers of 2 (see the exercises below).

8.1.9. CONJECTURE. *Let G be a k-regular bipartite graph on $2n$ points. Then the number of colorations of the lines of G with k given colors is at least*

$$\left(\frac{(k!)^2}{k^k}\right)^n.$$

8.1.10. EXERCISE. Prove that if Conjecture 8.1.8 is true for $k = a$ and $k = b$ then it is also true for $k = ab$. Consequently, the conjecture is true whenever k is a power of 2.

8.1.11. EXERCISE. Prove, using the result of Exercise 8.1.10, that the number of perfect matchings in a k-regular bipartite graph on $2n$ points is at least $(k/e)^n$.

This lower bound is only slightly worse than the bound in Theorem 8.1.3.

The existence of this lower bound on $\Phi(G)$ is equivalent to the following weaker version of Theorem 8.1.1: If A is a doubly stochastic matrix, then $\operatorname{per} A \geq e^{-n}$. This result, which represented the first significant step toward the solution of van der Waerden's Conjecture, is due to Bang (1979) and Friedland (1979). The graph-theoretic proof contained in these exercises is due to Schrijver and W.G. Valiant (1980).

8.2. The Method of Variables

Let G be a bipartite graph as in the introduction to this chapter and for each line $e \in E(G)$, let x_e be a (formal) variable. Define the matrix $A(\boldsymbol{x}) = (a_{ij})$ by

$$a_{ij} = \begin{cases} x_e, & \text{if } e = u_i w_j, \\ 0, & \text{if } u_i \text{ and } w_j \text{ are non-adjacent.} \end{cases}$$

Then $\det A(\boldsymbol{x})$ is a polynomial in the variables x_e $(e \in E(G))$. Every expansion term of $\det A(\boldsymbol{x})$ corresponds to a perfect matching of G and vice versa. Furthermore, different expansion terms are products of different sets of variables and so they do not cancel each other. Therefore we have:

8.2.1. THEOREM. *A bipartite graph G has a perfect matching if and only if* $\det A(\boldsymbol{x})$ *is not identically* 0. ■

To illustrate the use of Theorem 8.2.1, let us prove the non-trivial part of P. Hall's Theorem 1.1.3. (This proof is due to Edmonds (1967a). A more general result using similar techniques from linear algebra, but also using Hall's theorem itself, was obtained by Perfect (1966).)

Let G be a bipartite graph as before. We want to show that if $|\Gamma(X)| \geq |X|$ for every $X \subseteq W$, then G has a perfect matching. Suppose not. Then by Theorem 8.2.1, $\det A(\boldsymbol{x}) = 0$ for every $\boldsymbol{x}$. Set $x_e = \theta_e$, where θ_e are algebraically independent transcendentals. Now $A = A(\theta)$ is a matrix whose entries are real numbers and its non-zero entries are algebraically independent. Since $\det A = 0$, it follows that the columns of A are linearly dependent.

Let, say, the first k columns $a_1, \ldots, a_k$ of A form a minimal set of linearly dependent columns. Then $a_1, \ldots, a_{k-1}$ are linearly independent and a_k is a linear combination of them:

$$a_k = \sum_{j=1}^{k-1} \lambda_j a_j.$$

We also may assume that the first $k-1$ rows of the matrix $[a_1, \ldots, a_{k-1}]$ are linearly independent (since row rank = column rank). Then from the equations

$$a_{ik} = \sum_{j=1}^{k-1} \lambda_j a_{ij} \qquad (i = 1, \ldots, k-1)$$

we can determine the coefficient λ_j. In fact, Cramer's Rule gives each λ_j as a ratio of two determinants whose entries are among the numbers a_{ij}, $1 \leq i \leq k-1$, $1 \leq j \leq k$. Thus λ_j belongs to the field generated by the numbers a_{ij}, $1 \leq i \leq k-1$, $1 \leq j \leq k$. Also from

$$a_{kk} = \sum_{j=1}^{k-1} \lambda_j a_{kj},$$

it follows that a_{kk} belongs to the field generated by the numbers a_{ij} ($1 \leq i \leq k$, $1 \leq j \leq k$, $(i,j) \neq (k,k)$). But by the algebraic independence of non-zero entries of A, we conclude that $a_{kk} = 0$. Among these k columns, any one is a linear combination of the other $k-1$, so it follows that $a_{k1} = a_{k2} = \cdots = a_{kk} = 0$. On the other hand, we may permute rows k through n and repeat the above argument to show that $a_{\mu\nu} = 0$

for $k \le \mu \le n$, $1 \le \nu \le k$. Thus we have shown that $\Gamma(\{w_1,\dots,w_k\}) \subseteq \{u_1,\dots,u_{k-1}\}$, which is a contradiction, and the proof is complete.

The use of determinants can be extended to non-bipartite graphs, although it involves somewhat more complicated linear algebra. Let G be a simple graph, $V(G) = \{v_1,\dots,v_p\}$ and consider any orientation $\vec{G}$ of G. To each line e assign a variable x_e and define the matrix $B(\boldsymbol{x})$ as follows:

$$B(\boldsymbol{x}) = (b_{ij})_{p\times p}$$

$$\text{where}\quad b_{ij} = \begin{cases} x_e, & \text{if } e = (v_i, v_j), \\ -x_e, & \text{if } e = (v_j, v_i), \\ 0, & \text{if } v_i \text{ and } v_j \text{ are non-adjacent.} \end{cases}$$

If G is bipartite (with bipartition sets of the same size), there is a simple correspondence between $A(\boldsymbol{x})$ and $B(\boldsymbol{x})$. To see this, let the bipartition of G be (U, W) where $U = \{u_1,\dots,u_n\}$ and $W = \{w_1,\dots,w_n\}$. Then the matrix $B(\boldsymbol{x})$ has the following structure:

$$B(\boldsymbol{x}) = \begin{pmatrix} 0 & A(\boldsymbol{x}) \\ -A(\boldsymbol{x}) & 0 \end{pmatrix}$$

and hence it follows immediately that $\det B(\boldsymbol{x}) = (\det A(\boldsymbol{x}))^2$.

If G is not bipartite, we lose this nice relationship between $A(\boldsymbol{x})$ and $B(\boldsymbol{x})$. (In fact, of course, we lose $A(\boldsymbol{x})$!) But in the following classical result due to Tutte (1947), we see that $B(\boldsymbol{x})$ still provides us with a useful criterion for the existence of a perfect matching in the non-bipartite case.

But first we shall need a definition and an identity from linear algebra. Let B be a skew symmetric matrix of dimension $p = 2n$. For each partition $P = \{\{i_1, j_1\},\dots,\{i_n, j_n\}\}$ of the set $\{1,\dots,2n\}$ into pairs, form the expression

$$b_P = \operatorname{sgn}\begin{pmatrix} 1 & 2 & \cdots & 2n-1 & 2n \\ i_1 & j_1 & \cdots & i_n & j_n \end{pmatrix} b_{i_1 j_1} \cdots b_{i_n j_n},$$

where

$$\begin{pmatrix} 1 & 2 & \cdots & 2n-1 & 2n \\ i_1 & j_1 & \cdots & i_n & j_n \end{pmatrix}$$

is a permutation of the elements $1,\dots,2n$ and sgn denotes the sign of this permutation. Note that b_P depends neither on the order in which the classes of the partition are listed nor on the order of the two elements of a class. So b_P indeed depends only on the choice of the partition P.

Now we define the **Pfaffian** of matrix B by:

$$\operatorname{pf} B = \sum_P b_P.$$

For example, if

$$B = \begin{pmatrix} 0 & b_{12} & b_{13} & b_{14} \\ -b_{12} & 0 & b_{23} & b_{24} \\ -b_{13} & -b_{23} & 0 & b_{34} \\ -b_{14} & -b_{24} & -b_{34} & 0 \end{pmatrix}$$

then $\operatorname{pf} B = b_{12}b_{34} - b_{13}b_{24} + b_{14}b_{23}$.

We need the following identity from linear algebra. (See Muir (1882, 1906).)

8.2.2. LEMMA. *If B is a skew symmetric matrix,* $\det B = (\operatorname{pf} B)^2$. ■

8.2.3. THEOREM. *Let G be any graph and let $B(\boldsymbol{x})$ be defined as above. Then graph G has a perfect matching if and only if $\det B(\boldsymbol{x})$ is not identically 0.*

PROOF. The difficulty is that the expansion terms in $\det B(\boldsymbol{x})$ may cancel, since the same variable occurs twice as an entry of $B(\boldsymbol{x})$. If $p = |V(G)|$ is odd then G has no perfect matching, of course, and moreover, $\det B = 0$ since the determinant of a skew symmetric matrix of odd dimension is 0. Thus the theorem is true for graphs with an odd number of points. So assume now that $p = |V(G)| = 2n$.

Now each non-zero term b_P in the definition of $\operatorname{pf} B(\boldsymbol{x})$ corresponds to a perfect matching of G and vice versa. Furthermore, different perfect matchings of G correspond to terms consisting of different variables. Hence $\operatorname{pf} B(\boldsymbol{x}) \equiv 0$ if and only if G has no perfect matching. Using Lemma 8.2.2, Theorem 8.2.3 now follows. ■

So we have a determinant criterion for the existence of a perfect matching. In fact, Tutte's original proof of his perfect matching theorem (Theorem 3.1.1) made use of this criterion. In the next section we will pursue a determinant criterion for counting *all* perfect matchings.

Finally, let us point out that the existence criterion given in Theorem 8.2.3 can be used to obtain a *probabilistic algorithm* for finding a perfect matching in a graph. (See Box 8A.)

8.3. The Pfaffian and the Number of Perfect Matchings

In the previous section, we saw that the matrix of indeterminates $B(\boldsymbol{x})$ provides a method of deciding if an arbitrary graph has a perfect matching. In this section we will show that, with suitable modification, this matrix actually helps us count the number of different perfect matchings.

As before, adopt an arbitrary orientation $\vec{G}$ of G. Suppose $n = |V(G)|$. Let us define yet another matrix related to $\vec{G}$, the skew **adjacency matrix** of $\vec{G}$, $A_s(\vec{G})$ as follows:

$$A_s(\vec{G}) = (a_{ij})_{n\times n},$$

$$\text{where}\quad a_{ij} = \begin{cases} 1, & \text{if } (u_i, u_j) \in E(\vec{G}), \\ -1, & \text{if } (u_j, u_i) \in E(\vec{G}), \\ 0, & \text{otherwise.} \end{cases}$$

Now recall that each term in $\mathrm{pf}(A_s(\vec{G}))$ is of the form $\pm a_{i_1j_1}\cdots a_{i_nj_n}$ and hence equals 1, -1 or 0. Moreover, there is an obvious 1:1 correspondence between the perfect matchings of G and the non-zero terms of $\mathrm{pf}(A_s(\vec{G}))$. (It may be helpful to point out once again that the order in which the endpoints of the line are listed is irrelevant in the Pfaffian.)

If F is a perfect matching, let us define the **sign** of F, $\mathrm{sgn}\, F$, to be equal to the sign of the corresponding term in the Pfaffian. It is immediate by definition — but important to realize — that we always have

$$|\mathrm{pf}(A_s(\vec{G}))| \le \Phi(G). \tag{8.3.1}$$

It is then natural to ask when, in fact, equality holds, or equivalently, when does the Pfaffian enumerate all the perfect matchings of G. If equality holds, we will call the digraph $\vec{G}$ **Pfaffian** and, if there exists such a **Pfaffian orientation** for a graph G, we shall simply call G **Pfaffian**.

Let C be any even undirected cycle in $\vec{G}$. Now regardless of which of the two possible routings around C is chosen, if C contains an even number of oriented lines whose orientation agrees with the routing (we will call these "forward" lines), then C also contains an even number of lines whose orientation is opposite to the routing (we will call these "backward" lines). Hence the following definition is independent of the routing chosen.

If C is an even undirected cycle in $\vec{G}$, we shall say C is **evenly oriented** if it has an even number of lines oriented in the direction of the routing. Otherwise C is **oddly oriented**.

8.3.1. LEMMA. *Let $\vec{G}$ be an arbitrary orientation of an undirected graph G. Let F_1 and F_2 be any two perfect matchings of G and let k denote the number of evenly oriented alternating cycles formed in $F_1 \cup F_2$. Then* $\mathrm{sgn}\, F_1 \cdot \mathrm{sgn}\, F_2 = (-1)^k$.

PROOF. We first observe that if this equation holds for one orientation it holds for all. To see this suppose one line x is reversed in a given orientation. If x lies in neither F_1 nor in F_2, then there is no effect on either side of the equation. If x lies in both F_1 and F_2, then the sign of each changes, the value of k does not change and again the equation holds. Finally, if $x \in E(F_1) - E(F_2)$, it lies in some alternating cycle and hence $\operatorname{sgn} F_1$ changes, $\operatorname{sgn} F_2$ does not, but k changes by one. Thus the equation does hold in all instances.

Since we are now free to orient G as we like, we use the following scheme: if line x is in $F_1 \cap F_2$, orient it arbitrarily and similarly for $x \notin F_1 \cup F_2$. For $x \in E(F_1) \oplus E(F_2)$ (i.e., for x in an alternating cycle), orient x so that the alternating cycle containing it becomes a directed cycle. (Either of the two possible orientations of this cycle are equally satisfactory for us).

Now we must make the following remark about relabelling the points of G. A relabelling of G corresponds to permuting the rows (and columns) of the matrix $A_s(\vec{G})$ by some permutation π and one readily observes that this has the effect on the Pfaffian of multiplying each term by the sign of π. But then each perfect matching has its sign multiplied by $\operatorname{sgn} \pi$; that is, all $\operatorname{sgn} F$'s change or all remain unchanged. So certainly the equation contained in the statement of this lemma remains invariant under any relabelling of $V(G)$.

So we adopt the following method of relabelling: first take those oriented lines of $F_1 \cap F_2$ and label their endpoints with consecutive positive integers such that the label of the "head" of the oriented line equals the label of the "tail" plus one. For those oriented lines on an alternating cycle, choose any line in the cycle belonging to F_1 and number its tail and head with the next two (unused) consecutive positive integers. Continue until all remaining alternating cycles (and hence all of $V(G)$) are labelled.

Now b_{F_1}, the term of the Pfaffian corresponding to F_1, contains the sign of the identity permutation (+1) and each a_{mn} in the rest of the product is +1. So $\operatorname{sgn} F_1 = +1$. In the term b_{F_2}, all the a_{mn} are again +1 (by means of our special orientation and labelling) so $\operatorname{sgn} F_2$ is just the sign of the corresponding permutation ϵ. On the other hand, since each alternating (*even*) cycle of G corresponds to an *even* cycle of the permutation ϵ and the cycles are disjoint, it follows that $\operatorname{sgn} F_2 = (-1)^k$ and hence the lemma is proved. ∎

We now address the question: "When is a *given* orientation $\vec{G}$ of G Pfaffian?". The preceding lemma implies the next result which does give an answer of a sort, but, unfortunately, not a *good* characterization. Recall that a cycle C of G is nice if $G-V(C)$ contains a perfect matching.

8.3.2. THEOREM. *Let G be any even graph and $\vec{G}$, an orientation of G. Then the following four properties are equivalent:*

(1) *$\vec{G}$ is a Pfaffian orientation of G.*
(2) *Every perfect matching of G has the same sign relative to $\vec{G}$.*
(3) *Every nice cycle in G is oddly oriented relative to $\vec{G}$.*
(4) *If G has a perfect matching, then for some perfect matching F, every F-alternating cycle is oddly oriented relative to $\vec{G}$.* ■

We will now proceed to prove Kasteleyn's classical result (1963, 1967) that every planar graph has a Pfaffian orientation. Note that if a connected graph G is embedded in the plane, then it makes sense to speak of routing all the cycles of G in the "clockwise sense".

Let F be any face. We say that a line e on the boundary of F is oriented in the clockwise sense if when proceeding in the direction of the arrow on e, F lies to the right. Note that if F is a finite face the boundary of which is a cycle, then clockwise oriented lines are just those lines which are forward with respect to the clockwise routing.

8.3.3. LEMMA. *If $\vec{G}$ is a connected plane digraph such that every boundary face — except possibly the infinite face — has an odd number of lines oriented clockwise, then in every cycle the number of lines oriented clockwise is of opposite parity to the number of points of $\vec{G}$ inside the cycle. Consequently, $\vec{G}$ is Pfaffian.*

PROOF. Let C be any cycle in $\vec{G}$. Suppose there are f faces inside C and let c_i denote the number of clockwise lines on the boundary of face i for $i=1,\ldots f$. Each c_i is an odd number by hypothesis and hence

$$f \equiv \sum_{i=1}^{f} c_i \pmod 2.$$

Let v be the number of points inside C, e, the number of lines inside C, and k, the number of lines on C. Then by Euler's formula,

$$(v+k)-(e+k)+(f+1)=2,$$

and hence

$$e=v+f-1.$$

Now if c is the number of clockwise lines on cycle C, we have $\sum_{i=1}^{f} c_i = c+e$, for each interior line gets counted as a clockwise line exactly once when visiting the faces interior to C. So we have

$$f \equiv \sum_{i=1}^{f} c_i = c + e = c + (v + f - 1) \pmod 2$$

and hence $c + v - 1 \equiv 0 \pmod 2$; that is, $c \equiv (v-1) \pmod 2$. In particular, if C is a cycle which alternates with respect to a perfect matching, then it must contain an even number of points in its interior since said points must be matched pairwise. Thus C contains an odd number of forward lines and so condition (3) in Theorem 8.3.2 is satisfied. Thus the given orientation is Pfaffian. ∎

Next we show how we can actually construct an orientation satisfying the condition in Lemma 8.3.3, for any connected planar graph G. We use induction on the number of lines in G. If G is a tree then any orientation will do. Suppose that G is not a tree and choose any line on the boundary of the infinite face which belongs to a cycle. Let F_0 be the finite face containing this line e.

The graph $G - e$ has an orientation so that each finite face boundary has an odd number of clockwise oriented lines, by the induction hypothesis. Reinsert e and orient it so that the boundary of F_0 has an odd number of clockwise oriented lines. Since all the finite face boundaries different from F_0 remain unchanged, this orientation of G will have the desired properties.

In summary then, we can state the following theorem.

8.3.4. THEOREM. (*Kasteleyn's Theorem*). *Every planar graph has a Pfaffian orientation. Such an orientation can be constructed in polynomial time.* ∎

The procedure is illustrated on the graph in Figure 8.3.1. The skew symmetric matrix belonging to the orientation obtained is

$$A_s(\vec{G}) = \begin{pmatrix} 0 & 1 & 0 & -1 & -1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 \end{pmatrix}$$

Then $\det A_s(\vec{G}) = 9$ and hence by Lemma 8.2.2, $\mathrm{pf}(A_s(\vec{G})) = 3$. The reader may easily check that G does in fact have three perfect matchings.

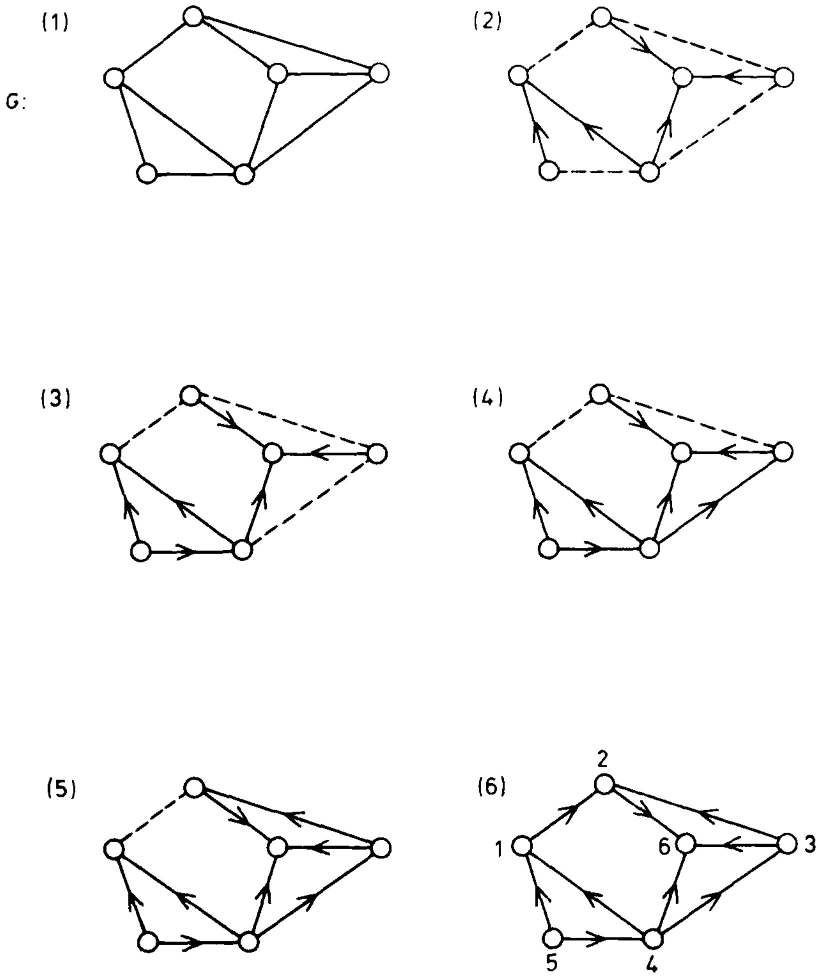


FIGURE 8.3.1. Construction of a Pfaffian orientation

Actually, Kasteleyn's method can be used to obtain the number of perfect matchings for a larger class of graphs than just those which are planar. Recall the famous theorem of Kuratowski (1930), which says that a graph is planar if and only if it contains no subdivision of K_5 or $K_{3,3}$. This implies that the following result of Little (1974a) is a generalization of Kasteleyn's Theorem 8.3.4. We omit the proof.

8.3.5. THEOREM. *If G contains no subdivision of $K_{3,3}$, then G has a Pfaffian orientation.* ■

8.3.6. EXERCISE. Show that $K_{3,3}$ has no Pfaffian orientation.

Now recall that Theorem 8.3.2 gave us several equivalent conditions for an orientation to be Pfaffian. By Kasteleyn's result we know that planar graphs always have such an orientation. But what about general graphs? We proceed to give some necessary and sufficient conditions for a graph to have a Pfaffian orientation, following Little (1973).

Suppose G is any graph, $\vec{G}$, any orientation of G and F, any perfect matching of G. Then define the function $l(F)$ with values in $\{0,1\}$ by $\operatorname{sgn} F = (-1)^{l(F)}$. Now to each line e in $E(G)$ assign a variable x_e.

8.3.7. THEOREM. *If G, $\vec{G}$ and the x_e's are as above, then the following are equivalent:*

(1) *G has a Pfaffian orientation.*
(2) *There is a set $S \subseteq E(G)$ such that for all perfect matchings F of G, $|S \cap F| \equiv l(F) \pmod 2$.*
(3) *The system of equations (one for each perfect matching F of G): $\sum_{e\in F} x_e = l(F)$ has a solution over $GF(2)$.*
(4) *There is no set of perfect matchings $F_1, \ldots, F_r$ such that $\sum_{j=1}^{r} F_j \equiv 0 \pmod 2$ and $\sum_{j=1}^{r} l(F_j) \equiv 1 \pmod 2$.*

PROOF. Let $\vec{\vec{G}}$ be any other orientation of G, and let S be the set of lines which must be reversed to obtain $\vec{\vec{G}}$ from $\vec{G}$. Now if F is any perfect matching, then

$$\operatorname{sgn} F = \operatorname{sgn} F(-1)^{|S\cap F|} = (-1)^{l(F)+|S\cap F|}.$$

Hence G is Pfaffian (that is, $\operatorname{sgn} F = 1$ holds for every perfect matching F) if and only if $|S \cap F| \equiv l(F) \pmod 2$. So the existence of a Pfaffian re-orientation is equivalent to the existence of a set S with property (2). Thus (1) and (2) are indeed equivalent.

Next note that (3) is just a rephrasing of (2) in terms of the incidence vector x of the set S. Finally, (4) is obtained form (3) by applying the criterion for the solvability of a system of linear equations. ■

In the special case when G is bipartite there is a characterization similar in spirit to the above, but simpler in that it does not involve the sign of a perfect matching and hence we can dispense with orienting the graph (see Pla (1965) and Little and Pla (1972)).

8.3.8. THEOREM. *Let G be a bipartite graph and let F be any perfect matching in G. Then G has a Pfaffian orientation if and only if there exists no collection of an odd number of F-alternating cycles, the* mod 2 *sum of which is zero.*

PROOF. Let $G = (U, W)$ be the bipartition of G. Orient G by directing all lines of F toward U and all other lines toward W. Note that every F-alternating cycle in G becomes a directed cycle in the resulting $\vec{G}$ and vice versa.

Now first suppose that G does not have a Pfaffian orientation. Then by Theorem 8.3.7, there is a set of perfect matchings $F_1, \ldots, F_r$ such that $\sum_{j=1}^{r} F_j \equiv 0 \pmod 2$, but $\sum_{j=1}^{r} l(F_j) \equiv 1 \pmod 2$. Let $\mathcal{A}$ be the family of all alternating cycles formed by $F \oplus F_j$ for $j = 1, \ldots, r$. Also let k_j denote the number of alternating cycles formed by $F \oplus F_j$. We may assume that the points of G are labelled so that $\operatorname{sgn} F = 1$. Then by Lemma 8.3.1,

$$\operatorname{sgn}(F_j) = (-1)^{k_j} \operatorname{sgn} F = (-1)^{k_j},$$

and hence $k_j \equiv l(F_j) \pmod 2$. Hence

$$|\mathcal{A}| = \sum_{j=1}^{r} k_j \equiv \sum_{r=1}^{r} l(F_j) \equiv 1 \pmod 2.$$

Further, consider the sum of cycles in $\mathcal{A}$ modulo 2. If we consider a line e not in F, then the number of cycles in $\mathcal{A}$ containing e is the same as the number of perfect matchings among $F_1, \ldots, F_r$ containing e, and this number is even by the choice of $F_1, \ldots, F_r$. Thus the sum of cycles in $\mathcal{A}$ is a subset of F. But since the mod 2 sum of cycles must be an Eulerian subgraph, it follows that the mod 2 sum of cycles in $\mathcal{A}$ is 0. This proves the sufficiency of the condition given in the theorem.

Conversely, assume that G contains a collection $\mathcal{A}$ of an odd number of F-alternating cycles, the mod 2 sum of which is 0. Suppose $\mathcal{A} = \{A_1, \ldots, A_t\}$ and let F_i be obtained from F by alternating on A_i. Then $l(F) = 0$ and $l(F_i) = 1$, by Lemma 8.3.1, and $\sum_{i=1}^{t} F_i \equiv \sum_{i=1}^{t} (A_i \oplus F) \equiv tF \equiv F \pmod 2$. So the set $F, F_1, \ldots, F_t$ of perfect matchings violates condition (4) in Theorem 8.3.7 and thus G does not have a Pfaffian orientation. ∎

In spite of the many equivalent conditions for the existence of a Pfaffian orientation of a graph, this property is not well-characterized.

The problem is that no polynomial algorithm is known for checking whether or not a given orientation of a graph G is Pfaffian. We do not even know if this property is in NP. (It is trivially in co-NP; to prove that a given orientation is non-Pfaffian, it suffices to exhibit two perfect matchings with different signs.) Similarly, we do not know whether the property of an undirected graph that it has a Pfaffian orientation is in NP. (This property too is in co-NP; to prove that G does not have a Pfaffian orientation, it suffices to consider an arbitrary orientation and then exhibit a set of perfect matchings violating condition (4) in Theorem 8.3.7.)

One approach to deciding whether or not an undirected graph G has a Pfaffian orientation is to use condition (3) in Theorem 8.3.7. A natural approach is then to compare the $GF(2)$ rank of the matrix of the system with the $GF(2)$ rank of the augmented matrix. Unfortunately, it is not known how to compute the $GF(2)$ rank of this matrix, that is, to determine the maximum number of perfect matchings linearly independent over $GF(2)$. (Recall that the corresponding problem over the real field was solved in Section 7.6.)

One special type of graph for which enumeration of perfect matchings is of considerable "real-world" interest is the class of rectangular lattice graphs or chessboards (perhaps more accurately "go boards"). Consider an $m \times n$ chessboard with mn even. In how many ways can this board be covered by dominoes (dimers)? This problem, known as the **dimer problem**, has applications in statistical mechanics, and in fact these applications prompted Kasteleyn, a physicist, to solve it. (For more on this, see Section 8.7.)

In graph theoretic terms, the dimer problem can be formulated as follows. Let $L(m,n)$ denote the graph whose points are the squares of an $m \times n$ chessboard, two being adjacent if and only if they have a boundary line in common. The problem is to determine $\Phi(L(m,n))$.

There are many formulas known for $\Phi(L(m,n))$. Here we confine ourselves to deducing one from Kasteleyn's Theorem 8.3.4 and then using it to obtain an asymptotic result. For a more detailed account on this subject, see Montroll (1964).

In order to apply Theorem 8.3.4, we first construct an orientation satisfying the conditions of Lemma 8.3.3. It is evident that the orientation shown in Figure 8.3.2 has this property. The skew symmetric adjacency matrix of this orientation is the following:

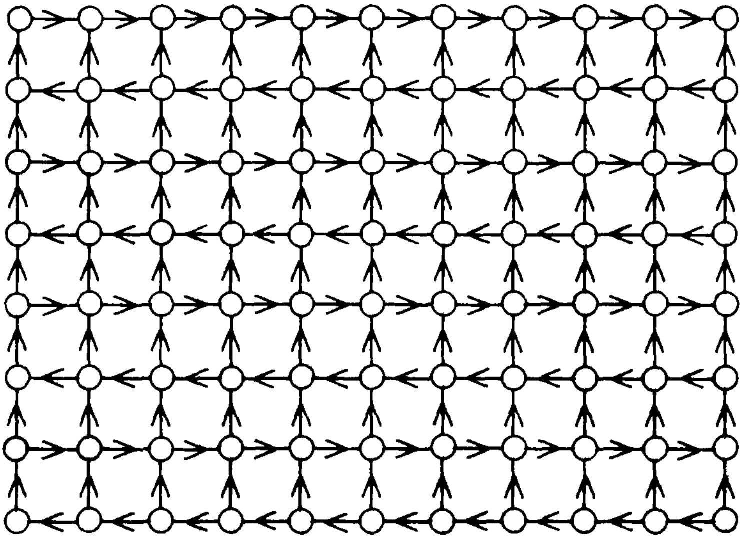


FIGURE 8.3.2. A Pfaffian orientation of the lattice graph

$$A_s(\vec{G}) = \begin{pmatrix} A & I & & & \\ -I & -A & I & & \\ & -I & & & \\ & & \ddots & & \\ & & & & I \\ & & & -I & (-1)^{n-1}A \end{pmatrix} \tag{8.3.2}$$

where A denotes the $m \times m$ matrix

$$A = \begin{pmatrix} 0 & 1 & & & \\ -1 & 0 & 1 & & \\ & -1 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & -1 & 0 \end{pmatrix}.$$

Now multiplying the first column, then the third and fourth row, then the fourth and fifth column, then the seventh and eighth row, etc. of the partitioned matrix $A_s(\vec{G})$ by -1, we do not change the absolute value of the determinant and we obtain matrix M where

$$M = \begin{pmatrix} -A & I & & & & \\ I & -A & I & & & \\ & I & & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & \cdot & I \\ & & & & I & -A \end{pmatrix}.$$

If we denote by B the $n \times n$ matrix

$$B = \begin{pmatrix} 0 & 1 & & & & \\ 1 & 0 & 1 & & & \\ & 1 & & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & \cdot & \\ & & & & 0 & 1 \\ & & & & 1 & 0 \end{pmatrix}$$

then we can write

$$M = -I \otimes A + B \otimes I,$$

where $\otimes$ denotes the Kronecker product of matrices. (For a general reference for Kronecker products see Halmos (1958).)

Let A have eigenvalues $\lambda_1 \cdots \lambda_m$ with eigenvectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$, and let B have eigenvalues $\mu_1, \ldots, \mu_n$ with eigenvectors $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n$. Then $\boldsymbol{b}_i \otimes \boldsymbol{a}_j$ is an eigenvector of M with eigenvalue $\mu_i - \lambda_j$, since

$$\begin{aligned}(-I \otimes A + B \otimes I)(\boldsymbol{b}_i \otimes \boldsymbol{a}_j) &= (-I\boldsymbol{b}_i \otimes A\boldsymbol{a}_j) + (B\boldsymbol{b}_i \otimes I\boldsymbol{a}_j) \\ &= -\lambda \boldsymbol{a}_j(\boldsymbol{b}_i \otimes \boldsymbol{a}_j) + \mu_i(\boldsymbol{b}_i \otimes \boldsymbol{a}_j) \\ &= (\mu_i - \lambda_j)(\boldsymbol{b}_i \otimes \boldsymbol{a}_j).\end{aligned}$$

It is not difficult to determine the eigenvalues of A and B. (See Lovász (1979c).) The eigenvalues of A are $2i\cos((\pi k)/(m+1))$, $(k = 1, \ldots, m)$, and the eigenvalues of B are $2\cos((\pi k)/(n+1))$, $(k = 1, \ldots, n)$. Thus the eigenvalues of matrix M are

$$2\left(\cos\left(\frac{\pi k}{m+1}\right) - i\cos\left(\frac{\pi l}{n+1}\right)\right), \qquad (k = 1, \ldots, m;\ l = 1, \ldots, n).$$

Hence the determinant of matrix M is the product of these numbers. Since we are interested in the absolute value of this determinant, we may replace these mn factors by their absolute values, and so the absolute value of the determinant of matrix (8.3.2) is

$$2^{mn}\prod_{k=1}^{m}\prod_{l=1}^{n}\left(\cos^2\left(\frac{\pi k}{m+1}\right)+\cos^2\left(\frac{\pi l}{n+1}\right)\right)^{1/2}.$$

Hence

$$\Phi(L(m,n)) = 2^{mn/2}\prod_{k=1}^{m}\prod_{l=1}^{n}\left(\cos^2\left(\frac{\pi k}{m+1}\right)+\cos^2\left(\frac{\pi l}{n+1}\right)\right)^{1/4}.$$

This formula, among others, can be used to determine the asymptotic behavior of the logarithm of $\Phi(L(m,n))$ as m and $n \to \infty$:

$$\begin{aligned}\frac{\log\Phi(L(m,n))}{mn} &= \frac{\log 2}{2} + \\ &\frac{1}{4}\cdot\frac{1}{mn}\sum_{k=1}^{m}\sum_{l=1}^{n}\log\left(\cos^2\left(\frac{\pi k}{m+1}\right)+\cos^2\left(\frac{\pi l}{n+1}\right)\right)\\ &\to \frac{\log 2}{2}+\frac{1}{4\pi^2}\int_0^{\pi}\int_0^{\pi}\log(\cos^2 x+\cos^2 y)dx\,dy\\ &= 1/\pi(1-1/9+1/25-1/49+1/81-1/121+\cdots)\\ &= .29\end{aligned}$$

So the number $\Phi(L(m,n))$ grows as $\exp(.29mn)$ as m and $n \to \infty$.

8.3.9. EXERCISE. Let G be a graph which has a Pfaffian orientation $\vec{G}$. Let $c : E(G) \to \Re$ be any weighting of its lines and let $h_{ij} = c(ij)a_{ij}$ where (a_{ij}) is the skew symmetric matrix belonging to this orientation. Then

$$\sum_F\Big(\prod_{e\in F} c_e\Big) = \mathrm{pf}(h_{ij}),$$

where F ranges over all perfect matchings of G.

8.3.10. EXERCISE. Let G be a planar graph, $T \subseteq V(G)$, $|T|$ even and let $c : E(G) \to \Re$ be any weighting of its lines. Find a skew symmetric matrix B of size $4q-2p$ such that

$$\sum_J\Big(\prod_{e\in J} c_e\Big) = \mathrm{pf}\,B,$$

where J ranges over all T-joins of the graph. (Hint: Use the construction of Exercise 8.0.4 to reduce this problem to the preceding exercise.)

8.4. Probabilistic Enumeration of Perfect Matchings

The following theorem establishes a somewhat surprising connection betwen the number of perfect matchings in a graph and the determinant of its orientations. A similar result, with *weighted* rather than *oriented* lines was proved by Godsil and Gutman (1981).

8.4.1. THEOREM. *Let G be a graph. Let $\vec{G}$ be a random orientation of G, obtained by orienting each line in G independently of the others with probability $1/2$ in either direction. Let $A_s(\vec{G})$ be the skew symmetric adjacency matrix of G. Then the expected value of* $\det A_s(\vec{C})$ *is* $\Phi(G)$.

PROOF. Denote $A_s(\vec{G})$ by the $p \times p$ matrix (a_{ij}) and write out the definition of its determinant:

$$\det A_s(\vec{G}) = \sum \operatorname{sgn}(\pi) a_{1\pi(1)} \cdots a_{p\pi(p)},$$

where the summation extends over all permutations π of $\{1, \ldots, p\}$. Note that the a_{ij} are random variables, which assume the values 1 and -1 with probability $1/2$, and which are mutually independent, except that $a_{ij} = -a_{ji}$. Computing the expectation, we have:

$$E\big(\det A_s(\vec{G})\big) = \sum \operatorname{sgn}(\pi) E(a_{1\pi(1)} \cdots a_{p\pi(p)}).$$

If π is a permutation having a fixed point, then the corresponding term in the sum is 0 since $a_{ii} = 0$. If π is a permutation such that i and $\pi(i)$ are non-adjacent for some i, then again the corresponding term is 0. If π is a permutation such that π^2 is not the identity, then there is an i, $1 \le i \le p$, such that $\pi(\pi(i)) \ne i$. Hence, the random variable $a_{i\pi(i)}$ occurs in the product corresponding to this permutation, but the random variable $a_{\pi(i)i}$ does not. So the factor $a_{i\pi(i)}$ of the product corresponding to π is independent of the other factors and hence $E(a_{1\pi(1)} \cdots a_{p\pi(p)}) = E(a_{i\pi(i)})E(a_{1\pi(1)} \cdots a_{(i-1)\pi(i-1)} a_{(i+1)\pi(i+1)} \cdots a_{p\pi(p)}) = 0$. So we are left with the terms corresponding to those permutations which have no fixed points, but for which π^2 is the identity and for which $i\pi(i) \in E(G)$ for all i. Clearly these permutations are in 1:1 correspondence with the perfect matchings of G and an easy computation shows that the term corresponding to such a permutation is always 1. The desired equality follows. ∎

Let us remark that we have seen (by Lemma 8.2.2 and inequality (8.3.1)) that

$$\det A_s(\vec{G}) \le (\Phi(G))^2.$$

So from Theorem 8.4.1 it follows that every graph G has an orientation $\vec{G}$ such that

$$\Phi(G) \le \det A_s(\vec{G}) \le (\Phi(G))^2.$$

However, it is not known how to *construct* such an orientation!

8.4.2. EXERCISE. Let G be a graph and $\vec{G}$ a random orientation of G. For any two perfect matchings F_1, F_2 of G, let $a(F_1, F_2)$ denote the number of alternating $F_1 - F_2$ cycles in G. Prove that if variance is denoted by D, then

$$D^2(\det A_s(\vec{G})) = \sum_{F_1 \neq F_2} 3^{a(F_1,F_2)},$$

where the sum is over all pairs of perfect matchings (F_1, F_2) such that $F_1 \neq F_2$.

8.4.3. EXERCISE. Prove that $D(\det A_s(\vec{G}))/\Phi(G)$ is bounded by a polynomial in $p = |V(G)|$ if G is a complete graph, but that this is not so for every graph.

BOX 8A. Probabilistic Methods in Graph Theory

The proof of Theorem 8.1.6 has the peculiar feature that even though it proves the existence of a k-regular bipartite graph with few perfect matchings, it does not explicitly construct such a graph. Another — and probably nicer — way to formulate this argument is as follows. Let us choose the partitions P_1 and P_2 at random (with uniform distribution from among all partitions of kn elements into n classes of size k). Then the graph $H(P_1, P_2)$ may be viewed as a random k-regular bipartite graph on $2n$ points (i.e., a random variable whose values are k-regular bipartite graphs on $2n$ points). Furthermore, what the computation in the proof yields is, in fact, the expected number of perfect matchings in this random graph:

$$E(\Phi(H(P_1, P_2))) = k^{2n} \Big/ \binom{kn}{n}.$$

Similar non-constructive proofs are now fairly common and the methods involved are often very powerful tools in many branches of combinatorics. This method, developed mainly by P. Erdős, has come to be called the *probabilistic method*.

We do not encounter this method as often in matching theory as, say, in Ramsey theory or in the theory of the chromatic number. The reason is that in matching theory we usually have good characterizations which tell us all possible reasons why a certain property can hold (or not hold) for a graph, whereas the probabilistic method is needed to prove the existence of a graph with a certain property, precisely when no simple reason can be given as to why this property holds for a graph, and so it would be difficult or impossible to prove that a proposed counterexample is indeed a counterexample. Since the enumeration problem for perfect matchings is NP-hard, the probabilistic approach is a quite natural thing to try.

Even a sketch of the details and successes of the probabilistic method would be beyond the scope of this book. Fortunately, we can refer the reader to the monograph of Erdős and Spencer (1974), and also to the more recent monograph of Graham, Rothschild and Spencer (1980) on Ramsey Theory, where the probabilistic method is extensively used and, accordingly, surveyed in detail.

However, probabilistic arguments occur not only in proofs, but in algorithms as well. Consider Theorem 8.2.3. How does this necessary and sufficient condition help us determine whether or not a given graph has a perfect matching? We all learned in high school that it is easy to determine whether or not a polynomial is identically zero; all that is required is to expand all quantities in brackets and carry out all possible cancellations. If we are left with the 0 polynomial, then of course our polynomial is identically 0 and conversely (but slightly less trivially), if we are left with a polynomial which is not *formally* 0, then it is not *identically* 0. But this procedure would be impractical to apply in the case of the polynomial $\det B(x)$, since if $B(x)$ is $n \times n$, to write it out in bracket-free form would take up to $n!$ terms.

Fortunately, however, we can do the following. Let us substitute some randomly generated numbers for the variables x_e, and evaluate the determinant with these entries. If we obtain a non-zero number, then obviously $\det B(x)$ cannot be identically 0. Suppose we find that the determinant is 0. Then it does not follow that it is *identically* 0, since we just might have been unlucky enough to have chosen a root. However, the probability of this is very small. So with considerable confidence we may conclude that $\det B(x)$ is identically 0!

But just how small is the probability of an error, that is, what is the probability of hitting a root? This depends on the way we generate the random numbers. If each x_e is, say, uniformly distributed in $[0,1]$, then the probability of hitting a root is 0. However, we have to do calculations with these x_e's and so a more realistic model is to choose them uniformly from the finite set $\{1,\dots,N\}$. Then the probability of hitting a root is not necessarily 0, but it is very small. In fact it is at most $(2/N)^q$, where q is the number of lines. (See J. T. Schwartz (1979, 1980).)

Thus we have designed an algorithm to determine whether or not a given graph has a perfect matching, which runs in polynomial time, but unfortunately may give the wrong answer! Fortunately, it may only err when it asserts that a graph does not have a perfect matching and in that case the probability of error can be made arbitrarily small. (For a graph with 100 lines, and with a choice of $N = 20$, the probability of an error is so small that if we run this algorithm once every second no error is likely to occur within the life span of the universe.)

In practice this *probabilistic* algorithm runs somewhat slower than the *combinatorial* matching algorithms (see Chapter 9). But in a slightly more complicated application of this idea to the matroid matching problem (see Chapter 11), the generalization of this probabilistic algorithm has much better running time bounds than the generalization of the combinatorial matching algorithms. (See Lovász (1979b).)

Algorithms which use randomly generated numbers and then give an answer which is correct with large probability (or if it is a number, with small expectation of error), have been used in numerical analysis for a long time (cf. "Monte Carlo methods"). Recently more and more applications of such methods are arising in discrete mathematics as well. Such algorithms have been applied successfully to primality testing (Rabin (1976, 1980), Solovay and Strassen (1977)), graph isomorphism testing (Babai (1979), Galil, Hoffman, Luks, Schnorr and Weber (1982)) and elsewhere. For a survey of complexity classes and other computational problems involving such algorithms, see Welsh (1983).

The formula for $\Phi(G)$ derived in Section 8.4 also suggests a Monte Carlo type algorithm for the computation of $\Phi(G)$. Fix some number N. Generate N random orientations of G, say $G^1, \ldots, G^N$. Determine

$$\Psi(G) = \frac{1}{N} \sum_{k=1}^{N} \det A_s(G^k).$$

Then if N is large enough, $\Psi(G)$ will be close to $\Phi(G)$ with large probability. The value of N which gives small relative error with large probability is $O(D/\Phi(G))$, where D is the variance of the random variable $\det A_s(G)$. Unfortunately, this ratio $D/\Phi(G)$ may be exponentially large in the size p of the problem, rendering this method useless in the general case. On the other hand, there are some graphs (for example, the complete graphs) for which $D/\Phi(G)$ is polynomially bounded. Whether there are any *more interesting* classes of such graphs is not known to the authors. It is certainly true, however, that the application of probabilistic methods to combinatorial enumeration problems in the spirit of this example is largely unexplored, and potentially very fruitful, territory.

8.5. Matching Polynomials

Let G be a graph and let $\Phi_k(G)$ denote the number of k-element matchings in G. We set $\Phi_0(G) = 1$ by convention. Thus $\Phi_1(G) = q$ is the number of lines, and if p is even then $\Phi_{p/2}(G) = \Phi(G)$ is the number of perfect matchings of G. Clearly, $\Phi_k(G) = 0$ for $k > \nu = \nu(G)$.

In this section we are concerned with properties of the sequence $(\Phi_0(G), \Phi_1(G), \dots)$. A rather standard approach to such questions is to form a generating function for the sequence and study analytic properties of this generating function. In this case, the following generating function, called the **matching defect polynomial**, turns out to have some nice properties:

$$m(G;x) = \sum_{k=0}^{\lfloor p/2 \rfloor} (-1)^k \Phi_k(G) x^{p-2k}.$$

As a motivation for the name, we write this polynomial as follows:

$$m(G;x) = \sum_M (-1)^{|M|} x^{\operatorname{def}(M)},$$

where the sum is over all matchings M in G and, as before in this book, $\operatorname{def}(M)$ denotes the defect of the matching M. (The reader should be warned that the name "matching defect polynomial" is our invention. At least three additional names for this expression are to be found in the literature; namely, "matching polynomial", "acyclic polynomial" and "reference polynomial".)

Another, perhaps more natural, polynomial to study is the **matching generating polynomial**:

$$g(G;x) = \sum_{k=0}^{\nu} \Phi_k(G) x^k.$$

Since we have the identity

$$m(G;x) = x^p g(G;(-x)^{-2}),$$

there is no essential difference between these two polynomials and we shall use them both. (Some properties of $m(G;x)$ do turn out to be somewhat nicer, however; see in particular Theorem 8.5.3 below.) In the case of bipartite graphs, $g(G;x)$ is also called the **rook polynomial**. (See Riordan (1958).) Let us note that trivially $g(G;1)$ is the total number of matchings in G. Furthermore, if G has connected components $G_1, \dots, G_r$, then

$$m(G;x) = m(G_1;x) \cdots m(G_r;x) \tag{8.5.1}$$

and a similar product identity holds for $g(G;x)$.

The starting point for our study of the matching defect polynomial is the following simple recurrence relation for the numbers $\Phi_k(G)$:

8.5.1. LEMMA. *Let G be any graph and suppose $uv \in E(G)$. Then*

$$\Phi_k(G) = \Phi_k(G-uv) + \Phi_{k-1}(G-u-v)$$

and

$$m(G;x) = m(G-uv;x) - m(G-u-v;x).$$

PROOF. $\Phi_k(G-uv)$ counts those k-element matchings in G which miss the line uv and $\Phi_{k-1}(G-u-v)$ counts those k-element matchings in G which contain the line uv. Hence the first relation holds. The second follows upon substitution into the definition of $m(G;x)$. ■

Repeated application of the above lemma yields the following result.

8.5.2. LEMMA. *Let G be a graph with $u \in V(G)$ and suppose the neighborhood of u, $\Gamma(u) = \{v_1, \ldots, v_k\}$. Then*

$$m(G;x) = x \cdot m(G-u;x) - \sum_{i=1}^{k} m(G-u-v_i;x).$$ ■

Let G be a graph and A its adjacency matrix. The **characteristic polynomial** of G, $P(G,x)$, is defined to be $\det(A-xI)$ and the set of zeros of this polynomial (i.e., the set of eigenvalues of the matrix A) is known as the **spectrum** of G. The spectrum of a graph has been extensively studied for some years by graph theorists and many interesting results have been obtained. The reader is referred to the comprehensive book on the subject by Cvetković, Doob and Sachs (1979).

An important attribute of the matching defect polynomial is that in some cases it is easy to evaluate. One such instance occurs when the graph is a tree. (See Mowshowitz (1972).)

8.5.3. THEOREM. *If G is a forest then $m(G;x)$ is equal to the characteristic polynomial of the adjacency matrix of G.*

PROOF. We may assume that G has at least one line, otherwise the assertion is trivial. Let u be an endpoint of G and v its unique neighbor. Assume that the points are labelled so that u is the first and v is the second. Let A, A' and A'' be the adjacency matrices of G, $G-u$ and $G-u-v$, respectively, and let I, I' and I'' denote the identity matrices of size $p, p-1$ and $p-2$, respectively. Then we have

$$\det(xI - A) = \begin{vmatrix} x & -1 & \cdot & \cdot & \cdot & \cdot & 0 \\ -1 & x & & & & & \cdot \\ 0 & & \cdot & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & x \end{vmatrix}.$$

If we expand this determinant by its first row, and then expand the second term by its first column, we obtain

$$\det(xI - A) = x \cdot \det(xI' - A') - \det(xI'' - A'').$$

Since the polynomial $m(G;x)$ satisfies this same recurrence relation by Lemma 8.5.2, the theorem follows by induction. ∎

8.5.4. EXERCISE. Show that the matching generating (or matching defect) polynomial of an outerplanar graph can be determined in polynomial time. (See Exercise 8.0.3.)

In general, the matching defect polynomial of a graph is different from the characteristic polynomial of its adjacency matrix. The important formula in the next theorem, due to Godsil (1981a), relates the matching polynomial of a general graph to the matching polynomial of a certain associated tree, and thereby to certain determinants. Unfortunately, this tree will be too large in general to be of any help in concrete computations, but the result will have important theoretical consequences.

Let G be any graph and suppose $u \in V(G)$. We define the **path-tree associated with G having root $\overline{u}$** as the tree whose points are all paths in G starting from u, two being adjacent (as points in the tree) if and only if one arises from the other in G by deleting its endpoint different from u. Then $\overline{u}$ will denote the path consisting of the singleton u. Similarly if $u \cdots v$ is any path, then $\overline{u \cdots v}$ denotes this path as a *point* of the path-tree. (See Figure 8.5.1.)

Note that the path-tree is always a tree and if G is a tree, its path-tree is the graph itself.

8.5.5. THEOREM. *Let G be any graph with $u \in V(G)$, and let T denote the path-tree associated with G having root $\overline{u}$. Then*

$$\frac{m(G-u;x)}{m(G;x)} = \frac{m(T-\overline{u};x)}{m(T;x)}.$$

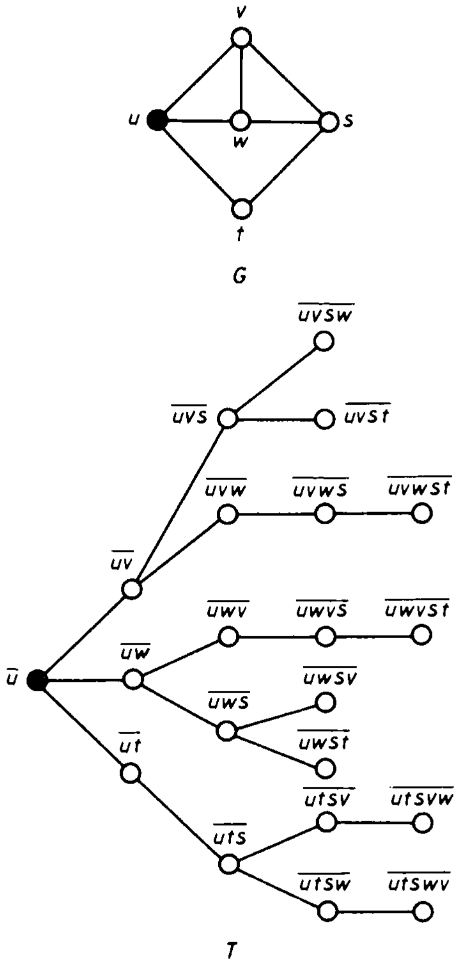


FIGURE 8.5.1. A graph and one of its path-trees

PROOF. We proceed by induction on p. If $p \leq 1$, the assertion is trivial. Now suppose $\Gamma(u) = \{v_1, \ldots, v_k\}$. Then in the path-tree T we have $\Gamma(\overline{u}) = \{\overline{uv_1}, \ldots, \overline{uv_k}\}$. Let T_i denote the branch of T with root $\overline{uv_i}$. Then we may observe the following facts.

1. T_i is isomorphic to the path-tree associated with the graph $G - u$ with root v_i. Hence by induction,

$$\frac{m(G-u-v_i;x)}{m(G-u;x)} = \frac{m(T_i - \overline{uv_i};x)}{m(T_i;x)}. \tag{8.5.2}$$

2. $T - \overline{u}$ is the disjoint union of $T_1, \ldots, T_k$, and hence

$$m(T - \overline{u}; x) = \prod_{j=1}^{k} m(T_i; x). \tag{8.5.3}$$

3. $T - \overline{u} - \overline{uv_i}$ is the disjoint union of all the T_j's, for $j \neq i$, and $T_i - \overline{uv_i}$. Hence

$$m(T - \overline{u} - \overline{uv_i}; x) = m(T_i - \overline{uv_i}; x) \cdot \prod_{j \neq i} m(T_j; x). \tag{8.5.4}$$

Based on these three formulas, we obtain from Lemma 8.5.2,

$$\frac{m(G; x)}{m(G - u; x)} = x - \sum_{i=1}^{k} \frac{m(G - u - v_i; x)}{m(G - u; x)}$$

(by (8.5.2))

$$= x - \sum_{i=1}^{k} \frac{m(T_i - \overline{uv_i}; x)}{m(T_i; x)}$$

(by (8.5.3) and (8.5.4))

$$= x - \sum_{i=1}^{k} \frac{m(T - \overline{u} - \overline{uv_i}; x)}{m(T - \overline{u}; x)}$$

(by Lemma 8.5.2)

$$= \frac{m(T; x)}{m(T - \overline{u}; x)}.$$

■

8.5.6. THEOREM. *Suppose G is a connected graph with $u \in V(G)$, and suppose T is the path-tree associated with G with root $\overline{u}$. Then T has a subforest T' such that*

$$m(G; x) = \frac{m(T; x)}{m(T'; x)}.$$

PROOF. Again we use induction on $|V(G)|$. Let $G_1, \ldots, G_k$ be the connected components of $G - u$ and let v_i be a neighbor of u in G_i. Let T_i denote the branch of T with root uv_i. Note that T_i is isomorphic to the path-tree associated with G_i having root v_i. Thus by the induction hypothesis, there exist subforests $T_i' \subseteq T_i$ such that

$$m(G_i; x) = \frac{m(T_i; x)}{m(T_i'; x)},$$

for $i = 1, \ldots, k$. Furthermore, $T - \overline{u}$ has components $T_1, \ldots, T_k$ and possibly even more, say, $T_{k+1}, \ldots, T_r$. Hence by Theorem 8.5.5,

$$\begin{aligned}
m(G;x) &= \frac{m(T;x)}{m(T-\overline{u};x)} \cdot m(G-u;x) \\
&= \frac{m(T;x)}{\prod_{i=1}^{r} m(T_i;x)} \prod_{i=1}^{k} m(G_i;x) \\
&= \frac{m(T;x)}{\prod_{i=1}^{r} m(T_i;x)} \cdot \frac{\prod_{i=1}^{k} m(T_i;x)}{\prod_{i=1}^{k} m(T'_i;x)} \\
&= \frac{m(T;x)}{\prod_{i=1}^{k} m(T'_i;x) \prod_{i=k+1}^{r} m(T_i;x)} \\
&= \frac{m(T;x)}{m(T'_1 \cup \cdots \cup T'_k \cup T_{k+1} \cup \cdots \cup T_r;x)}.
\end{aligned}$$

■

REMARK 1. It is not difficult to derive the following description of an appropriate subforest T': find a spanning tree T_0 of G by depth-first search from u, and delete from T all points corresponding to a path contained in T_0.

REMARK 2. Substituting $i/\sqrt{x}$ for x in the formula of Theorem 8.5.6, we obtain a similar representation result for the matching generating polynomial.

The following corollary, which is an important consequence of the preceding theorem, was first proved by Heilman and Lieb (1970, 1972) and, independently, by Kunz (1970) via different methods. This corollary was motivated by applications in chemistry (see Section 8.7), but we shall see that it also has important mathematical consequences.

8.5.7. COROLLARY. *All roots of the matching defect polynomial (and hence of the matching generating polynomial) are real.*

PROOF. For trees, this follows immediately from Theorem 8.5.3. But then it follows for all connected graphs by Theorem 8.5.6, and hence for all graphs by equation (8.5.1). ■

Let us add that, trivially, all roots of the matching generating polynomial are negative, while the roots of the matching defect polynomial lie symmetrically placed with respect to 0.

By a similar method we can obtain the following bound on the roots of $m(G;x)$, also due to Heilman and Lieb (1972):

8.5.8. THEOREM. *Let G be a graph with $\Delta(G) > 1$ and let t be any root of $m(G;x)$. Then*

$$t \leq 2\sqrt{\Delta(G)-1}.$$

PROOF. We first prove the result in the case when G is a tree. By Theorem 8.5.3, the roots of $m(G;x)$ are then the eigenvalues of the characteristic polynomial of the adjacency matrix of G (which we simply call "the eigenvalues of G"). Since this adjacency matrix is non-negative, its largest eigenvalue is also largest in *absolute* value. In this case we will show that the largest eigenvalue of any tree G is less than $2\sqrt{\Delta(G)-1}$.

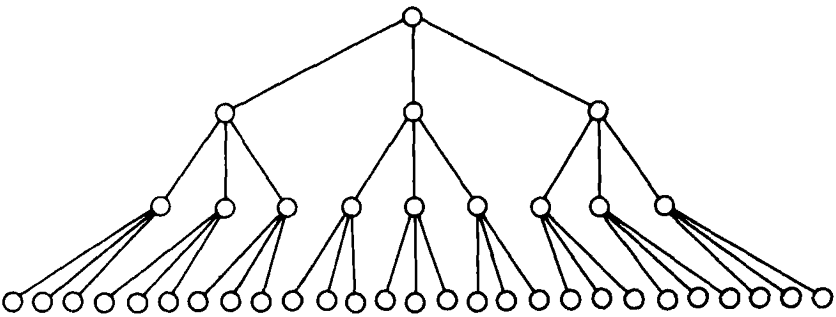


FIGURE 8.5.2. A full 3-ary tree of depth 3

The largest eigenvalue of a graph is not less than the largest eigenvalue of any subgraph. So it suffices to prove the inequality in the case when G is a full d-ary tree of depth k, where $d = \Delta(G) - 1$ and $k \geq 2$, since obviously every tree with maximum degree $d+1$ can be embedded in a full d-ary tree. (See Figure 8.5.2.)

The eigenvalues of a full d-ary tree of depth k can be calculated using the same method used to calculate the eigenvalues of a path, that is, a "1-ary tree" (cf. Lovász (1979c, Exercise 11.5)). We find that they are $2\sqrt{d}\,cos(m\pi/(k+1))$, $m = 1, \ldots, k$. Hence the largest eigenvalue of G is less than $2\sqrt{d}$ as claimed.

The general case now follows easily using Theorem 8.5.6. For any $u \in V(G)$, we have $m(G;x) \mid m(T;x)$. Obviously, $\Delta(T) \leq \Delta(G)$ and, since any root of $m(G;x)$ is also a root of $m(T;x)$, it follows that every root of $m(T;x)$ is less than $2\sqrt{\Delta(G)-1}$, by the special case already proved above. But then the same inequality holds for every root of $m(G;x)$. ∎

8.5.9. EXERCISE. Prove that the polynomial $m(G;x)$ has a root t with $t \geq \sqrt{\Delta(G)}$.

8.5.10. EXERCISE. Use Corollary 8.5.7 to prove that the sequence $\Phi_0(G), \Phi_1(G), \dots$ is **log-concave**. That is, show for every graph G and for all $i \geq 1$ that: $\Phi_{i-1}(G)\Phi_{i+1}(G) \leq (\Phi_i(G))^2$.

There are many similarly defined sequences in combinatorics which are believed to be log-concave. Two examples are the sequence $a_0, a_1, \dots$, where a_k denotes the number of k-element independent sets in a matroid and the sequence $W_0, W_1, \dots$, where W_k is the number of rank k flats in a matroid. (See, for example, Brylawski (1982).) Most of these conjectures seem to be very difficult. The fact that the log-concavity of $\Phi_k(G)$ can be proved fairly easily using Corollary 8.5.7 illustrates the strength of the latter result.

A more significant application of Corollary 8.5.8 is due to Godsil (1981b). Before stating this result, a bit of groundwork must be laid. Let G be any graph and let us choose a matching M of G at random. (Here we assume that every matching of G, including $\emptyset$, has equal probability of being selected. Then $\xi = |M|$ is a random variable. Let $\mu(G)$ and $\sigma^2(G)$ denote the expectation and the variance of ξ, respectively. It is not difficult to see that $\mu(G)$ and $\sigma^2(G)$ can be expressed in terms of the numbers $\Phi_k(G)$ as follows:

$$\mu(G) = \frac{\sum_{k=1}^{\lfloor p/2 \rfloor} k\Phi_k(G)}{\sum_{k=1}^{\lfloor p/2 \rfloor} \Phi_k(G)},$$

$$\sigma^2(G) = \frac{\sum_{k=1}^{\lfloor p/2 \rfloor} (k-\mu(G))^2\Phi_k(G)}{\sum_{k=1}^{\lfloor p/2 \rfloor} \Phi_k(G)}.$$

We derive two more formulas for $\mu(G)$ and $\sigma^2(G)$. Let us write the matching generating polynomial as follows:

$$g(G;x) = (1+r_1x)\cdots(1+r_\nu x) \tag{8.5.7}$$

where $-1/r_1,\ldots,-1/r_\nu$ are the roots of this polynomial. Thus by Corollary 8.5.8, $r_1,\ldots,r_\nu$ are positive real numbers. Clearly, $\Phi_k(G)$ is the kth elementary symmetric polynomial in the variables r_i. Furthermore, we have

$$\mu(G) = \sum_{i=1}^{\nu} \frac{r_i}{r_i+1} \tag{8.5.8}$$

and

$$\sigma^2(G) = \sum_{i=1}^{\nu} \frac{r_i}{(r_i+1)^2}. \tag{8.5.9}$$

In fact, the number of matchings in G is

$$g(G;1) = (1+r_1)\cdots(1+r_\nu)$$

and

$$\begin{aligned}\sum_{k=1}^{\nu} k\Phi_k(G) &= g'(G;1)\\ &= \sum_{i=1}^{\nu}(1+r_1)\cdots(1+r_{i-1})r_i(1+r_{i+1})\cdots(1+r_\nu)\\ &= g(G;1)\sum_{i=1}^{\nu}\frac{r_i}{1+r_i}.\end{aligned}$$

Hence equation (8.5.8) follows.

Equation (8.5.9) can be derived in a similar way.

8.5.11. EXERCISE. Prove that $\sigma^2(G) < \mu(G)$.

8.5.12. EXERCISE. Prove that $\mu(G) \geq \nu(G)/3$. (Babai; see Godsil (1981b)).

8.5.13. EXERCISE. Let G be the union of k point-disjoint copies of the n-star. Show that $g(G;1) = (n+1)^k, \mu(G) = nk/(n+1)$, and $\sigma(G) = \sqrt{nk}/(n+1)$.

8.5.14. EXERCISE. Suppose $G = K_n$. Define the sequence $a_0 = a_1 = 1, a_2, a_3, \ldots$ by the recurrence $a_{n+1} = a_n + na_{n-1}$. Prove that $g(G;1) = a_n, \mu(G) = na_{n-1}/a_n, E(\xi_G^2) = n$ and also that $\sigma^2(G) = n(1-na_{n-1}^2/a_n^2)$. Also show that $\mu(G) \sim n/2$ and $\sigma(G) \sim n^{1/4}/2$. (Difficult! See Godsil (1981b).)

Let $G_1, G_2, \ldots$ be a sequence of graphs. We say that the distribution of matching sizes in G_n is **asymptotically normal** if the distribution of ξ_{G_n} is asymptotically normal as $n \to \infty$, that is, if

$$P\left(\frac{\xi_{G_n} - \mu(G_n)}{\sigma(G_n)} < x\right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt \tag{8.5.10}$$

for each real number x. It is easy to see that a necessary condition for this to hold is that $\sigma(G_n) \to \infty$ as $n \to \infty$; in fact, the left hand side of (8.5.10) is a step function whose jumps are at least $1/\sigma(G_n)$ apart. Such functions can tend to a strictly monotone function only if $1/\sigma(G_n) \to 0$.

The next result shows that this condition is also sufficient. In fact, if $\sigma(G_n) \to \infty$ then not only does (8.5.10) hold, but the following stronger assertion holds as well:

$$\Phi_k(G_n) \cdot \frac{\sigma(G_n)}{g(G_n;1)} \to \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \tag{8.5.11}$$

if $k, n \to \infty$ such that $k - \mu(G_n) \sim x\sigma(G_n)$ and x is fixed. If (8.5.11) holds, we shall say that the distribution ot matching sizes in G_n is **asymptotically locally normal**. It is easy to see that conditions (8.5.10) imply condition (8.5.11). The following important result is due to Godsil (1981b).

8.5.15. THEOREM. *Let $G_1, G_2, \ldots$ be any sequence of graphs such that $\sigma(G_n) \to \infty$. Then the distribution of matching sizes in G_n is asymptotically locally normal.*

PROOF. Let $G = G_n$ and let $r_1, \ldots, r_\nu$ be as in equation (8.5.7). Let $\eta_1, \ldots, \eta_\nu$ be independent random variables such that

$$P(\eta_i = 0) = \frac{1}{r_i + 1}, \quad P(\eta_i = 1) = \frac{r_i}{r_i + 1}.$$

Claim. $\eta_1 + \cdots + \eta_\nu$ has the same distribution as ξ_G. In fact,

$$\begin{aligned} P(\eta_1 + \cdots \eta_\nu = k) &= \sum_{\substack{I \subseteq \{1,\ldots,\nu\} \\ |I|=k}} \prod_{i \in I} \frac{r_i}{1 + r_i} \cdot \prod_{i \notin I} \frac{1}{1 + r_i} \\ &= \frac{1}{\prod_{i=1}^{\nu}(1 + r_i)} \sum_{\substack{I \subseteq \{1,\ldots,\nu\} \\ |I|=k}} \prod_{i \in I} r_i \\ &= \frac{\Phi_k(G_n)}{g(G_k;1)} \quad = \quad P(\xi_{G_n} = k), \end{aligned}$$

by the definition of the r_i. This proves the claim.

Thus ξ_{G_n} can be represented as the sum of independent random variables, each taking only values 0 or 1, such that the variance of ξ_{G_n} tends to ∞. By a version of the Central Limit Theorem (see Feller (1950)) the assertion follows. ■

Several combinatorial sequences are known to be asymptotically normally distributed: the binomial coefficients $\binom{n}{0}, \binom{n}{1}, \ldots$, the absolute values of the Stirling numbers of the first kind, $|s(n,0)|, |s(n,1)|, \ldots$, and the Stirling numbers of the second kind, $S(n,0), S(n,1), \ldots$. Theorem 8.5.15 is a remarkably general result of this kind which implies the asymptotically locally normal distribution of several sequences by considering special sequences of graphs. For example, let G_n be the bigraph on $2n$ points $u_1, \ldots, u_n, v_1, \ldots, v_n$ where $u_i v_j \in E(G_n)$ if and only if $i \le j$. Then $\Phi_k(G_n) = |s(n,k)|$ and so the asymptotic local normality of the distribution of the sequence $\{|s(n,k)|\}$ can be derived using Theorem 8.5.15.

The condition that $\sigma(G_n) \to \infty$ is sometimes quite difficult to verify. One noteworthy special case was found by Godsil (1981b). It depends on the following lower bound for $\sigma(G)$:

8.5.16. LEMMA. *For every graph G having q lines and maximum degree $\Delta(G)$,*

$$\sigma(G) \quad \ge \quad \frac{\sqrt{q}}{4\Delta(G)-3}.$$

PROOF. The assertion is trivial if $\Delta(G) \le 1$, so suppose that $\Delta(G) > 1$. Using the notation of equation (8.5.7), we have by equation (8.5.9) and by Theorem 8.5.8,

$$\sigma^2(G) \quad = \quad \sum_{i=1}^{\nu} \frac{r_i}{(1+r_i)^2}$$

$$\ge \quad \sum_{i=1}^{\nu} \frac{r_i}{(4\Delta(G)-3)^2} \quad = \quad \frac{q}{(4\Delta(G)-3)^2}.$$ ■

8.5.17. COROLLARY. *Let G_n be a sequence of graphs such that $\Delta^2(G_n)/|E(G_n)| \to 0$. Then the distribution of matching sizes in G_n is asymptotically locally normal.* ■

REMARK. The sufficient condition given in Corollary 8.5.17 is not necessary. For example, if G_n is the complete n-graph, $(n \geq 3)$, then by Exercise 8.5.14, $\sigma(G_n) \to \infty$, but $\Delta^2(G_n)/|E(G_n)| > 1$. A trivial *necessary* condition for $\sigma(G_n) \to \infty$ is that $\nu(G) \to \infty$ (since otherwise ξ_{G_n} is the sum of a bounded number of 0–1 variables and so its distribution cannot tend to a continuous distribution). But this condition is not sufficient, as shown by Exercise 8.5.13. No simple graph-theoretic equivalent of $\sigma(G_n) \to \infty$ is known.

8.6. More on the Number of Perfect Matchings

In the preceding sections we have derived exact formulas, as well as various upper and lower bounds, for the number of perfect matchings in special classes of graphs, (in fact, mostly for the case of regular graphs). In this section we derive some additional lower bounds, this time by using more elementary methods, in which we make use of the structure theory developed in the preceding chapters.

The starting point of our discussion is a result of Kotzig (1959a) which asserts that if a 2-connected graph has a perfect matching then it has at least two. (The proof, which is now very easy using the structure theory developed in Chapter 5, is left to the reader.) Beineke and Plummer (1967) generalized this result by showing that any n-connected graph containing a perfect matching must have, in fact, at least n such. Zaks (1969, 1971) improved this result by proving the following.

8.6.1. THEOREM. *Let G be a k-connected graph which contains at least one perfect matching. Then the number of perfect matchings in G is at least*

$$k!! = \prod_{i=0}^{\lfloor (k-2)/2 \rfloor} (k-2i).$$

If k is odd, the complete graph K_{k+1} serves to show that the above bound is sharp. In fact, if k is odd, K_{k+1} is the *unique* extremal graph (Mader (1973)).

On the other hand, if k is even, the above bound is not sharp, but the best possible bound *is known.* No simple closed formula for this bound is available, but the bound can be recursively defined as follows. Since k is even, write $k = 2m$ and let S_m denote the graph obtained from K_{2m+2} by deleting a perfect matching. It is easy to see that $\Phi(S_0) = 0$, $\Phi(S_1) = 2$, and in general, $\Phi(S_m) = 2m(\Phi(S_{m-1}) + \Phi(S_{m-2}))$. Mader (1973) showed, in fact, that if G is as in Theorem 8.6.1 with k even, then

$\Phi(G) \geq \Phi(S_{k/2})$. Moreover, he showed that the bound is sharp and that $S_{k/2}$ is the *unique* extremal graph for $k \geq 4$. (See also Bollobás (1978a) and (1978b, pp. 63–67).)

If a k-connected graph G is not bicritical or if it has enough points, the lower bound in Theorem 8.6.1 can be dramatically improved. More precisely we have the next result.

8.6.2. THEOREM. *Let G be a k-connected non-bicritical graph containing a perfect matching. Then G contains at least $k!$ perfect matchings.*

Mader (1976) proved that in Theorem 8.6.2 the hypothesis of k-connectivity may be weakened to assuming only that the graph is 2-connected and that all degrees are at least k.

The definition of the bicritical property seems to imply that graphs with this property must have a large number of perfect matchings. It is therefore surprising that just the converse is true; they are the exceptional class in Theorem 8.6.2! Recall, however, from Corollary 7.6.10 that we do have the lower bound of $p/2+1$ on the number of perfect matchings in a bicritical graph. Together with Theorem 8.6.2 this implies the following result.

8.6.3. COROLLARY. *There exists a function of k, $p_0(k)$, such that if G is a k-connected graph containing a perfect matching and having at least $p_0(k)$ points, then G contains at least $k!$ perfect matchings.* ∎

It is easy to construct, for every $n \geq k$, a k-connected graph on $2n$ points which contains exactly $k!$ perfect matchings: take a complete bipartite graph with k and $2n-k$ points in the two color classes respectively, and add $n-k$ disjoint lines in the color class of size $2n-k$.

The optimal value of $p_0(k)$ is not known. The argument before the corollary above gives $p_0(k) = 2k!$, which is probably a very poor value.

The proof of Theorem 8.6.1 will depend on a result also due to Zaks (1971), which is interesting in itself. A point of a graph G is called **totally covered**, if every line incident with this point belongs to some perfect matching of G.

8.6.4. LEMMA. *Let G be a 2-connected graph which contains a perfect matching. Then G has a totally covered point.*

PROOF. We assume without loss of generality that G is saturated. Let G_0 be defined as in the proof of the Cathedral Theorem 5.3.8. Then G_0 is a saturated elementary graph by Theorem 5.3.8, and so by Exercise 5.3.5, $\mathcal{P}(G)$ has a class which is a singleton set $\{x\}$. This point x cannot be the base of a tower as G is 2-connected and so it is totally covered. ∎

For more recent improvements on the above result, see Gabow (1979).

PROOF (of Theorem 8.6.1). We use induction on k. For $k \leq 1$ the assertion is trivially true, so suppose that $k \geq 2$. By Lemma 8.6.4, G has a totally covered point x. Since G is k-connected, x has at least k neighbors $y_1, \ldots, y_d$ $(d \geq k)$. Since x is totally covered, every line xy_i is contained in a perfect matching or, equivalently, $G - x - y_i$ has a perfect matching. But $G - x - y_i$ is at least $(k-2)$-connected, so by the induction hypothesis $G - x - y_i$ contains at least $(k-2)!!$ perfect matchings. Hence

$$\Phi(G) = \sum_{i=1}^{d} \Phi(G - x - y_i) \geq d(k-2)!! \geq k(k-2)!! = k!!.$$

■

To prove Theorem 8.6.2 we shall need an estimate on the number of perfect matchings in a bigraph, due to M. Hall (1948).

8.6.5. LEMMA. *Let G be a simple bipartite graph with bipartition (A, B), and assume that each point in A has degree at least k. Then if G has at least one perfect matching, it has at least $k!$ perfect matchings.*

PROOF. The proof is similar to the Halmos-Vaughn proof of P. Hall's Theorem 1.1.3. We use induction on $|A|$, and distinguish between two cases.

Case 1. Assume that there exists a set $X \subset A$ such that $X \neq \emptyset$ and $|\Gamma(X)| = |X|$. Let G_1 be the subgraph spanned by $X \cup \Gamma(X)$ and let G_2 be the subgraph spanned by $V(G) - X - \Gamma(X)$. Just as in the Halmos-Vaughn proof of P. Hall's theorem, it follows that each of G_1 and G_2 contains a perfect matching. Furthermore, G_1 satisfies the same conditions as G, and so by the induction hypothesis G_1 contains at least $k!$ perfect matchings. Taking the union of these $k!$ perfect matchings with any perfect matching of G_2, we obtain $k!$ perfect matchings of G.

Case 2. Assume that $|\Gamma(X)| > |X|$ for every $X \subset A$, $X \neq \emptyset$. Let $x \in A$ and let $y_1, \ldots, y_d$ be the neighbors of x $(d \geq k)$. Each of the graphs $G - x - y_i$ satisfies the condition in P. Hall's theorem and so each has a perfect matching. Furthermore, each point in $A - x$ has degree at least $k-1$ in $G - x - y_i$, and so $G - x - y_i$ has at least $(k-1)!$ perfect matchings by the induction hypothesis. Hence

$$\Phi(G) = \sum_{i=1}^{d} \Phi(G - x - y_i) \geq d(k-1)! \geq k(k-1)! = k!.$$

■

PROOF (of Theorem 8.6.2). Without loss of generality we may assume that G is saturated. Consider the Cathedral structure of G described in Section 5.3 and let G_0 be the foundation. We claim that G_0 cannot be bicritical. This is trivial if $G = G_0$. Otherwise, there is at least one tower in the Cathedral structure and since G is 2-connected, the base of this tower has at least two elements. Thus $\mathcal{P}(G_0)$ has a class with at least two elements and so G_0 is not bicritical.

Let $S \in \mathcal{P}(G_0)$, and suppose $|S| \geq 2$. Then S is a cutset of G and so $|S| \geq k$. Furthermore, if T is any connected component of $G_0 - S$ then T must be adjacent to at least k points of S, since the neighbors of T in S separate G. Thus if we form the bipartite graph G_S' as in Section 5.2, every point in the color class $V(G_S') - S$ will have degree k. So by Lemma 8.6.5, G_S' has at least $k!$ perfect matchings. By Theorem 5.2.2(d) we obtain at least $k!$ perfect matchings of G_0. Since by Theorem 5.3.8 each of these extends to a perfect matching of G, we conclude that G contains at least $k!$ perfect matchings. ■

8.6.6. EXERCISE. Prove that every k-connected graph ($k \geq 2$) with a perfect matching contains at least k totally covered points.

8.6.7. EXERCISE. Use Theorem 5.4.6 (the Two Ear Theorem for 1--extendable graphs) to show that every 1-extendable graph with q lines and p points contains at least $(q-p)/2+2$ perfect matchings, and show that this bound could replace that of Corollary 7.6.10 in the proof of Corollary 8.6.3.

8.6.8. EXERCISE. Prove that if G is connected and has a unique perfect matching, then it has a cutline which belongs to this perfect matching.

8.6.9. EXERCISE. If G has a unique perfect matching then it has a point with degree $\log_2(p+2)$. Also prove that it has at most $p^2/4$ lines.

8.6.10. EXERCISE. Prove that if M is any matching in a 2-connected graph G, then G has a cycle C such that every line of M is either a line of C or point-disjoint from C. Give a solution to Exercise 8.6.7 based on this observation.

Let r and n be positive integers with $r < n$. An $r \times n$ array is called a **Latin rectangle** if each row contains each of the integers $1, \ldots, n$ exactly once and each column contains each of $1, \ldots, n$ at most once. If $r = n$, the corresponding square array is called a **Latin square**. The study and application of Latin squares has a long and fruitful history in classical combinatorics. (See Dénes and Keedwell (1974).)

8.6.11. EXERCISE. (a) Using Lemma 8.6.5, show that each $r \times n$ Latin rectangle can be extended to an $(r+1) \times n$ Latin rectangle in at least $(n-r)!$ different ways. (b) If $L(n)$ denotes the number of $n \times n$ Latin squares, use part (a) to prove that

$$L(n) \geq \prod_{k=1}^{n} k!.$$

(c) Using Theorem 8.1.3, prove that

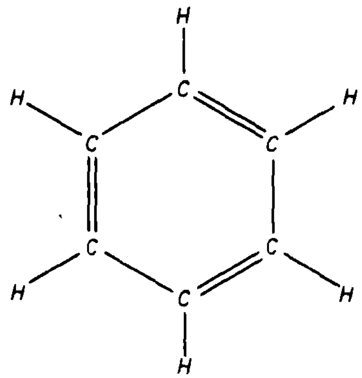
$$L(n) \geq \frac{(n!)^{2n}}{n^n}.$$

(d) Which is the better bound?

8.7. Two Applications to Physical Science

For many years chemists have realized that the study of certain properties of a chemical compound can be effectively pursued by studying the **topological model** of the molecule. This is nothing more than the underlying undirected graph of the molecule in which atoms are represented by points and chemical bonds by lines. Moreover, hydrogen atoms are frequently ignored when they correspond to points of degree one in the graph. In particular, organic chemists have begun to study the graphs of "conjugated" compounds; that is, compounds the molecules of which possess an alternating pattern of single and double bonds. They have christened the graphs of such compounds **Hückel graphs**. A simple example is furnished by the hydrocarbon benzene, which is depicted in Figure 8.7.1, together with its topological representation (in which degree one hydrogen atoms have been suppressed).

Trivially, every chemical compound has a Hückel graph. It is more interesting, however, to consider a converse question. When does an abstract graph represent a chemical compound which exists in the real world? This existence question is naturally one of the basic concerns of chemists as is the related question of stability and reactivity of chemical compounds. Various measures of stability have been defined and studied over the years. One of the most significant bodies of work in this area is that known as "resonance theory". We cannot present more than a glimpse of this theory, but we recommend the surveys of Herndon (1974) and Gutman (1982) and the recent book of Trinajstić (1983). Each contains exhaustive bibliographies on the subject as well.



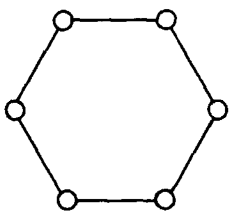


FIGURE 8.7.1. Benzene and its associated Hückel graph

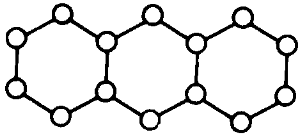
The most important facts to be gleaned from resonance theory for our purposes are two in number. Not surprisingly, they involve perfect matchings (which are called **Kekulé structures** by chemists).

(1) Certain classes of chemical compounds have been successfully synthesized only when the graph of the compound has a perfect matching.

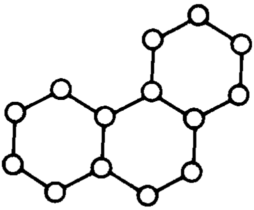
(2) It has been observed that for members of some families which have been synthesized and/or found in nature, the more perfect matchings possessed by their graphs, the more stable are the corresponding compounds.

The relationship between stability and the number of perfect matchings is not fully understood. One attempt to ascertain the nature of this dependence is due to Carter (1949) — and independently some time later to Swinborne-Sheldrake, Herndon and Gutman (1975) — who conjectured a dependence in which the "resonance energy" is proportional to $\log \Phi(G)$.

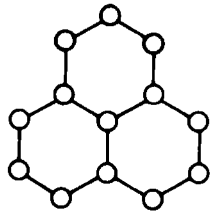
One of the most extensively studied families of chemical structures, with respect to resonance energy, is the family of **benzenoids**. The graphs of this familiy are especially simple to describe. They are obtained by arranging congruent hexagons in the plane in such a way that any two hexagons are either point-disjoint or possess exactly one common line.



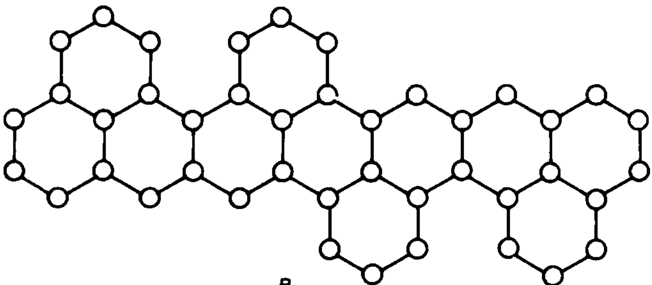
B_1
anthracene



B_2
phenanthrene



B_3
benzo[*de*]napthalene



B_4
tetrabenza[*de*, *hi*, *st*, *wx*]heptacene

FIGURE 8.7.2. Four benzenoid graphs

We illustrate with the graphs in Figure 8.7.2. The reader may have encountered these structures elsewhere in graph theory under the rubric "hexagonal animals". (See Harary (1968).)

It is easy to see that benzenoids B_1 and B_2 have four and five perfect matchings respectively, while B_3 and B_4 have none. Experimentally speaking, B_2 has been determined to be more stable than B_1, but both have been synthesized whereas B_3 and B_4 have not. We have included both B_3 and the much larger B_4 for the following reason. We see readily that all benzenoids are bipartite and hence in order to have perfect matchings their two color classes must be the same size. Hence B_3 cannot have a perfect matching. On the other hand, B_4 destroys a possible conjecture that equal size color classes is a *sufficient* condition for the existence of a perfect matching in a benzenoid. (See Exercise 1.1.6.) Of course the existence of a perfect matching in a benzenoid can be tested using a bipartite matching algorithm, but a much simpler procedure has been developed for this special family. (See Sachs (1984).)

One significant attempt at a precise mathematical formulation for resonance energy may be described as follows. (See Gutman, Milun and Trinajstić (1975), Aihara (1976) and Gutman (1977).) First recall that the roots of the characteristic polynomial $P(G,x)$ are real since A is a real symmetric matrix, so let us order them $x_1 \le x_2 \le \cdots \le x_p$. From Corollary 8.5.7 we know that the roots of the matching defect polynomial $m(G;x)$ are also real, so let us order them similarly as $x_1^d \le x_2^d \le \cdots \le x_p^d$. Now we define the **topological resonance energy** (*TRE*) of G (and its parent chemical compound) as:

$$TRE(G) = \sum_{j=1}^{p} g_j(x_j - x_j^d). \tag{8.7.1}$$

Here each g_j is a parameter known as the *occupation number of the jth molecular orbital* and is equal to the number of electrons (0, 1 or 2) in that orbital.

Experience shows that the *TRE* is closely related to the stability of the molecule.

Note that $TRE(G)$ is defined for any undirected graph G, but its behavior in such a general setting is far from understood. However, for the benzenoid hydrocarbons, it has been shown that $TRE(G)$ is mainly determined by $\Phi(G)$ and the number of hexagons $h(G)$ in the graph G. (See Gutman and Mohar (1981) and Gutman (1982).)

We shall not venture any further into the labyrinths of chemistry, but suffice it to say that the fact that *TRE* is now known to be a real number lends support to the validity of its use to measure chemical stability. Next we discuss another application of matching theory, the so-called **Ising model** of magnetic materials. This model has been studied extensively and has led to important results on ferromagnetism, phase transitions and other phenomena in statistical mechanics. It is obvious that we shall have to restrict ourselves to the simplest possible versions of these models, and we shall try to emphasize the graph-theoretic background of the results.

Now let us describe this model of a ferromagnetic material due to Ising (1925). Let the atoms forming the material be the points of a graph. We assume that only "nearby" atoms interact directly. (Precisely what "nearby" means is irrelevant at this point.) We obtain a graph G by connecting two atoms if their interaction is non-negligible. In most applications, we assume that the graph G has a very simple structure, usually a planar or 3-dimensional rectangular grid, but this too is irrelevant at the moment.

Suppose that each atom has a "spin" which is pointing either "up" or "down". We describe this spin by a value $\sigma_i \in \{-1, +1\}$. If i and j are two adjacent points (that is, two atoms whose interaction is not negligible), then the interaction energy of this pair is defined to be $J_{ij}\sigma_i\sigma_j$. That is, the energy is J_{ij} if the two atoms have the same spin and $-J_{ij}$ if their spins differ. Here the coefficient J_{ij} depends on the nature of the atoms i and j. Then we define the energy of a given spin configuration $S = \{i \mid \sigma_i = 1\} \subseteq V(G)$ to be

$$H(S) = -\sum_{ij \in E(G)} J_{ij}\sigma_i\sigma_j.$$

This term is known as the **Hamiltonian** of the system.

One can give the following physical interpretation of the direction of spins. If every σ_i is equal, that is, if the spins of all atoms point in the same direction, then the entire material is in a magnetic state. Consider a material in which every J_{ij} is positive. Then a spin configuration with all states equal minimizes $H(S)$ and hence is called a "ground state" of the material. At low temperatures, the material will be in this state and hence will be magnetic. Such materials are called **ferromagnetic**.

But what happens if some of the J_{ij}'s are negative? Suppose for example that the material is composed of two kinds of atoms, and J_{ij} is positive only between two atoms of the first kind. How will the material behave at very low temperatures? Which spin configuration minimizes $H(S)$?

The connection between this question and matching theory was discovered by Bieche, Maynard, Rammal and Uhry (1980) and further exploited by Barahona (1982). Let us transform the Hamiltonian $H(S)$ as follows:

$$H(S) = -\sum_{ij \in E(G)} J_{ij} + 2\sum_{ij \in \nabla(S)} J_{ij},$$

where S is the set of points in G with $\sigma_i = 1$. Since the first term is constant over all spin configurations S, it suffices to find a spin configuration (or equivalently, a set $S \subseteq V(G)$) which minimizes the second term.

Thus the problem of finding a **ground state**, that is, a state in which $H(S)$ is minimum, is equivalent to finding a minimum weight cut in the graph G with the values J_{ij} considered as weights. If all weights are non-negative, then of course $\emptyset$ is a minimum weight cut. (Here we allow $S = \emptyset$ and so $\nabla(S) = \emptyset$.) But if negative weights are allowed then this problem is NP-hard, as was remarked in Section 6.6. In that same section, however, we also saw how to use matching theory to solve this problem in polynomial time in the case when G is planar. Since planar grids are often used as models for very thin layers of material, this special case is of considerable interest in physics. This approach has been applied successfully to determine ground states and to describe the "morphology" of all ground states by Barahona, Maynard, Rammal and Uhry (1982).

In statistical physics, the **partition function** of the system plays a most important role. In our model, this function is defined by

$$f(T) = \sum_{S \subseteq V(G)} \exp(-H(S)/KT),$$

where K is the Boltzmann constant and the variable T is interpreted as the temperature.

We shall not go into the applications of the partition function, but instead we shall show that the following question can be answered using matching theory: how can we determine the value of the partition function for a given positive real number T.

Let us transform this expression as follows:

$$\begin{aligned} f(T) &= \sum_{S \subseteq V(G)} \exp(-H(S)/KT) \\ &= \prod_{e \in E(G)} \exp(J_e/KT) \cdot \sum_{S \subseteq V(G)} \prod_{e \in \nabla(S)} \exp(-2J_e/KT) \\ &= A \sum_{S \subseteq V(G)} \prod_{e \in \nabla(S)} x_e, \end{aligned}$$

where $A = \exp\left(\prod_{e \in E(G)} J_e/KT\right)$ and $x_e = \exp\left(-2J_e/KT\right)$. Here A is easy to compute, but what about the second term? To be able to say something about it, let us assume that the graph is planar. As we have seen before, this is a very restrictive assumption, but on the other hand, it includes important real-world situations.

Let G^* be the planar dual graph of G. Then, as seen already in Section 6.6, the sets of the form $\nabla(S)$ $(S \subseteq V(G))$, (i.e., the cuts of G), correspond to the Eulerian subgraphs (i.e., the $\emptyset$-joins) of G^*. Thus we have

$$f(T) = A \sum_{\substack{J \text{ is a} \\ \emptyset-\text{join in } G^*}} \prod_{e \in J} x_e.$$

We have seen before that an expression like this can be written as the Pfaffian of an appropriate matrix, and can therefore be evaluated for any given x in polyomial time using Kasteleyn's method for enumerating perfect matchings in planar graphs. (See Exercise 8.3.9.)

For further reading about the Ising problem the reader is directed to Brush (1964) and to Kasteleyn (1967).

This Page Intentionally Left Blank

9

Matching Algorithms

9.0. Introduction

In Chapter 1, we presented an algorithm for finding a maximum matching in a bipartite graph. From a mathematical point of view, this algorithm was essentially no more involved than the proof of König's Minimax Theorem. In fact, it followed the same lines as König's original proof.

For non-bipartite graphs the situation is quite different. Known polynomial-time algorithms for finding a maximum matching in a general graph are among the most involved of combinatorial algorithms. Most of them are based on augmentation along alternating paths just as are most proofs of Tutte's Theorem and the Berge Formula. But important new ideas are needed to turn these minimax results into a polynomial-time algorithm. The first polynomial-time matching algorithm for non-bipartite graphs was constructed by Edmonds (1965a). In this algorithm the key idea of "shrinking" certain odd cycles was introduced. Up to the present time most matching algorithms — certainly the most successful ones — are based (implicitly or explicitly) on this idea. We discuss a version of Edmonds' Algorithm in Section 9.1. As in Chapter 1, our main interest lies in the underlying mathematical ideas rather than in details of implementation, but this time we shall delve more deeply into the latter (especially into the associated data structures), for the problems arising here are far from trivial.

In Section 9.2 we extend this algorithm to the weighted case. The first polynomial algorithm to find a maximum weight matching in a graph was also found by Edmonds (1965b).

In Sections 9.3 and 9.4 we discuss matching algorithms based on ideas other than shrinking odd cycles. The algorithm presented in Section 9.3 is motivated by the Gallai-Edmonds Structure Theorem. While somewhat less efficient than the Edmonds Algorithm, it is based on somewhat more general ideas and may serve as an example of how a structure theorem can be turned into an efficient algorithm. In Section 9.4 we use the Ellipsoid Method to solve the matching problem. The resulting algorithm, although polynomial, is very slow. On the other hand, this method is based on ideas so general that they not only solve the matching problem in polynomial time, but also show the polynomial solvability of the majority of other combinatorial optimization problems for which polynomial algorithms are known! Such a general approach is important in classifying combinatorial problems according to their complexity, and thereby singling out those which should be targets of successful attack by special combinatorial methods.

For some large-scale applications of matching theory, even the $O(p^3)$ or $O(p^{2.5})$ implementations of the matching algorithm are too time-consuming. Hence simpler heuristic algorithms giving "almost optimal" matchings have also been developed. Such algorithms have at least one advantage, even from a theoretical point of view. The expected weight of their output can be determined in a situation where the weights of the lines are random variables. In this way bounds on the expected weight of optimum matchings can be obtained. (Edmonds' Algorithm is too involved for such a probabilistic analysis.) We shall not discuss matching heuristics in this book, however. The interested reader is directed to the survey paper of Avis (1983).

9.1. The Edmonds Matching Algorithm

We begin with a lemma which will enable us to reduce the size of the graph under consideration in many cases. Although this lemma will not be needed in the final analysis of the algorithm, it does help us understand the crucial step of "cycle shrinking" and lends us confidence that we are not losing necessary information when carrying out such shrinking.

9.1.1. LEMMA. (*The Cycle Shrinking Lemma*). *Let G be a graph, M, a matching in G and let Z be a cycle of length $2k+1$ which contains k lines of M and is point-disjoint from the rest of M. Construct a new graph G' from G by shrinking Z to a single point. Then $M' = M - E(Z)$ is a maximum matching in G' if and only if M is a maximum matching in G.*

PROOF. Assume first that M is not a maximum matching in G. Then there exists an augmenting path P relative to M, by Berge's Theorem 1.2.1. If P is disjoint from Z, then P is also an augmenting path in G' relative to M', and so M' is not maximum. So suppose that P intersects Z. At least one of the endpoints of P — say x — is not on Z. Starting at x, let z be the first point encountered on Z by traversing P. Then $P[x, z]$

is mapped onto an augmenting path relative to M' if Z is contracted. Thus M' is not maximum.

Assume now that M' is not a maximum matching in G' and let N' be a matching in G' of greater cardinality than M'. Let us again "blow up" Z to recover G. Then N' will correspond to a matching in G which covers at most one point in Z. But then N' can be augmented, using k lines of Z, to yield a matching N of size $|N| = |N'| + k > |M'| + k = |M|$. So M is not a maximum matching in G. ∎

Notice (using the same notation as in the above proof) that if we can find a matching in G' larger than M', then not only can we conclude that M is not a maximum matching in G, but we can also easily find a larger one!

We now turn to an informal description of the Edmonds Matching Algorithm. Suppose that we are given a graph G in which we have found a matching M. If M is perfect we have nothing to do, so suppose that the set S of points not covered by M is non-empty. Construct a forest F such that every connected component of F contains exactly one point of S, every point of S belongs to exactly one component of F, and every line of F which is at an odd distance from a point in S belongs to M. It then follows that every point of F which is at an odd distance from S has degree 2 in F. Such points will be called **inner** points, while the remaining points in F will be called **outer**. (In particular, all points of S are outer.) Such a forest is called an **M-alternating forest**. Clearly, the forest with point set S and no lines is an M-alternating forest.

Next, we consider the neighbors of outer points. If we find an outer point x adjacent to a point y not in F, then consider the line $yz \in M$ and let $F' = F + xy + yz$. This subgraph is clearly an M-alternating forest which is larger than F.

If F has two outer points x, y, which belong to different components of F and which are adjacent in G, then the roots of these two components are connected by an M-augmenting path consisting of xy and the unique paths from x and y to the root of each tree. So by alternating on this path we can obtain a matching larger than M. Now begin again using the larger matching.

If F has two outer points x, y, in the same connected component which are adjacent in G, then let C be the cycle formed by line xy and by the xy-path in F. Let P denote the (unique) path in F connecting C to a root of F. (We allow C to pass through a root, in which case path P consists of a single point.) Clearly P is an M-alternating path, so if we switch on P, we obtain another matching M_1 of the same size as M.

But M_1 and C now satisfy the conditions of Lemma 9.1.1, and so if we shrink C to a single point to obtain a new graph G', we have reduced the task of finding a matching larger than M in G to the task of finding a matching larger than $M_1 - E(C)$ in the smaller graph G'.

Finally, if every outer point has only inner points as neighbors, then we claim that the matching M is already maximum. For suppose that F contains m inner points and n outer points. Clearly $n - m = |S|$. Furthermore, if we delete all the inner points of F from G, the remaining graph will contain all the outer points of F as isolated points. Hence $\operatorname{def}(G) \geq n - m = |S|$. But M misses exactly $|S|$ points, and so it must be a maximum matching.

In summary then, we can always do one of the following: enlarge F, enlarge M, decrease $|V(G)|$ or stop with a maximum matching. Thus it is clear that this algorithm terminates in polynomial time with a maximum matching in G. However, it is worth-while to undertake a direct analysis of the execution of the algorithm in which we do not appeal to Lemma 9.1.1. Such an analysis will have two valuable consequences. First, we shall be able to derive better bounds for the running time and second, it will turn out that the algorithm yields new proofs of many fundamental results in matching theory. (See the next four exercises.)

9.1.2. EXERCISE. Prove that the set A of inner points at termination is the same set as the set $A(G)$ in the Gallai-Edmonds Structure Theorem 3.2.1.

9.1.3. EXERCISE. Let G be a bipartite graph with positive surplus. By applying the Edmonds Matching Algorithm to G, obtain a new proof of Theorem 1.3.8.

9.1.4. EXERCISE. Give an alternate proof of the Gallai-Edmonds Theorem using the Edmonds Matching Algorithm.

9.1.5. EXERCISE. Give a new proof of the ear structure theorem for factor-critical graphs (Theorem 5.5.1) based upon the Edmonds Matching Algorithm.

To facilitate our analysis, let us write down the algorithm a bit more formally.

(A) THE DATA STRUCTURE

The main difficulty in the Edmonds Algorithm lies in the implementation of shrinking. It does not suffice simply to construct the graph obtained by shrinking an appropriate odd cycle. We must later recover the original graph in order to construct an augmented matching of the original graph from the augmented matching in the shrunken graph. Repeated odd cycle shrinking will map factor-critical graphs onto points; we call such factor-critical graphs "blossoms". When we make an augmentation step, we must have enough information at hand about the structure of the blossoms to be able to easily find a near-perfect matching in each of them which misses any given point.

One way to achieve this is to store an ear decomposition of the blossoms. So let us discuss how ear decompositions of factor-critical graphs can be stored in an economical way. Suppose that $G = P_0 + P_1 + \cdots + P_k$ is an ear decomposition of a factor-critical graph G where $P_0 = \{r\}$ is a singleton and P_{i+1} is an odd ear attached to $P_0 + P_1 + \cdots + P_i$ for each $i = 0, \ldots, k-1$. Let M be the near-perfect matching of G which consists of every second line of each P_i (and so M fails to cover r). Then for every point $x \neq r$, there is a unique first ear P_t containing x. This P_t has two lines incident with x and precisely one of these belongs to M. Let $\mu(x)$ denote the other endpoint of the line in M incident with x and let $\phi(x)$ denote the other neighbor of x on P_i. Also let $\phi(r) = r$.

9.1.6. EXERCISE. Show that the mappings μ and ϕ determine the ear decomposition uniquely.

Construct the sequences

$$x^0 = x,\ x^1 = \mu(x),\ x^2 = \phi\mu(x),\ x^3 = \mu\phi\mu(x), \ldots$$

and

$$x^{-1} = \phi(x),\ x^{-2} = \mu\phi(x),\ x^{-3} = \phi\mu\phi(x), \ldots.$$

Note that the sequence $x^0, x^1, x^2, \ldots$ never cycles. It proceeds along the first ear containing x, P_t, then it continues along a P_s where $s < t$, and so on. Hence it terminates with r. This yields an M-alternating path from x to r and if we switch along this path, we obtain a maximum matching of G missing x. (See Figure 9.1.1.)

9.1.7. EXERCISE. Let G be a graph, let μ, $\phi : V(G) \to V(G)$ be two mappings and suppose $r \in V(G)$. Assume that the following conditions hold:

(E–1) $\phi(r) = \mu(r) = r$ and ϕ and μ have no other fixed points;
(E–2) μ^2 is the identity mapping;
(E–3) for each $x \in V(G)$, $(\mu\phi)^n(x) = r$ for sufficiently large n.

Prove that μ and ϕ are mappings associated with an ear decomposition of G described above.

Returning to the Edmonds Algorithm now, we shall not explicitly carry out the shrinking, but only indicate the points to be identified. Thus we may work in the same graph G throughout. The price we have to pay is that instead of alternating forests, we shall have to work with more complicated subgraphs which we call "blossom forests". Given a matching M, a **blossom forest with respect to** M is a subgraph B with the

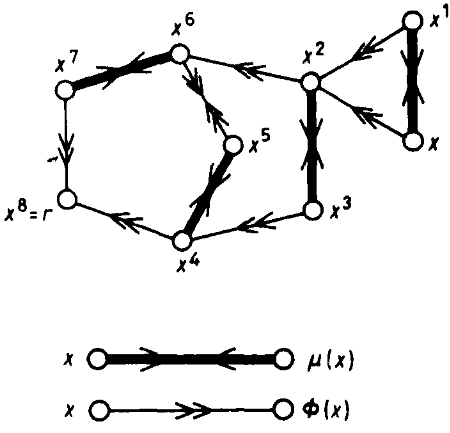


FIGURE 9.1.1. How to store an ear decomposition

following properties. It contains point-disjoint factor-critical subgraphs called **blossoms** such that M contains a near-perfect matching of every blossom and, upon shrinking every blossom to a single point, we obtain an alternating forest with respect to the image of M. Moreover, the images of blossoms are outer points of this alternating forest. (Note that if all blossoms are singletons, that is, if our blossom forest is an alternating forest, then blossom points are just outer points.)

It follows that every blossom has a unique point which is either unmatched by M or matched to a point outside the blossom. This point will be called the **base** of the blossom. Since inner points of the alternating tree obtained by blossom shrinking correspond to single points in the blossom forest, we retain the name **inner** for them. Note that the inner points all have degree two. Points not in the forest will be called **out-of-forest.** Note that all out-of-forest points are paired up by the matching M. (See Figure 9.1.2.)

The matching M will be described by a function $\mu : V(G) \to V(G)$ such that $\mu(x)$ is the other endpoint of the line of M incident with x if x is matched and $\mu(x) = x$ if x is unmatched.

For each blossom of the blossom forest, we shall store an ear decomposition starting with its base, using the function ϕ defined above. We let $\phi(x) = x$, if x is a base point of a blossom or an out-of-forest point. For inner points x, however, we may use the same pointer $\phi(x)$ to denote the other endpoint of the unique line of $E(B) - M$ incident with x.

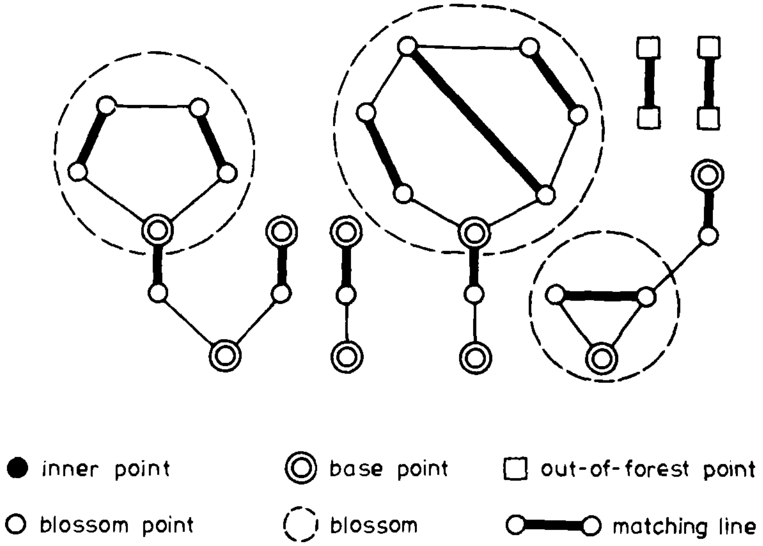


FIGURE 9.1.2. A blossom forest

Instead of explicitly shrinking each blossom to a single point, we shall direct a pointer $\rho(x)$ from each point of a blossom to the base of that blossom. (We may interpret $\rho(x)$ informally as "point x is currently shrunk to point $\rho(x)$.") We let $\rho(x) = x$ for all inner and out-of-forest points.

We need one more item in our data structure. Since we are going to scan blossom points for lines joining them to out-of-forest points or other blossoms, we shall need to know which points have been "scanned" and which points are still "unscanned". This will be indicated by a label $\sigma(x) \in \{\text{SCANNED}, \text{UNSCANNED}\}$.

These considerations motivate the *data structure* which we set up as follows. Our graph will be given by listing the set of neighbors $\Gamma(x)$ for each point x. During the execution of the algorithm, other information

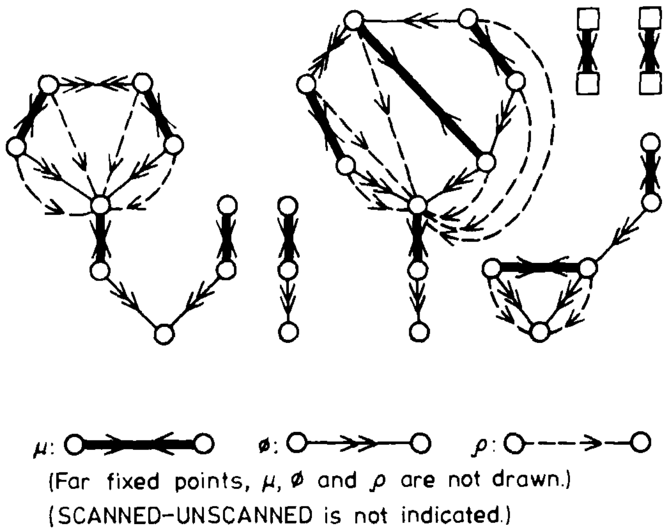


FIGURE 9.1.3. Data structure for the algorithm

concerning points has to be stored, so for each point x we shall maintain a list

$$(\Gamma(x), \mu(x), \phi(x), \rho(x), \sigma(x)),$$

where

$$\Gamma(x) \subseteq V(G) - \{x\}, \mu(x), \phi(x), \rho(x) \in V(G),$$

and

$$\sigma(x) \in \{\text{SCANNED}, \text{UNSCANNED}\}.$$

REMARK. The pointers μ, ϕ and ρ defined above will satisfy the following conditions throughout the execution of the algorithm.

(P–1) If $\mu(x) \neq x$, then $x\mu(x) \in E(G)$ and $\mu(\mu(x)) = x$. (Hence the lines $x\mu(x)$ form a matching M.)

(P–2) If $\phi(x) \neq x$, then $x\phi(x) \in E(G)$.

(P–3) The lines of type $x\mu(x)$ and of type $x\phi(x)$ form a blossom forest with respect to matching M.

(P–4) When restricted to the points of a blossom, μ and ϕ define an ear decomposition of the blossom starting with its base point. $\phi(x) = x$ for base points and out-of-forest points.

(P–5) $\rho(x)$ is the base point of the blossom containing x if x is a blossom point and $\rho(x) = x$, otherwise.

vskip 4 pt

It is easy to read off the type of point x is from its list. In fact, x is

(i) an unmatched point if and only if $\mu(x) = x$ (and then also $\phi(x) = \rho(x) = x$),
(ii) a blossom point if and only if it is either unmatched or $\phi\mu(x) \neq \mu(x)$,
(iii) an inner point if and only if $\phi\mu(x) = \mu(x)$, but $\phi(x) \neq x$,
(iv) a base point if and only if it is a blossom point with $\phi(x) = x$, and
(v) an out-of-forest point if and ony if $\mu(x) \neq x$, but $\phi\mu(x) = \mu(x)$ and $\phi(x) = x$.

Now for each point x, set

$$x^0 = x, x^1 = \mu(x), x^2 = \phi\mu(x), x^3 = \mu\phi\mu(x), \ldots$$

and set $S(x) = (x^0, x^1, x^2, \ldots)$. It is immediate from (P–1) – (P–5) that for every blossom point x, the sequence $S(x)$ forms an M-alternating path from x to some unmatched point and then repeats this point. We shall delete this constant part of the sequence, so that from now on, for every blossom point x the sequence $S(x)$ will be finite without repetition and will end with an unmatched point.

It is also clear that $S(x)$ leaves any blossom through its base point and enters any blossom from an inner point. Once it leaves a blossom it never returns.

It will also prove useful to discuss the situation in which two blossom points x and y have $S(x) \cap S(y) \neq \emptyset$. Clearly they must then belong to the same connected component of the blossom forest and $S(x)$ and $S(y)$ terminate with the same unmatched point. Let u be the first base point in $S(x)$ which also belongs to $S(y)$ and let x^i and y^j be the first points of $S(x)$ and $S(y)$ in the blossom containing u. (Obviously, i and j are even.) Set

$$A(x,y) = \{x^0, \ldots, x^{i-1}, y^0, \ldots, y^{j-1}\}.$$

Observe that $A(x,y)$ is easily computable. It is trivial to compute $S(x)$ and $S(y)$. Furthermore, x^i is the first point x^k in $S(x)$ with $\rho(x^k) \in S(y)$ and y^j is the first point y^l in $S(y)$ with $\rho(y^l) \in S(x)$.

We are now prepared to describe the algorithm.

(B) THE ALGORITHM

We begin with $\mu = \phi = \rho =$ identity and $\sigma =$ UNSCANNED. So our initial blossom forest consists of $V(G)$ and no lines.

1^0. If every blossom is SCANNED, stop. The set of lines $\{x\mu(x)\}$ is a maximum matching in G. If there is at least one UNSCANNED blossom point, select one of them, say x.

2^0. Consider the points in $\Gamma(x)$.

Case 1. Every neighbor y of x is either an inner point or a blossom point with $\rho(y) = \rho(x)$. Then relabel x as SCANNED and return to 1^0.

Case 2. Point x has a neighbor y which is a blossom point with $\rho(y) \neq \rho(x)$. Then let us form the sets $S(x)$ and $S(y)$ and distinguish two cases:

Case 2a. If $S(x)$ and $S(y)$ are disjoint then we can *augment*. This is done by setting:

$\mu(z) = x^{2i}$, if $z = x^{2i-1}$,
$\mu(z) = x^{2i-1}$, if $z = x^{2i}$,
$\mu(z) = y^{2j}$, if $z = y^{2j-1}$,
$\mu(z) = y^{2j-1}$, if $z = y^{2j}$,
$\mu(x) = y$, $\mu(y) = x$, leaving
$\mu(z)$ unchanged otherwise, and setting
$\phi = \rho =$ identity, $\sigma =$ UNSCANNED.

Return to 1^0.

Case 2b. $S(x)$ and $S(y)$ are not disjoint. Then we can *shrink* a blossom as follows. Let u be the first base point on $S(x)$ which belongs to $S(y)$. Set

$\phi(z) = x^{2i-1}$, if $z = x^{2i} \in A(x,y)$, $i \geq 1$,
$\phi(z) = y^{2j-1}$, if $z = y^{2j} \in A(x,y)$, $j \geq 1$,
$\phi(x) = y$ and $\phi(y) = x$, leaving
$\phi(z)$ unchanged otherwise,
$\rho(z) = u$, if $\rho(z) \in A(x,y)$ and leaving
$\rho(z)$ unchanged otherwise.
Leave μ and σ unchanged.

Return to 2^0.

Case 3. Point x has a neighbor y which is out-of-forest. Then the blossom tree will *grow* if we set

$\phi(y) = x$,
leave $\phi(z)$ unchanged otherwise, and
leave μ, ρ and σ unchanged.

Return to 2^0.

End.

9.1.8. THEOREM. *The Matching Algorithm presented in (B) above terminates in $O(p^3)$ time When it terminates, the lines of the form $x\mu(x)$ form a maximum matching.*

PROOF. Let us call the execution of one augmentation, one shrinking or one expansion (that is, given x and y, one manipulation of the data structure) an iteration.

First of all, we have to show that the hypotheses (P–1) through (P–5) which we made about the data structure remain valid throughout the execution of the algorithm. This follows easily from the analysis in (A).

Next we show that the algorithm must terminate in $O(p^3)$ time. To this end, we estimate separately the time needed to carry out the iterations (that is, to manipulate the data) and the time needed to search for a pair (x, y) of points satisfying the conditions of Cases 2 or 3 in the algorithm.

First, notice that the execution of any iteration takes no more than $O(p)$ time. The number of augmentations is $O(p)$, since the number of unmatched points drops at each augmentation. The number of other iterations (expansions and shrinkings) between any two augmentations is $O(p)$, since the number of fixed points of ϕ drops at each such interation. Hence there are only $O(p^2)$ iterations and the total time spent manipulating data is $O(p^3)$.

Second, notice that each point x is scanned only once between any two augmentations. To scan a point takes $O(p)$ time. Thus the total time spent searching for x and y is $O(p^3)$. This proves that the algorithm does indeed terminate in $O(p^3)$ time.

Consider now the situation at termination. We want to show that the lines $x\mu(x)$ form a maximum matching. We need the following two observations.

Claim 1. When the algorithm terminates, every pair $\{x, y\}$ of adjacent blossom points belong to the same blossom; that is, they satisfy $\rho(x) = \rho(y)$.

In fact, when the algorithm terminates, both x and y must be labelled SCANNED. Consider the step when $\sigma(x)$ or $\sigma(y)$ changed for the final time. Then $\sigma(x)$ must have been labelled SCANNED and $\sigma(y)$ must have changed from UNSCANNED to SCANNED (say). Before this step, both x and y had been blossom points with $\rho(x) = \rho(y)$ (by the rule for changing a label to SCANNED). Since no augmentation step can follow this step, we must have $\rho(x) = \rho(y)$ throughout the rest of the algorithm.

Claim 2. When the algorithm terminates, no blossom point is adjacent to any out-of-forest point.

The argument here is quite similar to that for Claim 1.

The proof of Theorem 9.1.8 is now easy. Let M be the matching defined by μ when the algorithm terminates. Let X denote the set of inner points and let Y be the set of base points. Then every unmatched point belongs to Y and the matched points in Y are, in fact, matched with the points in X by M. Hence

$$|Y| = |X| + p - 2|M| \geq X + \operatorname{def}(G).$$

On the other hand, the blossoms in the final blossom tree are odd connected components of $G - X$, by Claims 1 and 2. Hence

$$|Y| \leq c_0(G - X) \leq |X| + \operatorname{def}(G),$$

by the "trivial" half of the Berge Formula. Hence equality must hold throughout and, in particular, M must be a maximum matching. ■

REMARK 1. Note that the data structure is redundant in the sense that the pointer $\rho(x)$ is easily computable from the pointers $\mu(x)$ and $\phi(x)$. In fact, $\rho(x)$ is the last point x^k in $S(x)$ such that $x^0, \ldots, x^k$ are all blossom points. Recall that the main use of $\rho(x)$ has been to tell whether or not two blossom points x and y belong to the same blossom. However, recomputing $\rho(x)$ every time we need it would add a factor of p to the running time. It is more economical to store it along with the other information.

Reference	Total Time Bound
1. Edmonds (1965a)*	p^4
2. Witzgall and Zahn (1965)*	p^3
3. Balinski (1969)*	p^3
4. Gabow (1973, 1976), Lawler (1976)	p^3
5. Kameda and Munro (1974)	pq
6. Even and Kariv (1975), Kariv (1976)	$p^{5/2}$
7. Bartnik (1978)	$p^{5/2}$
8. Micali and Vazirani (1980)	$qp^{1/2}$

*(Bound stated for algorithm not established in reference cited.)

TABLE 9.1.1. Cardinality Matching Algorithms

REMARK 2. It is possible to streamline this algorithm. The most obvious way to save time is by modifying the augmentation step. As we have presented the algorithm, after each augmentation it starts again "from scratch". Why not save all the components of the blossom forest built so far (except those two which contain x and y)? Of course this savings is significant only if these two trees are not too large. If they contain the majority of points then the savings is not significant. This savings can be achieved by special rules for selecting the next blossom point to be scanned. The above idea leads to an improved running time of $O(p^{5/2})$. (For a brief history of cardinality matching algorithms, see Table 9.1.1. For much additional information on implementation of these algorithms, see Bartnik (1978), Gondran and Minoux (1979 or 1984, Chapter 7), Burkard and Derigs (1980), Galil (1983) or Vazirani (1984).)

9.2. Weighted Matching

In Chapter 7 we treated a generalization of the matching problem by considering weights on the lines of the graph and by asking for a matching of maximum weight. We obtained a good characterization of this maximum by describing the facets of the matching polytope and then applying linear programming duality.

We now describe a polynomial time algorithm to find a matching with maximum weight. We might expect that since a good characterization was obtained from the duality theorem, this algorithm would arise from the application of a linear programming algorithm to the matching polytope. The situation, however, is not this simple because the description of the matching polytope as the solution set of a system of linear inequalities involves exponentially many constraints.

In Section 9.4 we shall discuss how a general method to solve linear programs, the so-called Ellipsoid Method, can also be used to obtain a polynomial-time weighted matching algorithm. This is far from being practical, however, and therefore we first describe a more efficient and more combinatorial weighted matching algorithm, which is essentially due to Edmonds (1965b). This algorithm will also make use of the linear programming dual of the maximum weight matching problem viewed as a linear program, but we shall not need Edmonds' description of the matching polytope. In fact, an analysis of this algorithm will yield a new proof of the matching polytope theorem. On the other hand, however, we shall need an algorithm for the maximum cardinality matching problem as a subroutine.

We start with some obvious reductions and transformations of the problem. Let G be a graph and $w : E(G) \to \mathbb{Q}$ a weighting of its lines. First, we may assume that $w \geq 0$ since the lines with negative weight do not occur in any maximum weight matching. Second, we may assume that $|V(G)|$ is even, since if not, we may add an isolated point. Third, we may assume that G is a complete graph, since we may join all pairs of non-adjacent points by lines having weight zero. But then every matching can be extended to a perfect matching without decreasing its weight, and so it suffices to look for a *maximum weight perfect matching* in a complete graph with weights on its lines.

Finally, by replacing each weight by its negative, it suffices to solve the problem of finding a *minimum weight perfect matching* in a complete graph with weights on its lines.

The algorithm which we describe below yields the following result.

9.2.1. THEOREM. *The minimum weight perfect matching problem can be solved in polynomial time.*

PROOF. By Theorem 7.3.4, this problem can be formulated as a linear program:

$$\begin{array}{lll} \text{minimize} & \boldsymbol{w}^{\mathrm{T}} \cdot \boldsymbol{x} & \\ \text{subject to} & x_e \geq 0 & \text{(for each } e \in E(G)) \\ & \boldsymbol{x}(C) = 1 & \text{(for each trivial odd cut } C) \\ & \boldsymbol{x}(C) \geq 1 & \text{(for each non-trivial odd cut } C). \end{array}$$

Now consider the dual program. We have a variable y_C for each odd cut C and the following objective and constraints:

$$\begin{array}{lll} \text{maximize} & \sum_C y_C & \\ \text{subject to} & y_C \geq 0 & \text{(for each non-trivial odd cut } C) \\ & \sum_{\substack{C \\ e \in C}} y_C \leq w(e) & \text{(for every } e \in E(G))). \end{array}$$

Any vector y satisfying the constraints in this dual program will be called, briefly, a *dual solution.* We may view such a dual solution as a packing of odd cuts, where each line e has capacity $w(e)$ and each odd cut C occurs y_C times. Note, however, that y_C may be fractional and we may even have $y_C < 0$ for the trivial odd cuts.

For any dual solution $\boldsymbol{y}$, define

$$E_{\boldsymbol{y}} = \{e \in E(G) \mid \sum_{\substack{C \\ e \in C}} y_C = w(e)\}.$$

So $E_{\boldsymbol{y}}$ is the set of lines whose capacity is already "used up". Now set $G_{\boldsymbol{y}} = (V(G), E_{\boldsymbol{y}})$.

Our strategy will be to find an optimum dual solution first, and then use this to obtain a minimum weight perfect matching.

To motivate some of the considerations to follow, let us consider the so-called "complementary slackness" conditions. In our case, these can be derived as follows. Let M be any perfect matching and $\boldsymbol{y}$ any dual solution. Then

$$\begin{aligned} w(M) &= \sum_{e \in M} w(e) \geq \sum_{e \in M} \sum_{\substack{C \\ e \in C}} y_C \\ &= \sum_C |M \cap C| y_C \geq \sum_C y_C, \end{aligned}$$

since obviously $|M \cap C| \geq 1$ for each odd cut and $|M \cap C| = 1$ for each trivial odd cut. If M is a minimum weight perfect matching and $\boldsymbol{y}$ is an optimum dual solution then we have equality here. So in particular,

$$w(e) = \sum_{\substack{C \\ e \in C}} y_C$$

holds for each $e \in M$, that is, M is a perfect matching in $G_{\boldsymbol{y}}$. Moreover, we must have $|M \cap C| = 1$ whenever C is an odd cut with $y_C > 0$. Conversely, if M is a perfect matching in $G_{\boldsymbol{y}}$ and $|M \cap C| = 1$ for each odd cut C with $y_C > 0$, then $w(M) = \sum_C y_C$. Hence M must be a minimum weight perfect matching and $\boldsymbol{y}$ must be an optimum dual solution.

We shall restrict ourselves to dual solutions of a rather special form. Recalling the "uncrossing" technique from several proofs in Chapters 6 and 7, it is quite natural to consider only those dual solutions $\boldsymbol{y}$ for which $y_{C_1} > 0$ and $y_{C_2} > 0$ imply that C_1 and C_2 are non-crossing. In fact we shall need slightly more, which can be formulated as follows. We assume that there exists a family $\mathcal{H}$ of odd subsets of $V(G)$ such that

(P–1) $\mathcal{H}$ is nested; i.e., if $S, T \in \mathcal{H}$, then either $S \subseteq T$ or $T \subseteq S$ or $S \cap T = \emptyset$;

(P–2) $\mathcal{H}$ contains all singletons from $V(G)$;

(P–3) if C is a non-trivial odd cut, then $y_C > 0$ if and only if $C = \nabla(S)$ for some $S \in \mathcal{H}$;

(P–4) if $S \in \mathcal{H}$, and if $U_1, \ldots, U_h$ are the maximal members of $\mathcal{H}$ properly contained in S, then contracting each U_i to a single point in $G_{\boldsymbol{y}}[S]$, we obtain a factor-critical graph.

Note that the existence of a family $\mathcal{H}$ with properties (P–1) – (P–3) is equivalent to the property of $\boldsymbol{y}$ that cuts C with $y_C > 0$ are non-crossing. Property (P–1) implies that $|\mathcal{H}| \le 3p/2$ (we leave the proof of this to the reader). So in a dual solution with such a structure, at most $3p/2$ dual variables are non-zero. It would not be too difficult to see that among the optimal dual solutions, there is always one with this structure. But we shall not need this fact, which follows, incidently, as a by-product of the algorithm.

The algorithm, as we describe it, has two main phases.

Phase 1. Set $y_C = 0$ for all C and let $\mathcal{H}$ be the collection $\{\{v\} \mid v \in V(G)\}$. Suppose that we have some dual solution $\boldsymbol{y}$ and a family of odd sets $\mathcal{H}$ satisfying conditions (P–1) – (P–4). Let $S_1, \ldots, S_k$ be the (inclusionwise) maximal members of $\mathcal{H}$. It follows from (P–1) that $S_1, \ldots, S_k$ are mutually disjoint and by (P–2) they partition $V(G)$. Let $G'_{\boldsymbol{y}}$ denote the graph obtained from $G_{\boldsymbol{y}}$ by contracting each S_i to a single point s_i. Since $|V(G)|$ is even, but $|S_j|$ is odd, it follows that $k = |V(G'_{\boldsymbol{y}})|$ is even.

Case 1. Suppose $G'_{\boldsymbol{y}}$ has no perfect matching. Let us find the sets of the Gallai-Edmonds decomposition for $G'_{\boldsymbol{y}}$: $A(G'_{\boldsymbol{y}})$, $C(G'_{\boldsymbol{y}})$ and $D(G'_{\boldsymbol{y}})$. Let, say, $A(G'_{\boldsymbol{y}}) = \{s_1, \ldots, s_m\}$ and let $H_1, \ldots, H_{m+d}$ be the components of $G'_{\boldsymbol{y}}[D(G'_{\boldsymbol{y}})]$, where $d = \operatorname{def}(G'_{\boldsymbol{y}}) > 0$. Also let $T_i = \bigcup_{s_j \in V(H_i)} S_j$. Obviously, $|T_i|$ is odd.

Now modify the dual solution $\boldsymbol{y}$ as follows. Let

$$\begin{aligned} y^t_{\nabla(S_j)} &= y_{\nabla(S_j)} - t \quad (1 \le j \le m), \\ y^t_{\nabla(T_i)} &= y_{\nabla(T_i)} + t \quad (1 \le i \le m+d), \\ y^t_C &= y_C, \qquad \text{otherwise.} \end{aligned}$$

For a small — but positive — t, the vector $\boldsymbol{y}^t$ is also a dual solution. In fact, $y^t_C \ge y_C \ge 0$ for every non-trivial odd cut C, except perhaps if $C = \nabla(S_j)$ for some $1 \le j \le m$. But then $S_j \in \mathcal{H}$ and hence either $|S_j| = 1$ or else $y_{\nabla(S_j)} > 0$ and so $y^t_{\nabla(S_j)} \ge 0$ for a small t. Furthermore, we claim that

$$\sum_{\substack{C \\ e \in C}} y^t_C \le w(e)$$

also holds for each $e \in E(G)$. This is clear if $e \notin E_{\boldsymbol{y}}$, so consider any $e \in E_{\boldsymbol{y}}$. If $e \notin \nabla(T_1) \cup \cdots \cup \nabla(T_{m+d})$ then

$$\sum_{\substack{C \\ e \in C}} y^t_C \le \sum_{\substack{C \\ e \in C}} y_C = w(e),$$

so suppose that $e \in \nabla(T_i)$ for some $1 \le i \le m+d$. Then e corresponds to a line e' of $G'_{\boldsymbol{y}}$ such that e' connects H_i to a point outside H_i. By the definition of H_i and by the basic properties of the Gallai-Edmonds decomposition, e' must connect H_i to one of $s_1, \ldots, s_m$. So $e \in \nabla(S_j)$ for some $1 \le j \le m$, and hence

$$\begin{aligned} \sum_{\substack{C \\ e \in C}} y^t_C &= (y_{\nabla(T_i)} + t) + (y_{\nabla(S_j)} - t) + \sum_{\substack{C \\ e \in C \\ C \ne \nabla(T_i), \nabla(S_j)}} y_C \\ &= \sum_{\substack{C \\ e \in C}} y_C = w(e). \end{aligned}$$

This shows that $\boldsymbol{y}^t$ is indeed a dual solution. If we set

$$\mathcal{H}^t = \{T_1, \ldots, T_{m+d}\} \cup \{S \in \mathcal{H} \mid |S| = 1 \text{ or } y^t_{\nabla(S)} > 0\},$$

then it is straightforward to check that $\mathcal{H}^t$ satisfies conditions (P–1) – (P–4) with respect to the dual solution $\boldsymbol{y}^t$. So $\boldsymbol{y}^t$ is a dual solution of the required form.

Also note that

$$\sum_C y_C^t = \sum_C y_C + td > \sum_C y_C,$$

so $\boldsymbol{y}^t$ is "better" than $\boldsymbol{y}$.

Choose t as large as possible so that $\boldsymbol{y}^t$ is still a dual solution. From the proof of the fact that this holds for small values of t, we see that the value of this largest t is

$$t = \min\{t_1, t_2, t_3\}, \tag{9.2.1}$$

where

$$t_1 = \min\{y_{\nabla(S_j)} \mid 1 \le j \le m,\ |S_j| > 1\},$$

$$t_2 = \min\{w(e) - \sum_{\substack{C \\ e \in C}} y_C \mid e \in \nabla(T_1) \cup \cdots \cup \nabla(T_{m+d}) - \nabla(S_1) - \cdots - \nabla(S_m)\},$$

$$t_3 = \tfrac{1}{2}\min\{w(e) - \sum_{\substack{C \\ e \in C}} y_C \mid e \in \nabla(T_i) \cap \nabla(T_j),\ 1 \le i < j \le m+d\}.$$

Now replace $\boldsymbol{y}$ by $\boldsymbol{y}^t$ and repeat.

Case 2. Suppose G' has a perfect matching. Then Phase 1 of the algorithm is completed.

Phase 2. Let M_0 be the matching in $G_{\boldsymbol{y}}$ corresponding to a perfect matching M' in $G'_{\boldsymbol{y}}$. We extend M_0 to a perfect matching M of $G_{\boldsymbol{y}}$ as follows. Consider any S_i, say S_1, such that $|S_i| > 1$. (If $|S_i| = 1$, for each i, then $M_0 = M'$ is already a perfect matching.) Let $U_1, \ldots, U_r$ be the maximal members of $\mathcal{H}$ properly contained in S_1. Obviously S_1 has exactly one point matched by M_0. Let v_1 be this point and assume without loss of generality that $v_1 \in U_1$. Contract each U_i to a single point u_i and delete all points of $V(G) - S_1$ from $G_{\boldsymbol{y}}$ to obtain a graph G''. By hypothesis (P–4), G'' is factor-critical, so $G'' - u_1$ has a perfect matching M'_1, which corresponds to a matching M_1 in $G_{\boldsymbol{y}}$. Then $M_0 \cup M_1$ is a matching in $G_{\boldsymbol{y}}$ which matches exactly one point of each of the sets $U_1, \ldots, U_r, S_2, \ldots, S_k$. Continuing in this way, we obtain a perfect matching M of G.

By the construction above, this perfect matching M satisfies the two conditions: $M \subseteq E_{\boldsymbol{y}}$ and $|M \cap C| = 1$ for each C with $y_C > 0$. Hence by the remark made at the beginning of this proof, M is necessarily a minimum weight perfect matching and we are done.

It remains to estimate the running time of this algorithm. In particular, this will show that it terminates in finite time. (This is not entirely obvious from the preceding discussion!) We shall estimate the running times of Phase 1 and Phase 2 separately.

The first of these two tasks is the more difficult. Let us investigate how the graph $G'_{\boldsymbol{y}}$ changes when $\boldsymbol{y}$ is replaced by the "better" dual solution $\boldsymbol{z} = \boldsymbol{y}^t$. Let us start with $t = 0$ (when $\boldsymbol{z} = \boldsymbol{y}$), and let us increase t to its largest possible value. Obviously, as soon as t becomes positive, all lines connecting a point of $A(G'_{\boldsymbol{y}})$ to a point of $A(G'_{\boldsymbol{y}}) \cup C(G'_{\boldsymbol{y}})$ drop out of G', and each connected component H_i spanned by $D(G'_{\boldsymbol{y}})$ is contracted to a single point h_i. When t reaches its maximum value as given by equation (9.2.1), one or more of the following things happen:

(1) a set S_j, $1 \le j \le m$, drops out of $\mathcal{H}$. Accordingly, the point s_j of $G'_{\boldsymbol{y}}$ is "inflated" to become a factor-critical subgraph (by property (P–4) of $\mathcal{H}$),

(2) a line connecting some point h_i to $C(G'_{\boldsymbol{y}})$ is added, or

(3) a line connecting two points h_i and h_j is added.

From this we see that $\operatorname{def}(G'_{\boldsymbol{z}}) \le \operatorname{def}(G'_{\boldsymbol{y}})$. Moreover, if $\operatorname{def}(G'_{\boldsymbol{z}}) = \operatorname{def}(G'_{\boldsymbol{y}})$, then the number of those points of G which correspond to points in $D(G'_{\boldsymbol{z}})$ is larger than the number of those points which correspond to points in $D(G'_{\boldsymbol{y}})$. Hence after at most p changes of the dual solution, $\operatorname{def}(G'_{\boldsymbol{y}})$ must decrease, and so the total number of times the dual solution is changed is at most p^2. To change a dual solution we must solve an unweighted maximum matching problem, which can be done by the algorithm of Section 9.1 in $O(p^3)$ time. So the total time used in Phase 1 is $O(p^5)$.

Phase 2 can be carried out by solving at most $|\mathcal{H}|$ unweighted matching problems and hence can be accomplished in $O(p^4)$ time. ■

9.2.2. EXERCISE. Show that with some care, Phase 1 of the weighted matching algorithm can be implemented in $O(p^4)$ time and Phase 2 can be implemented in $O(p^3)$ time.

9.2.3. EXERCISE. Prove that the dual solution $\boldsymbol{y}$ obtained in Phase 1 has values which are half-integral, provided the given weights $w(e)$ are integral.

9.2.4. EXERCISE. Notice that the theorem describing the perfect matching polytope (Theorem 7.3.4) was only used as motivation for the idea of the algorithm. Derive Theorem 7.3.4 from this algorithm.

9.2.5. EXERCISE. Formulate the algorithm for the special case of bipartite graphs, and show that it can be implemented in $O(p^3)$ time.

The original polynomial weighted matching algorithm of Edmonds (1965b) was of $O(p^4)$ complexity. Subsequently, Gabow (1973), Lawler (1976) and Cunningham and Marsh (1978) developed algorithms each of $O(p^3)$ complexity. The fastest weighted matching algorithms to date are $O(qp \log p)$ and are due to Galil, Micali and Gabow (1982) and to Ball and Derigs (1983).

9.3. An Algorithm Based Upon the Gallai-Edmonds Theorem

By far the largest number of known matching algorithms — certainly the most efficient ones — are based upon the idea of building alternating forests and upon Edmonds' idea of shrinking odd cycles. In the next two sections we describe two more matching algorithms based upon different ideas. They do not seem to be as efficient as the algorithms discussed in the preceding section, but they do have some interesting features of their own.

The Edmonds Matching Algorithm gives, as a by-product, the important Gallai-Edmonds Structure Theorem. The algorithm presented next turns this relationship around and shows that the Gallai-Edmonds result may serve as a starting point for a matching algorithm.

Let G be a graph in which we seek a maximum matching. Assume that we have already found a matching of size k and we wish to find a larger one, or to conclude that the matching at hand is maximum. Assume further that we have already found not only one k-element matching, but a non-empty sequence $\mathcal{L} = (M_1, \ldots, M_t)$ of such matchings. (It will turn out that the number t will never be larger than p.)

Associated with this sequence, we define the sets

$$D(\mathcal{L}) = \{x \in V(G) \mid x \notin V(M_i) \text{ for some } 1 \le i \le t\},$$

$$A(\mathcal{L}) = \{x \in V(G) - D(\mathcal{L}) \mid x \text{ is adjacent to some point in } D(\mathcal{L})\},$$

and

$$C(\mathcal{L}) = V(G) - D(\mathcal{L}) - A(\mathcal{L}).$$

Of course, if $k = \nu(G)$ and $\mathcal{L}$ consists of *all* maximum matchings of G, then $D(\mathcal{L}) = D(G), A(\mathcal{L}) = A(G)$ and $C(\mathcal{L}) = C(G)$.

The following simple lemma may be viewed as a converse (in a sense) to the Gallai-Edmonds Structure Theorem.

9.3.1. LEMMA. *Let $\mathcal{L}$ be a list of k-element matchings of a graph G and suppose $M \in \mathcal{L}$. Assume that M contains no line joining any point of $A(\mathcal{L})$ to any point of $A(\mathcal{L}) \cup C(\mathcal{L})$ and that M contains a near-perfect matching of every connected component of $G[D(\mathcal{L})]$. Then M is a maximum matching.*

PROOF. The hypothesis implies that every connected component of $D(\mathcal{L})$ is odd, and that each of these components contains exactly one point which is either unmatched or matched to a point in $A(\mathcal{L})$. It also follows that the lines of M between $A(\mathcal{L})$ and $D(\mathcal{L})$ cover $A(\mathcal{L})$. Hence $D(\mathcal{L})$ has exactly $|A(\mathcal{L})| + |V(G) - V(M)|$ connected components and so

$$\operatorname{def}(G) \geq c_0(G - A(\mathcal{L})) - |A(\mathcal{L})| = |V(G) - V(M)|.$$

Hence M is maximum as claimed. ■

Now we can describe the new algorithm. At the beginning, set $k = 0$ and $\mathcal{L} = \{\emptyset\}$. In a general step, select any $M_1 \in \mathcal{L}$ and check to see if M_1 satisfies the conditions in Lemma 9.3.1. If it does, then stop; M_1 is already a maximum matching and we have also found the sets $A(G)$, $C(G)$ and $D(G)$. (See Exercise 9.3.2.)

So suppose that M_1 violates the conditions of Lemma 9.3.1. We are going to show how to find a matching which either has cardinality greater than k or which misses a point not in $D(\mathcal{L})$. There are two cases to consider.

Case 1. Assume that M_1 contains a line xy with $x \in A(\mathcal{L})$ and $y \in A(\mathcal{L}) \cup C(\mathcal{L})$. By the definition of $A(\mathcal{L})$, point x is adjacent to some point z in $D(\mathcal{L})$. By the definition of $D(\mathcal{L})$, there is a matching $M_2 \in \mathcal{L}$ missing z. If M_1 itself misses z, then replacing the line xy by the line xz in M_1, we create a new matching of size k which misses y. So if we add this matching to the list $\mathcal{L}$, set $D(\mathcal{L})$ will increase in size. If M_1 does not miss z then consider the $M_1 - M_2$ alternating path P which starts at z. If this path P ends with a line of M_1 then "switching" on P, we obtain from M_2 a matching of size $k+1$. If P ends with a line of M_2 and does not go through the line xy, then switching on $P + zx + xy$, we obtain from M_1 a k-element matching which misses y and so we can increase $D(\mathcal{L})$ as before. Finally, if P passes through xy then switching on an appropriate portion of P, we obtain from M_1 a k-element matching, M'_1, which misses one of x or y. Again, if we add M'_1 to the list $\mathcal{L}$, $D(\mathcal{L})$ will increase.

Case 2. Assume next that $G[D(\mathcal{L})]$ has a connected component T such that $M_1 \cap E(T)$ is not a near-perfect matching of M_1. There are now some subcases to distinguish:

Case 2a. Suppose $M_1 \cap E(T)$ misses (at least) two points x, y of T, and x is missed by M_1 as well. (Point y may be missed by M_1 or matched by M_1 with a point in $A(\mathcal{L})$.) If x and y are adjacent, then add the line xy to M_1 and delete the line incident with y from M_1 (if M_1 contains such a line). In this way, we obtain either a $k+1$-element matching or a k-element matching missing a point in $A(\mathcal{L})$, and in both cases we are done.

So suppose that x and y are non-adjacent. Let P be a shortest path in T connecting x to y, and let z be the immediate successor to x on P. By the definition of $D(\mathcal{L})$, there exists a matching M_i in $\mathcal{L}$ which misses z. Let Q denote the alternating $M_1 - M_i$ path which starts at z. If Q does not end at x, then, depending on the parity of Q, either Q or $Q + xz$ is an augmenting path with respect to M_1 or M_i, and so we obtain a larger matching. If Q ends at x and also passes through y, then clearly it contains at least one point of $A(\mathcal{L})$. But then we can rearrange the lines of M_1 in the odd cycle $Q + xy$ so as to miss this point of $A(\mathcal{L})$, and so we can add a k-element matching to the list $\mathcal{L}$ such that $D(\mathcal{L})$ increases in size.

Finally, if Q ends at x, but does not go through y, then switching on Q, we obtain a new k-element matching M_1' such that z is missed by M_1' and y is missed by $M_1' \cap E(T)$. Since the distance between z and y in T is smaller than the distance between x and y, if we repeat this last procedure at most $|V(T)|$ times, we must arrive at a case previously treated.

Case 2b. Suppose $M_1 \cap E(T)$ misses at least two points x, y of T, but M_1 misses no point of T. Thus x and y are matched with two points u and v in $A(\mathcal{L})$. Let $M_i \in \mathcal{L}$ miss x, and let Q denote the alternating $M_1 - M_i$ path which starts at x. If Q ends with an M_1-line, then switching on Q, we obtain a $(k+1)$-element matching. So suppose that Q ends with an M_i-line. If Q does not go through y, then switching on Q, we obtain from M_1 a k-element matching M_1' which misses x and $M_1' \cap E(T)$ misses y. Thus we are back in Case 2a. If Q goes through y and y is at an even distance from its endpoints on Q, then switching on the appropriate portion of Q, we obtain from M_1 a k-element matching M_1' which misses y and for which $M_1' \cap E(T)$ misses x. Thus we are back to Case 2a once again.

Finally, if y is at an odd distance from the endpoints on Q, then switching on an appropriate portion of Q, we obtain a k-element matching M_1' which misses some v in $A(\mathcal{L})$. But then upon adding M_1' to $\mathcal{L}$, the set $D(\mathcal{L})$ increases.

Case 2c. Suppose $M_1 \cap E(T)$ is a perfect matching of T. Then $|T|$ is even. Suppose $M_i \in \mathcal{L}$ misses a point of T. Since T is even, $M_i \cap E(T)$ must miss another point of T as well, and so yet again we are back in Case 2a.

This completes the description of the algorithm. It is not as efficient as the algorithm of Edmonds. The main reason we have described it here is that it shows how a structural result like that of Gallai and Edmonds can motivate an algorithm.

9.3.2. EXERCISE. Prove that for the final list $\mathcal{L}$ in the preceding algorithm, we have

$$A(\mathcal{L}) = A(G),\ C(\mathcal{L}) = C(G) \text{ and } D(\mathcal{L}) = D(G).$$

9.3.3. EXERCISE. Prove that during the algorithm, one always has $|\mathcal{L}| \leq p$.

9.3.4. EXERCISE. Derive the Gallai-Edmonds Structure Theorem from an analysis of this algorithm.

9.3.5. EXERCISE. Show that the algorithm described in this section can be implemented in $O(p^4)$ time.

9.4. A Linear Programming Algorithm for Matching

One possible motivation for the study of the facets of the matching polytope is its possible application to matching algorithms. The matching polytope is given to us as the convex hull of a certain set. If we can find a description of it as the solution set of a system of linear inequalities, then we can apply the techniques of linear programming to maximize an arbitrary linear objective function over this polytope; that is, we can solve the *weighted* matching problem. Now Theorem 7.3.1 tells us the list of inequalities needed to solve this problem, so it is tempting to try to use this theorem to design a matching algorithm.

Until 1978, there were two main obstacles to accomplishing this task and to designing a polynomial time matching algorithm using this approach. First, there was no polynomial time algorithm known for linear programming. Second, even if we dismiss this obstacle as irrelevant from the practical point of view — after all, the simplex method is very efficient in practice — the difficulty remains that the list of inequalities given in Theorem 7.3.1 is exponentially long for general graphs. Recall from Box 7B, however, that the Ellipsoid Method is not only polynomial even in the worst case, but in addition, it does not need an explicit list of all the constraints!

Using the latter observation, Karp and Papadimitriou (1980, 1982), Padberg and Rao (1981) and Grötschel, Lovász and Schrijver (1981) showed that the Ellipsoid Method also helps to overcome the second difficulty mentioned above, and does indeed lead to a polynomial time matching algorithm based on Edmonds' description of the matching polytope.

We shall see that one more result will be needed. This is an algorithm due to Padberg and Rao (1982) for finding a minimum weight odd cut in a graph, a problem we have discussed previously in Section 6.6.

We shall describe the geometric idea behind this application of the Ellipsoid Method. It seems that at the present time no practical implementation is known and so we do not go into any details concerning the problems of rounding, storage, numerical instability, etc. This algorithm is interesting from a theoretical point of view, because it is based on very general ideas and it is not surprising that it cannot compete with special-purpose algorithms like that of Edmonds.

Let G be a graph and $w : E(G) \to Z_+$, a weighting of its lines. We want to find a matching in G having maximum weight. We may assume that G is a complete graph with an even number of points, for if necessary, we can add a new point and all lines missing from G with weight 0 and we clearly get an equivalent problem in this way. Furthermore, we may restrict ourselves to perfect matchings, since in a complete graph having non-negative weights every matching M extends to a perfect matching F the weight of which is at least the weight of M.

Let $PM(G)$ denote the perfect matching polytope of G. Our task is to maximize a linear objective function $\boldsymbol{w} \cdot \boldsymbol{x}$ over all $\boldsymbol{x} \in PM(G)$. Recall from Theorem 7.3.4 that $PM(G)$ can be described by the constraints

$$\boldsymbol{x} \geq 0 \tag{9.4.1}$$

$$\boldsymbol{x}(\nabla(v)) = 1 \qquad \text{(for all } v \in V(G)) \tag{9.4.2}$$

$$\boldsymbol{x}(\nabla(S)) \geq 1 \qquad \text{(for all odd } S \subseteq V(G)). \tag{9.4.3}$$

Also recall from Box 7B that in order to apply the Ellipsoid Method, all we need is a subroutine to check whether or not a given point $\boldsymbol{x} \in \Re^{E(G)}$ belongs to this polytope.

Let us try to determine just how difficult this task is! If $\boldsymbol{x} \in PM(G)$, this means that $\boldsymbol{x}$ satisfies constraints (9.4.1), (9.4.2) and (9.4.3). The first two groups of inequalities are only $p+q$ in number, so we can check whether or not $\boldsymbol{x}$ satisfies them by direct substitution. It is more difficult, however, to check whether constraints (9.4.3) are satisfied, since there are exponentially many of them. Observe, however, that if we view the vector $\boldsymbol{x}$ as a (non-negative) weighting of the lines of G, then (9.4.3) says that the weight of every $V(G)$-cut is at least 1. So the task of checking whether or not $\boldsymbol{x} \in PM(G)$ could be solved if we had a procedure to find a minimum weight $V(G)$-cut. As we have already mentioned, this last problem was solved by Padberg and Rao (1982). The solution they obtain is based on their theorem (see Theorem 6.6.11 of this book), which says that if F is any cut-equivalent tree with respect to the capacities $\boldsymbol{x}(e)$ ($e \in E(G)$), then the minimum weight of a $V(G)$-cut (i.e., a cut with an odd number of points on both shores) is equal to the minimum weight of the $V(G)$-cuts determined by F. Since there are only $p-1$ cuts determined by F, it is trivial to find a minimum weight $V(G)$-cut among them. Of course, we need a cut-equivalent tree, but this can be found by the algorithm of Gomory and Hu (see Section 2.3). So whether or not $\boldsymbol{x} \in PM(G)$ can be checked in polynomial time.

Thus the Ellipsoid Method can be applied to maximize an arbitrary linear objective function over the perfect matching polytope of a graph in polynomial time.

Another way to state the above argument is to observe that the Ellipsoid Method reduces the Maximum Weight Perfect Matching Problem to the Minimum Weight $V(G)$-Cut Problem in polynomial time. An argument very similar to the above can be presented to show that the Minimum Weight $V(G)$-Cut Problem can be reduced to the Maximum Weight Perfect Matching Problem. So the Ellipsoid Method establishes the *equivalence* of these two problems from the perspective of polynomial-time algorithms. We know combinatorial algorithms which solve these problems directly, namely the algorithms of Edmonds and of Padberg and Rao respectively. It is quite surprising that a general geometric method like the Ellipsoid Method tells us that the solution of either one of these problems automatically implies a solution of the other, at least in so far as we are only concerned with the polynomiality of the running time. Above, we used the Padberg-Rao algorithm to find a maximum weight perfect matching because this appeared to us to be simpler than using the Edmonds Algorithm. Moreover, this reduction leads to more general algorithms (see Box 11B).

Let us point out, however, that Padberg and Rao have proposed another way to apply their Minimum Weight $V(G)$-Cut Algorithm to obtain maximum weight matchings, and while this does not yield an algorithm which is polynomial, it does yield one which is practically efficient. They propose to start with the polytope described by constraints (9.4.1) and (9.4.2) (the so-called **Fractional Perfect Matching Polytope**) and optimize the given linear objective function over this polytope using, say, the Simplex Method. (We can even use a commercial LP-code for this.) If the optimum point $\boldsymbol{x}_1$ is in $PM(G)$, then we were lucky and can stop! If not, then find an inequality in (9.4.3) which is violated by $\boldsymbol{x}_1$, add this as a new constraint and solve the new linear program. We may repeat this until a maximum matching is found, or stop after a time and use the fractional matching found as an upper bound in a branch-and-bound procedure for finding a maximum matching. For a general reference for branch-and-bound methods, see Gondran and Minoux (1979, 1984).

This algorithm is not polynomial, but in practice it appears competitive with implementations of the Edmonds Algorithm. Furthermore, it has the advantage that an algorithm based on similar ideas can even be used to solve NP-complete problems, like the Travelling Salesman Problem, in non-polynomial — but still reasonable — time. For details of this method see Grötschel (1982).

10

The f-factor Problem

10.0. Introduction

A natural generalization of the matching problem can be formulated as follows. It is known as the **f-factor problem**. Assume that we assign an integer $f(v)$ to each point $v \in V(G)$. Does there exist an **f-factor**, that is, a spanning subgraph H of G such that $\deg_H(v) = f(v)$ for each point v? If not, how well can f be approximated by the degrees of a spanning subgraph? The latter — more nebulous — problem has come to be called the **degree-constrained subgraph problem**. The aim of this chapter is to answer these questions, by introducing a decomposition of $V(G)$ analogous to the Gallai-Edmonds decomposition defined in Chapter 3.

There is a problem closely related to this which is, in fact, equivalent in the case when $f(v) = 1$ for all v, but which should be distinguished carefully in the general case. We shall call this the **f-matching problem**. Given, as before, an integer-valued function f defined on $V(G)$, can we assign a non-negative integer $x(e)$ to every line e such that

$$\boldsymbol{x}(\nabla(v)) = \sum_{\{e \mid v \in e\}} x(e) = f(v)$$

holds for each point v? Such an assignment of values $x(e)$ will be called a **perfect f-matching**. Clearly, f-factors can be identified with those perfect f-matchings which satisfy the additional constraint $x(e) \leq 1$. On the other hand, a perfect f-matching may be viewed as a graph G' such that $V(G') = V(G)$, each line of G' is parallel to some line of G, and $\deg_{G'}(v) = f(v)$ holds for each point v.

The problem of the existence of a perfect f-matching is a special case of the problem of the existence of an f-factor: if we replace each line of G by a very large number of parallel lines, then the f-factors of the resulting graph may be viewed as perfect f-matchings of the original and vice versa. But considering f-matchings separately is sometimes a useful intermediate step toward the study of f-factors.

Obviously, in the case when $f(v) = 1$ for all v, 1-factors and perfect 1-matchings are the same.

If we want to investigate f-factors, then we may try to make use of information we might already have about 1-factors, i.e., perfect matchings. There are several ways to do this. First, we may try to reduce the f-factor problem to the 1-factor problem, by constructing from the given graph G another graph G' such that G' has a perfect matching if and only if G has an f-factor. That such a construction is possible, was first observed by Tutte (1954). This leads to a rather straightforward, but somewhat tedious, treatment of the f-factor problem. We shall sketch this approach in Section 10.1.

Second, we may generalize the methods developed to solve the 1-factor problem. The original solution of the f-factor problem, due to Tutte (1952), the work of Ore (1957), and in a sense the treatment of Graver and Jurkat (1980) all have this flavor. Third, there is the possibility of obtaining more general results which have no meaning in the special case of 1-factors. To illustrate this idea, let us remark only that G has an f-factor if and only if it has an $\hat{f}$-factor, where

$$\hat{f}(v) = \deg_G(v) - f(v).$$

This "complementarity" relation is not present if we restrict ourselves to the case of 1-factors.

The main body of our discussion will be the development of a structure theory, analogous to the Gallai-Edmonds decomposition. The methods which yield the decomposition are based upon a further generalization of the problem and upon the freedom gained from this greater generality (Lovász (1970a, 1970b, 1972c)). It is also possible to generalize some of the structure theory of Chapter 5 to the case of graphs having an f-factor. Since this extension is somewhat lengthy we do not repeat it here but refer the reader to another paper of Lovász (1972b).

In the last section of this chapter we treat the problem of realization of degree sequences. This problem is a special instance of the f-factor problem, corresponding to the case when the given graph is complete. There are, however, several more direct approaches for obtaining the fundamental result of Erdős and Gallai (1960) characterizing realizable degree sequences, as well as various extensions of their result. We also discuss the problem of realizing a degree sequence by a graph with a perfect matching.

10.1. Reduction Principles

We begin by sketching a method for reducing the f-factor problem to the 1-factor (or perfect matching) problem. The construction we present is due to Tutte (1954). It can be used to derive a necessary and sufficient condition for the existence of an f-factor, but we postpone this to a later section where a direct proof will be given, and omit the straightforward but tedious derivation from the simple construction that follows.

To begin with, let us show how the f-matching problem can be reduced to the matching problem. Let G be a graph and f an integer-valued function defined on $V(G)$. To exclude trivial cases, let us assume that $f \geq 0$.

Construct a graph G' as follows. For each $v \in V(G)$ let U_v be a set of $f(v)$ elements so that U_v and U_w are disjoint if $v \neq w$. Let

$$V(G') = \bigcup_{v \in V(G)} U_v.$$

Connect each element of U_v to each element of U_w, whenever v and w are adjacent in G.

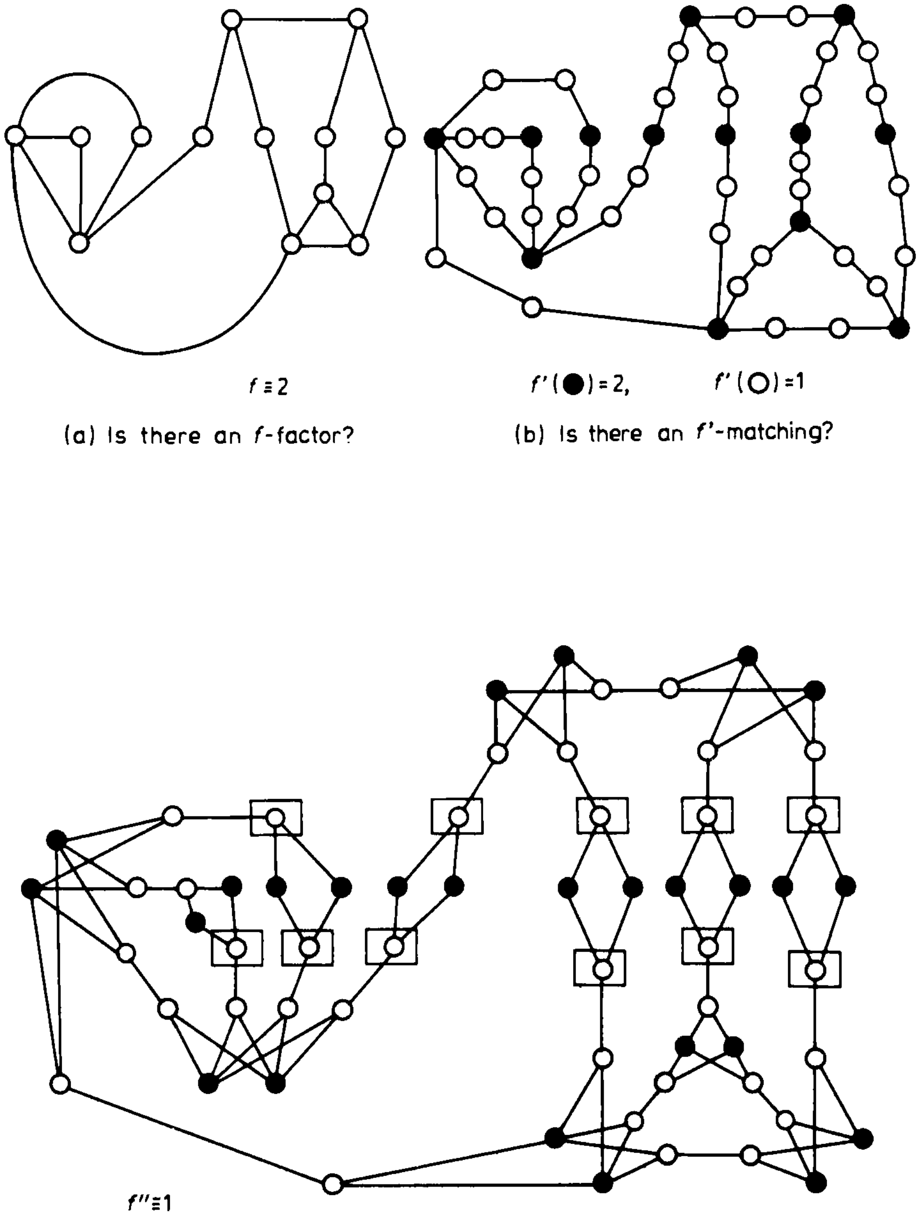
10.1.1. THEOREM. *The graph G' has a perfect matching if and only if G has a perfect f-matching.*

PROOF. We may assume without loss of generality that G has no multiple lines. Suppose now that G' has a perfect matching F. For each line $vw \in E(G)$, let $x(vw)$ denote the number of lines of F connecting U_v and U_w. Then the number of lines of F incident with U_v is exactly $\sum_w x(vw)$. But since F is a perfect matching of G', this number is also $|U_v| = f(v)$. This shows that the q-vector $\boldsymbol{x}$ is a perfect f-matching.

Conversely, every perfect f-matching of G yields a perfect matching of G' by a similar construction. ∎

Let us now regard the existence problem of perfect f-matchings as solved. Assume that we are given a graph G and an integer-valued function f on $V(G)$, and suppose we want to decide whether or not G has an f-factor. Let us construct a new graph G'' by subdividing each line of G with two new points. Extend the domain of f to $V(G'')$ by letting $f(v) = 1$ on each new point.

10.1.2. THEOREM. *The graph G'' has a perfect f-matching if and only if G has an f-factor.*



$f \equiv 2$

(a) Is there an f-factor?

$f'(\bullet) = 2$, $f'(\circ) = 1$

(b) Is there an f'-matching?

$f'' \equiv 1$

(c) Is there a perfect matching? No. See Tutte-set $\square$.

FIGURE 10.1.1. Reduction of the f-factor problem to the 1-factor problem

PROOF. The key to this equivalence is the observation that every perfect f-matching of G'' is automatically an f-factor. Hence it suffices to show that G'' has an f-factor if and only if G has one.

Suppose that F is an f-factor of G. Then we can construct an f-factor F'' of G'' as follows. For each line $e \in F$, take the two endlines of the 3-path in G'' corresponding to e; for each $e \in E(G) - F$, take the middle line of the corresponding 3-path. The set F'' of lines of G'' obtained in this way is clearly an f-factor of G''.

Conversely, we can construct an f-factor of G from any f-factor of G'' by reversing this procedure. ∎

On the basis of the two preceding theorems, we can reduce the problem of the existence of an f-factor to the problem of the existence of a perfect matching.

It is possible at this point to translate Tutte's 1-factor criterion into Tutte's f-factor criterion. But this is somewhat tedious and yields little insight into the structure of f-factors. Another shortcoming of this translation is that the matching algorithms are translated into f-matching algorithms which are polynomial only if the function f is given in the unary sense (cf. Box 2B). We shall not enter into the study of f-matching and f-factor algorithms here, but instead refer the interested reader to the following exercises. In addition, the reader is referred to Pulleyblank (1973), Marsh (1979) and Anstee (1983). A good up-to-date survey of the f-matching problem may be found in Schrijver (1983a).

10.1.3. EXERCISE. (For this exercise only, we allow loops!) Show that the f-factor problem for graphs with loops can be reduced to the f-factor problem for graphs without loops.

10.1.4. EXERCISE. Let G be a graph and $f : V(G) \to Z_+$. Show that the following two problems: "Find a maximal subgraph H with $\deg_H \leq f$" and "Find a minimal subgraph H' with $\deg_H \geq f$" are equivalent (in a sense) for every G and f. Also show that these problems can be reduced to the f-factor problem.

10.1.5. EXERCISE. Show that one can reduce the following to the f-factor problem: "Find a subgraph H with $\max\{|f(v) - \deg_H(v)| \mid v \in V(G)\}$ as small as possible."

10.2. A Structure Theory for f-factors

We start by further generalizing the f-factor problem. Assume that we are given two functions $f_1 \geq f_2$ on $V(G)$, and suppose we are looking for a spanning subgraph H such that

$$f_1(v) \geq \deg_H(v) \geq f_2(v)$$

for each point v. (The hypothesis that H is spanning means only that $V(H) = V(G)$, and does not mean that the lines of H span G. Thus $\deg_H(v) = 0$ is allowed.) To get a more symmetric formulation of this problem, set $f = f_1$, $g = \hat{f}_2 = \deg_G - f_2$, and for any spanning subgraph H of G, denote by $\overline{H}$ the spanning subgraph $(V(G), E(G) - E(H))$. Then the conditions on H can be written as follows:

$$\deg_H \leq f \quad \text{and} \quad \deg_{\overline{H}} \leq g. \tag{10.2.1}$$

To exclude trivial cases, we shall always assume that

$$f, g \geq 0 \quad \text{and} \quad f + g \geq \deg_G. \tag{10.2.2}$$

Let H be an arbitrary spanning subgraph of G and define the following quantities:

$$\delta(v; H; f, g) = \max\{0, \deg_H(v) - f(v), \deg_{\overline{H}}(v) - g(v)\},$$
$$\delta(H; f, g) = \sum_{v \in V(G)} \delta(v; H; f, g),$$

and

$$\delta(f, g) = \min_H \delta(H; f, g).$$

Trivially, a spanning subgraph H satisfying inequalities (10.2.1) exists if and only if $\delta(f, g) = 0$. Such a spanning subgraph will be called an (f, g)-**factor**. In particular, an f-factor exists if and only if $\delta(f, \hat{f}) = 0$. A further trivial observation is that

$$\delta(H; f, g) = \delta(\overline{H}, g, f)$$

and hence

$$\delta(f, g) = \delta(g, f).$$

We call a spanning subgraph H (f, g)-**optimal**, if $\delta(H; f, g) = \delta(f, g)$. Our goal is to derive a minimax formula for the value of $\delta(f, g)$ and to describe the structure of (f, g)-optimal subgraphs.

10.2.1. EXERCISE. Suppose $f+g = \deg_G$. Prove that $\delta(f,g) \equiv f(V(G))$ (mod 2).

In the case when $f(v) = 1$ and $g(v) = \deg_G(v)-1$ for all $v \in V(G)$ (i.e., in the 1-matching case) every maximum matching is (f,g)-optimal as is every minimum line cover. Moreover, if we extend a maximum matching to a minimum line cover (see Exercise 1.0.3), then every intermediate subgraph is also (f,g)-optimal. Furthermore, $\delta(f,g) = \mathrm{def}(G)$. These assertions can be verified easily using Theorem 10.2.5 below.

It should be emphasized also that the apparent greater generality gained by having upper and lower bounds on the degree, rather than one prescribed value, is really only technical. If fact, the (f,g) problem can be reduced to the f-factor problem (see the next exercise) and hence to the 1-matching problem. This reduction shows also that $\delta(f,g)$ can be computed in polynomial time.

10.2.2. EXERCISE. Given graph G, two functions f and g on $V(G)$ as above and a $k \geq 0$, construct a graph G_1 and a function f_1 on $V(G_1)$ such that $\delta(f,g) \leq k$ if and only if G_1 has an f_1-factor.

Let us make a few additional remarks about the generalization of the f-factor problem we are considering. Instead of prescribing a single value, we have allowed the degree of the subgraph to belong to an arbitrarily prescribed interval. We can further generalize the problem by prescribing an arbitrary set $I(v)$ of integers for every point v, and then look for an I-factor, that is, a subgraph H such that $\deg_H(v) \in I(v)$ for all $v \in V(G)$. Lovász (1972c) showed that 3-line-colorability is reducible to the case when $I(v) = \{1\}$ or $\{0,3\}$ at each point, and so by the result of Holyer (1981) it follows that this problem for arbitrary sets $I(v)$ is NP-complete. But in the special case when $I(v)$ is allowed to have only one-element gaps, most of the discussion presented in this section can be carried over.

Let us mention a few special cases of the problem. If $I(v)$ is a set of consecutive odd integers for $v \in T$ and a set of consecutive even integers for $v \in V(G) - T$, then an I-factor is a T-join, the degrees of which are bounded from above and below. The reader may verify that the question as to whether such a T-join exists can easily be reduced to the f-factor problem.

A second special case arises when $I(v)$ contains all integers *except exactly one* integer $h(v)$ for each $v \in V(G)$. In this case we are looking for a subgraph the degree of which is different from a prescribed value at each point. This problem is quite easily solved in polynomial time (see Lovász (1973)).

On the other hand, the following quite similar problem is not known to be polynomially solvable. Let $G = (A, B)$ be a bipartite graph. Does G have a spanning subgraph H such that $\deg_H(v) = 1$ for $v \in A$ and $\deg_H(v) \neq 1$ for $v \in B$?

We do know the following polynomially solvable special case of this problem, however. Let G_0 be any graph, $A = V(G_0)$ and let B be the set of all lines and triangles of G_0. Let G be the bigraph with bipartition (A, B) and all lines of the form ab with $a \in V(b)$. Then G has a subgraph H as above, if and only if G_0 contains a set of point-disjoint lines and triangles covering $V(G_0)$. The latter problem was solved in polynomial time by Cornuéjols, Hartvigsen and Pulleyblank (1982). (See Exercise 6.4.4.)

At this point, it is an easy matter to define a partition of the point set of G into four classes, analogous in a way to the Gallai-Edmonds partition. But it turns out to be a lot more work to derive the important properties of this partition! However, the definition of the partition is easy enough. In the 1-matching case, we based our decomposition upon whether or not a given point is covered by all maximum matchings. In the present case an (f, g)-optimal subgraph may violate either one of the inequalities in (10.2.1). Accordingly we have four natural possibilities.

$$C = C(G; f, g) = \{v \in V(G) \mid \deg_H(v) \leq f(v) \text{ and } \deg_{\overline{H}}(v) \leq g(v) \text{ for every } (f,g)\text{-optimal subgraph } H\},$$

$$A = A(G; f, g) = \{v \in V(G) - C \mid \deg_{\overline{H}}(v) \leq g(v) \text{ for every } (f,g)\text{-optimal subgraph } H\},$$

$$B = B(G; f, g) = \{v \in V(G) - C \mid \deg_H(v) \leq f(v) \text{ for every } (f,g)\text{-optimal subgraph } H\},$$

$$D = D(G; f, g) = V(G) - A - B - C.$$

When graph G is understood, we shall write $C(f, g)$, $A(f, g)$, $B(f, g)$ and $D(f, g)$ for these four sets respectively. This partition of the points will play an important role throughout the discussion to follow. Note that $C = V(G)$ if and only if G has an (f, g)-factor.

In the 1-matching case the structure introduced above is related to the Gallai-Edmonds decomposition as follows: $C(G; f, g) = C(G)$, $A(G; f, g) = A(G)$, $B(G; f, g)$ is the union of all singleton components of $D(G)$ and $D(G; f, g)$ is the union of the point sets of the non-singleton components of $D(G)$. (Recall that this distinction between singleton and non-singleton components of $D(G)$ arose earlier, for example, in the proof of Theorem 6.3.1.)

10.2.3. EXERCISE. Show that if the graph shown in Figure 10.2.1 has $f(v) = 2$ and $g(v) = \deg_G(v) - 2$ for each $v \in V(G)$, then G has the (A, B, C, D) decomposition shown.

10.2.4. EXERCISE. If sets A, B, C and D are as defined above, then these sets can be determined in polynomial time.

Before we can develop the properties of the partition $\{A, B, C, D\}$ we need a bit more preparation. Let H be any spanning subgraph of G

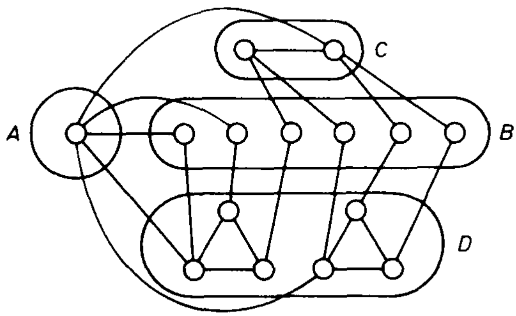


FIGURE 10.2.1.

and let v be a point such that $\deg_H(v) > f(v)$. Delete from H any line vw incident with v. The resulting subgraph H' has

$$\begin{aligned}
&\delta(v; H'; f, g) = \delta(v; H; f, g) - 1,\\
&\delta(w; H'; f, g) \leq \delta(w; H, f, g) + 1, \text{ and}\\
&\delta(u; H'; f, g) = \delta(u; H; f, g) \text{ for all } u \neq v, w.
\end{aligned}$$

Hence

$$\delta(H'; f, g) \leq \delta(H; f, g).$$

If, in particular, H is (f, g)-optimal, then so is H' and it follows that we must have equality in the preceding inequalities; that is, $\deg_{\overline{H}}(w) \geq g(w)$.

Similarly, if $\deg_{\overline{H}}(v) > g(v)$, then we may add to H any line of $\overline{H}$ incident with v and obtain a subgraph H' with $\delta(H'; f, g) \leq \delta(H; f, g)$.

The addition or deletion of a line so that $\delta(v;H';f,g)$ strictly decreases for at least one point v will be called a **local augmentation at v (with respect to (f,g)).**

Let us make several immediate observations. If H is (f,g)-optimal, and H' is obtained from H by a local augmentation, then it is also true that H can be obtained from H' by a local augmentation. If H is (f,g)-optimal, then no local augmentation can change any line incident with the set C. The remarks just prior to the definition of local augmentation above imply that, using local augmentations, we can construct from any spanning subgraph one which satisfies $\deg_H \leq f$. Consequently, we have the next result.

10.2.5. THEOREM. *For a graph G and two functions f and g satisfying conditions* (10.2.2), *there exists an (f,g)-optimal subgraph H satisfying* $\deg_H \leq f$. ■

Local augmentations turn out to play a role analogous to that of augmenting paths in the 1-matching case. The power of this technique is a consequence of the following lemma.

10.2.6. LEMMA. *Let J be an (f,g)-optimal subgraph and H any spanning subgraph. Then H can be transformed by local augmentations into an (f,g)-optimal subgraph H_0 such that $E(J) \oplus E(H_0)$ contains only lines spanned by C.*

PROOF. Consider two subgraphs H' and J' with the following three properties:

(a) H' can be obtained from H by local augmentations;
(b) J' can be obtained from J by local augmentations;
(c) $|E(H') \oplus E(J')|$ is minimal.

We derive several additional properties of the pair (H',J').

(1) For each point v, we have

$$\delta(v;H';f,g) \leq \delta(v;J';f,g).$$

For suppose, to the contrary, that $\delta(v;H';f,g) > \delta(v;J';f,g)$ holds for some point v. Suppose, for example, $\deg_{H'}(v) > f(v)$. Then obviously $\deg_{J'}(v) < \deg_{H'}(v)$. Therefore $E(H') - E(J')$ has a line e incident with v. The removal of e from H' is a local augmentation and decreases $|E(G') \oplus E(J')|$. This contradicts (c).

(2) The subgraph H' is (f,g)-optimal and

$$\delta(v;H';f,g) = \delta(v;J';f,g)$$

holds for each point v.

This follows immediately from (1) and from the fact that by (b), J' is (f,g)-optimal.

(3) If $\delta(v;J';f,g)>0$, then $\deg_{H'}(v)=\deg_{J'}(v)$.

Suppose, for example, that $\deg_{H'}(v)>\deg_{J'}(v)$. Property (2) can hold then only if the interval $[\hat{g}(v),f(v)]$ lies between the degrees $\deg_{H'}(v)$ and $\deg_{J'}(v)$. Thus if we delete any line of $E(H')-E(J')$ incident with v from H', we carry out a local augmentation, and so we get a contradiction of (c) as before.

(4) If $\delta(v;H';f,g)>0$, then no line of $E(H')\oplus E(J')$ is incident with v.

Suppose, to the contrary, that e is such a line and suppose, say, that $e\in E(H')-E(J')$. We know by (3) that $\deg_{H'}(v)=\deg_{J'}(v)$. Moreover, either $\deg_{H'}(v)>f(v)$ or $\deg_{\overline{H}'}(v)>g(v)$. In the first case the removal of e from H' and in the second, the addition of e to J' is a local augmentation which decreases $|E(H')\oplus E(J')|$. In both cases we get a contradiction of (c).

(5) If $v\notin C$ then no line of $E(H')\oplus E(J')$ is incident with v.

To prove this claim choose an (f,g)-optimal subgraph K such that $\delta(v;K;f,g)>0$. If there are other ways to choose H' and J' subject to (a), (b) and (c), but contradicting (5), choose this pair so that the following condition is also fulfilled:

(d) $|E(H')\oplus E(K)|$ is minimal.

Then no local augmentation can be carried out on H' which would decrease $E(H')\oplus E(K)$. For such a local augmentation at some point w could also be carried out on J' as a local augmentation by (4), and so it would yield a pair (H'',J'') of spanning subgraphs satisfying (a), (b) and (c), but still violating (5) (since $E(H'')\oplus E(J'')=E(H')\oplus E(J')$), thus yielding a contradiction of (d).

By the arguments for (1) and (2) this implies that $\delta(v;H';f,g)=\delta(v;K;f,g)>0$. But then we have a contradiction of (4). This completes the proof of (5).

Thus we know that H' and J' are (f,g)-optimal and $E(H')\oplus E(J')$ is spanned by C. As remarked before, we can go back from J' to J by local augmentations. But by (5), the same local augmentations can be carried out on H' as well, thus yielding a spanning subgraph H_0. Hence H_0 is (f,g)-optimal and $E(H_0)\oplus E(J)=E(H')\oplus E(J)$ is spanned by C.■

10.2.7. COROLLARY. *Every subgraph can be transformed into an (f,g)-optimal subgraph by local augmentations.* ■

10.2.8. COROLLARY. *Two (f,g)-optimal subgraphs can be transformed into each other by local augmentations if and only if any line spanned by C belongs either to both or to neither of the two subgraphs.* ∎

10.2.9. EXERCISE. Let G be a factor-critical graph and let M and M' be two near-perfect matchings in G. Then there exists a sequence $M_1 = M, M_2, \dots, M_k = M'$ of near-perfect matchings such that $|M_i \oplus M_{i+1}| = 2$ for $i = 0, \dots, k-1$.

As a further application of Lemma **10.2.6**, we prove the following theorem which will be used later, but which is of interest in its own right.

10.2.10. THEOREM. *Let G be a graph and let $f_1 \le f_2 \le \cdots \le f_k$ and $g_1 \le g_2 \le \cdots \le g_k$ be integer-valued functions defined on $V(G)$ such that $f_1 + g_1 \ge \deg_G$, $f_1 \ge 0$, $g_1 \ge 0$. Then G has a spanning subgraph which is simultaneously (f_i, g_i)-optimal for every $1 \le i \le k$.*

PROOF. We proceed by induction on k. For $k = 1$ the assertion is trivial. Let H' be a subgraph of G which is (f_i, g_i)-optimal for $1 \le i \le k-1$. By Corollary 10.2.7 we can transform H' into an (f_k, g_k)-optimal subgraph H by local augmentations with respect to (f_k, g_k). But obviously any local augmentation with respect to (f_k, g_k) is also a local augmentation with respect to (f_i, g_i) for $1 \le i \le k-1$, and so it preserves the (f_i, g_i)-optimality of H'. Hence H is (f_i, g_i)-optimal for every $1 \le i \le k-1$. ∎

The following simple, but important, observation is also an immediate consequence of Lemma 10.2.6.

10.2.11. THEOREM. *For any point $v \notin C$, the degrees of v in (f,g)-optimal subgraphs form a sequence of consecutive integers.*

PROOF. Suppose that both a and b occur as the degrees of v in two (f,g)-optimal subgraphs H and J, with, say, $\deg_H(v) = a$, $\deg_J(v) = b$. By Lemma 10.2.6, we can transform H by local augmentations into another (f,g)-optimal subgraph H' such that $\deg_{H'}(v) = \deg_J(v)$. Since local augmentations change the degree of v by at most 1, the degrees of v in the "intermediate" subgraphs will cover every integer in the interval $[a,b]$. ∎

Let us remark that the assertion of the above theorem is not valid for the points in C. In Figure 10.2.2 we show two functions f, g defined on the points of the triangle, so that the degree of the dark point in (f,g)-optimal subgraphs is either 0 or 2.

We are now prepared for the study of the partition $\{A, B, C, D\}$. We start with some simple properties which follow easily from Theorem 10.2.10.

It will be convenient to introduce the following idea. Given the integer-valued function f defined on $V(G)$ and a point $v \in V(G)$, define a new function f^v also on $V(G)$ as follows:

$$f^v(w) = \begin{cases} f(w)+1, & \text{if } v = w \\ f(w) \quad , & \text{if } v \neq w. \end{cases}$$

Note that we can compose such functions and let us denote $f^{uv} = (f^u)^v$ where u and v are arbitrary points in $V(G)$. More specifically, we shall make use of the function f^{vv} which clearly amounts to the following:

$$f^{vv}(w) = \begin{cases} f(w)+2, & \text{if } v = w \\ f(w) \quad , & \text{if } v \neq w. \end{cases}$$

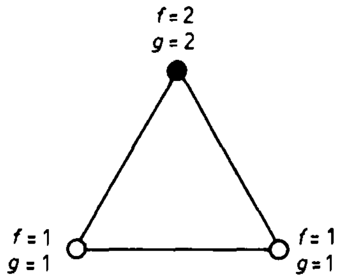


FIGURE 10.2.2.

10.2.12. THEOREM. *Let G be a graph and let f and g be two functions satisfying inequalities* (10.2.2). *Then:*

(a) *If $v \in A$, then $\deg_H(v) \geq f(v)$ for every (f, g)-optimal subgraph H.*
(b) *If $v \in B$, then $\deg_{\overline{H}}(v) \geq g(v)$ for every (f, g)-optimal subgraph H.*
(c) *If $v \in D$, then $f(v) = \hat{g}(v)$ and the set of degrees of v in (f, g)-optimal subgraphs is $\{f(v)-1, f(v), f(v)+1\}$.*

PROOF. (a) Suppose, to the contrary, that there exists an (f, g)-optimal subgraph H with $\deg_H(v) < f(v)$. Since $\deg_H(v) \geq \hat{g}(v)$ by the definition of A, this implies that $f(v) > \hat{g}(v)$. By Theorem 10.2.5 we can transform H by local augmentations into an (f,g)-optimal subgraph J such that $\delta(v;J;f,g)>0$, and so by the definition of A, $\deg_J(v)>f(v)$. At some step during this sequence of local augmentations the degree of v changes from $f(v)-1$ to $f(v)$. But this cannot happen for a local augmentation of an (f,g)-optimal subgraph.

The proof of (b), as well as the proof of the first assertion of (c), is analogous. It is also clear, by the definition of D and by Theorem 10.2.11, that $f(v)-1$, $f(v)$ and $f(v)+1$ must occur as the degrees of v in some three (f,g)-optimal subgraphs. Thus it suffices to prove (in view of Theorem 10.2.11) that $f(v)+2$ and $f(v)-2$ are not degrees of v in any (f,g)-optimal subgraph.

Assume, to the contrary, that $\deg_H(v)=f(v)+2$ for some (f,g)-optimal subgraph H. Then

$$\delta(H;f^{vv},g^v)=\delta(H;f,g)-2$$

and hence

$$\delta(f^{vv},g^v)\leq\delta(f,g)-2. \tag{10.2.3}$$

It follows by a similar argument that

$$\delta(f,g^v)\leq\delta(f,g)-1. \tag{10.2.4}$$

By Theorem 10.2.10, there exists a spanning subgraph J which is simultaneously (f,g)-optimal, (f,g^v)-optimal and (f^{vv},g^v)-optimal. But then inequality (10.2.3) implies that

$$\delta(J;f^{vv},g^v)\leq\delta(J;f,g)-2, \tag{10.2.5}$$

while inequality (10.2.4) implies that

$$\delta(J;f,g^v)\leq\delta(J;f,g)-1. \tag{10.2.6}$$

But inequality (10.2.6) implies that $\deg_J(v)<f(v)$ and, moreover, we clearly have equality in (10.2.6). However, then

$$\delta(J;f^{vv},g^v)=\delta(J;f,g^v)=\delta(J;f,g)-1,$$

which contradicts inequality (10.2.5). ∎

Our next theorem describes the status of lines incident with C with regard to being in optimal subgraphs.

10.2.13. THEOREM. *Let G be a graph, let f, g be two functions satisfying inequalities* (10.2.2) *and let $\{A, B, C, D\}$ be the partition of the point set $V(G)$ defined above. Then:*

(a) *No line connecting A to C or spanned by A belongs to any (f,g)-optimal subgraph.*
(b) *Every line connecting B to C or spanned by B belongs to every (f,g)-optimal subgraph.*
(c) *No line connects C to D.*

PROOF. (a) Let uv be any line with $u \in A$ and $v \in A \cup C$, and suppose, to the contrary, that there exists an (f,g)-optimal subgraph H containing uv. By the definition of A and by Lemma 10.2.6, H can be transformed by local augmentations into an (f,g)-optimal subgraph J such that $\deg_J(u) > f(u)$. Note that by the definition of A and C, the deletion of uv from any (f,g)-optimal subgraph K cannot increase $\delta(K;u;f,g)$ or $\delta(K;v;f,g)$ and so it cannot be a local augmentation. Thus uv must belong to J. But then the deletion of uv from J is a local augmentation, which contradicts the observation just made.

Parts (b) and (c) follow by similar arguments. ■

The following two lemmas are analogues of the Stability Lemma 3.2.2, but they deal with the more general case of f-factors.

10.2.14. LEMMA. *Suppose G is a graph with integer-valued functions f_1, f_2, g_1 and g_2 defined on $V(G)$. Furthermore, suppose $f_1 \le f_2$, $g_1 \le g_2$ and $f_1 + g_1 \ge \deg_G$. Then*

$$A(f_1, g_1) \supseteq A(f_2, g_2),$$
$$B(f_1, g_1) \supseteq B(f_2, g_2),$$

and

$$D(f_1, g_1) \supseteq D(f_2, g_2).$$

PROOF. We will show that if $v \notin C(f_2, g_2)$, then there exists an (f_1, g_1)-optimal subgraph H_1 with $\deg_{H_1}(v) > f_1(v)$ if and only if there exists an (f_2, g_2)-optimal subgraph H_2 with $\deg_{H_2}(v) > f_2(v)$. This will clearly imply the lemma.

I. Assume first that there exists an (f_2, g_2)-optimal subgraph H_2 with $\deg_{H_2}(v) > f_2(v)$. Then clearly

$$\delta(f_2^v, g_2) \le \delta(f_2, g_2) - 1.$$

Since $f_1 \leq f_2^v$ and $g_1 \leq g_2^v$, Theorem 10.2.10 implies that there exists a subgraph H which is simultaneously (f_1, g_1)-optimal and (f_2^v, g_2)-optimal. But then

$$\delta(H; f_2^v, g_2) = \delta(f_2^v, g_2) \leq \delta(f_2, g_2) - 1 \leq \delta(H; f_2, g_2) - 1.$$

This implies that $\deg_H(v) > f_2(v) \geq f_1(v)$.

II. Second, assume that there exists an (f_1, g_1)-optimal subgraph H_1 such that $\deg_{H_1}(v) > f_1(v)$. Then

$$\delta(f_1^v, g_1) \leq \delta(f_1, g_1) - 1.$$

Furthermore, we have by hypothesis that $v \in C(f_2, g_2)$ and so

$$\delta(f_2^v, g_2^v) \leq \delta(f_2, g_2) - 1.$$

Consider now a spanning subgraph H which is simultaneously (f_1^v, g_1)-optimal and (f_2^v, g_2^v)-optimal. (Such a subgraph exists again by Theorem 10.2.10). Then we have

$$\delta(H; f_1^v, g_1) = \delta(f_1^v, g_1) \leq \delta(f_1, g_1) - 1 \leq \delta(H; f_1, g_1) - 1,$$

and so $\deg_H(v) > f_1(v)$. Similarly,

$$\delta(H; f_2^v, g_2^v) = \delta(f_2^v, g_2^v) \leq \delta(f_2, g_2) - 1 \leq \delta(H; f_2, g_2) - 1, \tag{10.2.7}$$

and so $\deg_H(v) \notin [g_2(v), f_2(v)]$. But since $\deg_H(v) > f_1(v)$, this means that $\deg_H(v) > f_2(v)$.

To conclude, it suffices to show that H is (f_2, g_2)-optimal. But this is clear since trivially

$$\delta(J; f_2^v, g_2^v) \geq \delta(J; f_2, g_2) - 1,$$

for any spanning subgraph J, and so we must have equality in (10.2.7). ■

To formulate our second "stability lemma", let us define, using a given pair of functions (f, g), two further pairs as follows.

$$\bar{f}(v)=\begin{cases}\deg_G(v), & \text{if } v\in B\cup C,\\ f(v), & \text{if } v\in A\cup D,\end{cases}$$

$$\bar{g}(v)=\begin{cases}\deg_G(v), & \text{if } v\in A\cup C,\\ g(v), & \text{if } v\in B\cup D,\end{cases}$$

$$f'(v)=\begin{cases}\deg_G(v), & \text{if } v\in B\cup C,\\ 0, & \text{if } v\in A,\\ f(v), & \text{if } v\in D,\end{cases}$$

$$g'(v)=\begin{cases}\deg_G(v), & \text{if } v\in A\cup C,\\ 0, & \text{if } v\in B,\\ g(v), & \text{if } v\in D.\end{cases}$$

10.2.15. LEMMA. *If G is any graph and if all symbols are as defined above, then $A(f',g') = A(\bar{f},\bar{g}) = A(f,g)$, $B(f',g') = B(\bar{f},\bar{g}) = B(f,g)$, $C(f',g') = C(\bar{f},\bar{g}) = C(f,g)$ and $D(f',g') = D(\bar{f},\bar{g}) = D(f,g)$. Furthermore, the following equation holds:*

$$\delta(f,g)=\delta(\bar{f},\bar{g})=\delta(f',g')-f(A)-g(B).$$

PROOF. Considering any spanning subgraph H which is simultaneously (f,g)-optimal and $(\bar{f},\bar{g})$-optimal (and such must exist by Theorem 10.2.10), we see that

$$\delta(f,g)=\delta(\bar{f},\bar{g}).$$

By Lemma 10.2.14 we have $A(\bar{f},\bar{g}) \subseteq A(f,g)$, $B(\bar{f},\bar{g}) \subseteq B(f,g)$, and $D(\bar{f},\bar{g}) \subseteq D(f,g)$. We now claim, moreover, that $C(\bar{f},\bar{g}) \subseteq C(f,g)$. To prove this, suppose $v \notin C(f,g)$. Then there exists an (f,g)-optimal subgraph H such that, e.g., $\deg_H(v) > f(v)$. But then $\bar{f}(v) = f(v)$ by definition. Clearly H is also $(\bar{f},\bar{g})$-optimal, and this proves that $v \notin C(\bar{f},\bar{g})$. Thus we have proved all the statments in the lemma involving $(\bar{f},\bar{g})$.

Now we turn to (f',g'). Again, considering a subgraph H which is simultaneously $(\bar{f},\bar{g})$-optimal and (f',g')-optimal (and such must exist by Theorem 10.2.10), we see that

$$\delta(\bar{f},\bar{g})=\delta(f',g')-g(A)-f(B).$$

By Lemma 10.2.14 we have

$$A(\bar{f},\bar{g})\subseteq A(f',g'),\ B(\bar{f},\bar{g})\subseteq B(f',g') \text{ and } D(\bar{f},\bar{g})\subseteq D(f',g').$$

Furthermore, if $v \in C(\bar{f}, \bar{g})$, then $f'(v) = g'(v) = \deg_G(v)$ and hence by definition of C, we have $v \in C(f', g')$. So $C(\bar{f}, \bar{g}) \subseteq C(f', g')$. But then we must have equality and the lemma follows. ■

Our last task is to describe the structure of D. As preparation for this, we introduce and study the f-factor analogues of factor-critical graphs.

Let G be a graph and suppose f and g are integer-valued functions defined on $V(G)$. We say that G is (f, g)-**critical**, if it is connected and has $A = B = C = \emptyset$. By Theorem 10.2.12(c), in this case we must have $f = \hat{g}$. It is easy to show that a graph G is (f, g)-critical if and only if it is connected and $\delta(f^v, g)$, $\delta(f, g^v) < \delta(f, g)$ for all $v \in V(G)$. (In fact, this property suggested the name.)

The following result is a generalization of Gallai's Lemma 3.1.13. Note that in our presentation, it is a straightforward consequence of the preceding results.

10.2.16. THEOREM. *Let G be (f, g)-critical. Then $\delta(f, g) = 1$.*

PROOF. Choose any $v \in V(G)$ and consider the pair (f^v, g). By the Stability Lemma 10.2.14, we have

$$V(G) = D(f, g) = D(f^v, g) \cup C(f^v, g).$$

But by Theorem 10.2.13(c), no line connects $D(f^v, g)$ to $C(f^v, g)$. So since G is connected, we have either $D(f^v, g) = V(G)$ or $C(f^v, g) = V(G)$. The first of these two possibilities cannot occur as $f^v(v) = f(v) + 1 > \hat{g}(v)$. Thus $C(f^v, g) = V(G)$; that is, $\delta(f^v, g) = 0$. But then

$$\delta(f, g) \leq \delta(f^v, g) + 1 = 1.$$

Since $\delta(f, g) \neq 0$, the theorem follows. ■

10.2.17. COROLLARY. *If G is (f, g)-critical then $f(V(G))$ is odd.* ■

We are now prepared to prove the main structure theorem, which — together with Theorem 10.2.13 — describes the structure of (f, g)-optimal subgraphs and is the basis of the good characterization of $\delta(f, g)$ provided by the minimax result of Theorem 10.2.19. Define the functions f_B, g_A by $f_B(v) = f(v) - |\nabla(v, B)|$, $g_A(v) = g(v) - |\nabla(v, A)|$. Note that $f_B + g_A = \deg_{G[C \cup D]}$.

10.2.18. THEOREM. *Let G be any graph and f and g, non-negative integer-valued functions defined on $V(G)$ such that $f + g \geq \deg_G$. Let A, B, C and D be as defined throughout this section and let $D_1, \ldots, D_t$ be the connected components of $G[D]$. Then every D_i is (f_B, g_A)-critical and $G[C]$ has an (f_B, g_A)-factor. Furthermore, we have*

$$\delta(f,g) = t + |\nabla(A,B)| - f(A) - g(B).$$

PROOF. We may assume that $f(v) = 0$ for $v \in A$ and $g(v) = 0$ for $v \in B$, since otherwise we may consider f' and g' as previously defined instead of f and g, by Lemma 10.2.15.

Let H be any (f,g)-optimal subgraph. We know by Theorem 10.2.13 that H contains no line connecting A to C or spanned by A, but it contains all lines connecting B to C or spanned by B. Since the deletion of lines incident with A is a local augmentation by the hypothesis that $f(v) = 0$ for $v \in A$, we may assume that H contains no line incident with A. Similarly, we may assume that H contains all lines between B and D. A simple computation then shows that

$$\delta(H; f,g) = \delta(H[D]; f_B, g_A) + |\nabla(A,B)|. \tag{10.2.8}$$

Conversely, let J_0 be any (f_B, g_A)-optimal subgraph of $G[D]$, and let

$$J = J_0 \cup G[B] \cup \nabla(B,D) \cup \nabla(B,C) \cup H[C],$$

where H is any (f,g)-optimal subgraph. Then the same computation again shows that

$$\delta(J; f,g) = \delta(J_0; f,g) + |\nabla(A,B)|. \tag{10.2.9}$$

From equations (10.2.8) and (10.2.9) we see that

$$\delta(f,g) = \delta(f_B, g_A) + |\nabla(A,B)|. \tag{10.2.10}$$

It follows from (10.2.8) and (10.2.10) that if H is an (f,g)-optimal subgraph, then $H[D]$ is an (f_B, g_A)-optimal subgraph.

Next we show that every $v \in D$ belongs to $D(G[D]; f_B, g_A)$. To this end, let H be any (f,g)-optimal subgraph such that $\deg_H(v) > f(v)$. Add to H all lines of $\nabla(B,D)$. These are local augmentations, so we get another (f,g)-optimal subgraph H' and clearly we still have $\deg_{H'}(v) > f(v)$. Delete all lines incident with A from H'. These are again local augmentations, and since the "deficiency" decreases at the points of A, it must increase at the other endpoints of these lines. Hence $\deg_{H''}(v) > f(v)$. But then

$$\deg_{H''[D]}(v) = \deg_{H''}(v) - |\nabla(v,B)| > f(v) - |\nabla(v,B)| = f_B(v).$$

Furthermore, $H''[D]$ is (f_B, g_A)-optimal by the above observation. Similarly, we can construct an (f_B, g_A)-optimal subgraph in which the degree of v is less than $f(v)$. This proves that $D(G[D]; f_B, g_A) = D$. Consequently, every connected component of $G[D]$ is (f_B, g_A)-critical. Thus $\delta(f_B, g_A) = t$, and using equation (10.2.10) the proof of the theorem is complete. ∎

Thus we have proved a result generalizing the Gallai-Edmonds Structure Theorem. Just as in the case of perfect matchings, this theorem has many consequences and applications. In this book we do not treat all of these, but some of them are formulated as exercises. The single most important consequence of this structure theorem is the following minimax theorem characterizing $\delta(f,g)$ (Lovász (1970b)), which implies a necessary and sufficient condition for the existence of an f-factor due to Tutte (1952).

To formulate this result, we introduce some notation. Let G be a graph, let f and g be non-negative integer-valued functions defined on $V(G)$, and let X and Y be disjoint subsets of $V(G)$. We say that a connected component K of $G - X - Y$ is **(X,Y)-odd** (with respect to (f,g)), if (a) $f(v) = \hat{g}(v)$ for all $v \in V(K)$, and (b) $f(V(K)) + |\nabla(V(K), Y)|$ is odd.

The reader may verify without difficulty that K is (X,Y)-odd with respect to (f,g) if and only if it is (Y,X)-odd with respect to (g,f). We denote by $\tau(X,Y)$ the number of (X,Y)-odd components of $G-X-Y$. Note that every connected component of $D[G]$ is (A,B)-odd. Hence $\tau(A,B)$ is the number of connected components of $D[G]$.

10.2.19. THEOREM. *Let G be a graph and let f and g be two non-negative integer-valued functions on $V(G)$ such that $f + g \geq \deg_G$. Then*

$$\delta(f,g) = \max\{\tau(X,Y) + |\nabla(X,Y)| - f(X) - g(Y)\},$$

where (X,Y) ranges over all pairs of disjoint subsets of $V(G)$.

PROOF. Consider any spanning subgraph H and any two disjoint subsets X, Y of $V(G)$. We show that

$$\delta(H; f,g) \geq \tau(X,Y) + |\nabla(X,Y)| - f(X) - g(Y). \qquad (10.2.11)$$

Setting $Z = V(G) - X - Y$, we then have

$$\delta(H;f,g) \geq \sum_{v\in X}(\deg_H(v) - f(v)) + \sum_{v\in Y}(\deg_{\overline{H}}(v) - g(v)) + \sum_{v\in Z}\delta(v;H;f,g).$$

If K is any (X,Y)-odd component of $G-X-Y$, then either K contains a point v with $\delta(v;H;f,g) \neq 0$, or H has a line connecting K to X or $\overline{H}$ has a line connecting K to Y, by a simple parity argument. Hence

$$\sum_{v\in Z}\delta(v;H;f,g) \geq \tau(X,Y) - |\nabla_H(Z,X)| - |\nabla_{\overline{H}}(Z,Y)|,$$

and so

$$\delta(H;f,g) \geq \deg_{\overline{H}}(X) + \deg_H(Y) - f(X) - g(Y) + \tau(X,Y) - |\nabla_H(Z,X)| - |\nabla_{\overline{H}}(Z,Y)|.$$

But clearly

$$\deg_H(X) + \deg_{\overline{H}}(Y) \geq |\nabla_G(X,Y)| + |\nabla_H(Z,X)| + |\nabla_{\overline{H}}(Z,Y)|,$$

and hence inequality (10.2.11) follows. Since by Theorem 10.2.18 any (f,g)-optimal subgraph H, together with $X = A$ and $Y = B$, gives equality in (10.2.11), the theorem is proved. ■

Taking $g = \hat{f}$ in Theorem 10.2.19, we obtain the following important theorem due to Tutte (1952).

10.2.20. THEOREM. (*The f-Factor Theorem*). *A graph G has an f-factor if and only if for every two disjoint subsets X and Y of $V(G)$, the number of those connected components K of $G-X-Y$ for which $f(V(K))+|\nabla(V(K),Y)|$ is odd, does not exceed $f(X)+\hat{f}(Y)-|\nabla(X,Y)|$.* ■

From Theorems 10.2.19 and 10.2.20 several sufficient conditions for the existence of (f,g)-factors can be derived as special cases. Some of these are stated below as exercises.

10.2.21. EXERCISE. Use Theorem 10.2.5 to derive Gallai's Identity 1.0.2.

10.2.22. EXERCISE. Let G be a factor-critical graph. Construct a new graph M whose points are the near-perfect matchings of G, any two of which are adjacent if and only if they have all but one of their lines in common. Prove that M is connected.

10.2.23. EXERCISE. Show that Theorem 10.2.10 contains Lemma 3.1.5 as a special case.

10.2.24. EXERCISE. Prove that every factor-critical graph is also $(1, \deg_G -1)$-critical.

10.2.25. EXERCISE. Prove that a bipartite graph is never (f, g)-critical. Derive the necessary and sufficient condition for the existence of an f-factor in a bipartite graph (Theorem 2.4.2) from the f-factor Theorem.

10.2.26. EXERCISE. Let G be a graph and let a and b be integers such that $a + b \geq \Delta(G) + 1$. Prove that G has an (a, b)-factor.

10.2.27. EXERCISE. Let G be a graph such that any two odd cycles in G either have a point in common or are connected by a line. Prove that G has an f-factor if and only if $\sum_v f(v)$ is even and $|\nabla(X, Y)| \leq f(X)+f(Y)$ holds for any two disjoint subsets X, Y of $V(G)$.

10.2.28. EXERCISE. Suppose that $f + g > \deg_G$. Prove that G has an (f, g)-factor if and only if $|\nabla(X, Y)| \leq f(X) + g(Y)$ holds for any two disjoint subsets X, Y of $V(G)$.

10.2.29. EXERCISE. Deduce from the f-factor Theorem the following results of Petersen (1891): If G is a $(2d)$-regular graph then G has a 2-factor. If, in addition, G is connected and $|E(G)|$ is even, then G has a d-factor.

10.3. Realization of Degree Sequences

When is a given sequence of integers realizable as the degree sequence of some graph? It is not difficult to see that this question is a special case of the f-factor problem. In fact, it is just the f-factor problem for complete graphs. However, more transparent conditions can be formulated in this special case. Furthermore, solutions are known for the realization problem where the realizing graph is required to have certain additional properties (like being k-connected, or having a perfect matching). Such versions of the general f-factor problem are not well-solved; in fact they are often NP-complete.

We give only a sampling of realizability results; more examples can be found, for example, in Berge (1973). We shall restrict ourselves to realizability by simple graphs. (Realizability problems tend to be simpler if multiple lines are allowed. Some results of this type occur in the exercises.)

The principal result concerning realizability is due to Erdős and Gallai (1960).

10.3.1. THEOREM. *Let $d_1 \geq d_2 \geq \cdots \geq d_n \geq 0$ be integers. Then there exists a simple graph with degrees $d_1, \ldots, d_n$ if and only if*

(a) $\sum_{i=1}^{n} d_i$ *is even, and*

(b) $\sum_{i=1}^{k} d_i \leq \sum_{i=k+1}^{n} \min\{d_i, k\} + k(k-1)$

holds for each $1 \leq k \leq n$.

PROOF. The necessity of (a) is trivial. The necessity of (b) follows by counting the lines incident with the first k points in two different ways.

We prove the sufficiency of conditions (a) and (b) using Tutte's f-factor Theorem 10.2.20. Of course this is an example of "overkill", since the sufficiency could be proved by a fairly simple induction argument, as will be sketched later. But we want to use this proof also to illustrate the power of the f-factor Theorem.

So suppose now that (a) and (b) hold. Let K_n denote the complete graph on points $\{1, \ldots, n\}$. Then a graph with degrees $d_1, \ldots, d_n$ exists if and only if K_n has a d-factor (where, to be precise, d denotes the function defined by $d(i) = d_i$ for $1 \leq i \leq k$). By the f-factor Theorem it suffices to show that for every $X, Y \subseteq V(G)$, such that $X \cap Y = \emptyset$, we have

$$\begin{aligned} \tau(X,Y) &\leq d(X) + d(Y) - |\nabla(X,Y)| \\ &= d(X) - d(Y) + (n-1)|Y| - |X| \cdot |Y|, \end{aligned} \tag{10.3.1}$$

where $d(X) = \sum d_i$ and this sum is over all points in set X.

If $X = Y = \emptyset$, then the right hand side of inequality (10.3.1) is 0, but so is the left hand side by condition (a). Consider now the case when $X \cup Y \neq \emptyset$. Note that $\tau(X,Y) \leq 1$ since K_n is a complete graph. Furthermore, if we fix $|X| = r$ and $|Y| = k$ then the minimum of the right hand side is attained when X is the set of those r points with smallest prescribed degrees and Y is the set of those k points with largest prescribed degrees. Thus it suffices to prove that

$$\epsilon + \sum_{i=1}^{k} d_i \leq \sum_{i=n-r+1}^{n} d_i + (n-1-r)k, \tag{10.3.2}$$

where $\epsilon = 1$, if $\sum_{i=k+1}^{n-r} d_i + (n-r-k)k$ is odd and $\epsilon = 0$, otherwise.

To verify this inequality, let us write it in the form

$$\epsilon + \sum_{i=1}^{k} d_i \le \sum_{i=n-r+1}^{n} d_i + \sum_{i=k+1}^{n-r} k + k(k-1).$$

This inequality follows from (b) immediately unless $\epsilon = 1$, and $d_{n-r} \ge k$, but $d_{n-k+1} \le k$. However, in this latter case we must have

$$\sum_{i=1}^{k} d_i = \sum_{i=n-r+1}^{n} d_i + (n-1-r)k,$$

and hence

$$\sum_{i=k+1}^{n-r} d_i + (n-r-k)k = \sum_{i=1}^{n} d_i - 2\sum_{i=1}^{k} d_i + 2(n-r)k - k(k+1),$$

which is even and so $\epsilon = 0$. This contradiction proves the theorem. ∎

This theorem provides an easy procedure to check whether or not a sequence is the degree sequence of a simple graph. But there is an even simpler algorithm to do this due to V. Havel (1955) and to Hakimi (1962). It is based on the following lemma.

10.3.2. LEMMA. *Let $d_1 \ge d_2 \ge \cdots \ge d_n \ge 0$ be integers. Then this sequence is realizable as the degrees of a simple graph if and only if the numbers $d_1 - 1, \ldots, d_{d_n} - 1, d_{d_n+1}, \ldots, d_{n-1}$ are realizable.*

PROOF. The sufficiency of the condition is clear: if $d_1 - 1, \ldots, d_{d_n} - 1, d_{d_n+1}, \ldots, d_{n-1}$ are the degrees of a simple graph G', then introducing a new point and connecting it to the first d_n points of G' we obtain a simple graph with degrees $d_1, \ldots, d_n$.

To prove the necessity, it suffices to show that if $d_1, \ldots, d_n$ is the degree sequence of some simple graph, then it is, in fact, the degree sequence of one in which the point n is adjacent to the points $1, \ldots, d_n$. To show this, let G be a simple graph with degrees $d_1, \ldots, d_n$. Furthermore, suppose that there are points i and j in G such that $1 \le i < j \le n-1$, and n is adjacent to j but not to i. Since $d_j \le d_i$, there must be a point $k \ne i, j, n$ which is adjacent to i but not to j. Replacing the lines ki and nj by the lines ni and kj, we obtain another simple graph with the same degrees, but such that in the neighborhood of n, point j is replaced by i. Repeating this procedure if necessary, we end up with a graph with the given degrees in which the point n is adjacent to the first d_n points. This completes the proof of the lemma. ∎

The preceding lemma provides a very simple way to check for the realizability of a sequence as the degree sequence of a simple graph. In the case when the answer is affirmative, it also constructs the graph. It could also be used to obtain a direct proof of Theorem 10.3.1 by induction. The details of this are left to the reader as an exercise.

As we remarked at the beginning of this section, we may ask if the given degree sequence is realizable by a graph with certain special properties. Here we prove one result of this nature; further results of this kind are included as exercises.

The following theorem was conjectured by Grünbaum (1970a) and proved independently by Kundu (1973) and Lovász (1974).

10.3.3. THEOREM. *Let $d_1, d_2, \ldots, d_n$ be integers. Then $d_1, \ldots, d_n$ are the degrees of a simple graph with a perfect matching if and only if n is even and both sequences $d_1, \ldots, d_n$ and $d_1 - 1, \ldots, d_n - 1$ are degree sequences of simple graphs.*

PROOF. The necessity of the condition is trivial. Suppose, on the other hand, that $d_1, \ldots, d_n$ is the degree sequence of a simple graph G on points $\{1, \ldots, n\}$ and $d_1 - 1, \ldots, d_n - 1$ is the degree sequence of another simple graph G' on the same set of points. We choose G and G' so that the symmetric difference of $E(G)$ and $E(G')$ is minimal, and show that $G' \subseteq G$. This will clearly imply the theorem.

Suppose that $G' \nsubseteq G$, and let v be the point of $V(G)$ which is incident with the largest number r of lines of G' not in G. Obviously, there are $r+1$ lines in $E(G) - E(G')$ incident with v. Choose any $z \in V(G)$ such that $zv \in E(G') - E(G)$, and then any $w \in V(G)$ such that $zw \in E(G) - E(G')$. (Such a w exists since $\deg_G = \deg_{G'} + 1$).

We claim that for any $y \in V(G) - v - w$ for which $vy \in E(G) - E(G')$, we also have $wy \in E(G) - E(G')$. For, suppose first that $wy \notin E(G)$. Then replace the lines vy and wz with the lines vz and wy in G. This results in a graph with the same degrees, but with smaller symmetric difference with respect to G', contrary to the minimality assumption. Second, assume $wy \in E(G')$. Then replacing the lines vz and wy by the lines vy and wz in G', we get a contradiction as before.

But then we see that the number of lines in $E(G) - E(G')$ incident with w is strictly larger than the number of lines in $E(G) - E(G')$ incident with v. In fact, for every line $yv \in E(G) - E(G')$, we have the line $yw \in E(G) - E(G')$ $(y \neq v, w)$. Moreover, in addition, we also have the line wz. (The line vw may or may not be a line in $E(G) - E(G')$, but it contributes the same number to $\deg_G(v)$ and $\deg_G(w)$ in any case.) This is a contradiction of the choice of v. ■

The above proof also provides a simple algorithm for constructing a graph with a perfect matching and having given degrees.

10.3.4. EXERCISE. Prove that integers $d_1 \geq d_2 \geq \cdots \geq d_n \geq 0$ are the degrees of a graph with multiple lines, but no loops, if and only if $\sum d_i$ is even and $d_1 \leq d_2 + \cdots + d_n$.

10.3.5. EXERCISE. Prove that integers $d_1 \geq \cdots \geq d_n \geq 0$ are the degrees of a tree if and only if $d_n > 0$ and $d_1 + \cdots + d_n = 2n - 2$.

10.3.6. EXERCISE. Prove that integers $d_1 \geq \cdots \geq d_n \geq 0$ are the degrees of a connected simple graph if and only if they are the degrees of a simple graph, $d_n > 0$ and $d_1 + \cdots + d_n \geq 2n - 2$.

10.3.7. EXERCISE. Show that the problem as to whether or not a graph G has a *connected* f-factor is NP-complete.

10.3.8. EXERCISE. Prove that if $f, g : S \to Z_+$ are two functions defined on the same finite set S, and g takes its values from a set $\{k, k+1\}$ for some k, then there exists a simple graph on S with degrees f having a g-factor if and only if all three of f, g and $f - g$ are degree sequences of simple graphs (Kundu (1973)).

11

Matroid Matching

11.0. Introduction

We have remarked before that matching theory is a central part of graph theory, not only because of its applications, but also because it is the source of many important ideas developed during the rapid growth of combinatorics during the last several decades. Matching theory serves as an archetypal example of how a "well-solvable" problem can be studied. It also seems to be very near the limits of well-solvability in a sense. Most reasonable generalizations of the matching problem have turned out to be difficult — usually NP-hard — problems.

This is not quite the case with the *bipartite* matching problem, which has several non-trivial, but well-solved, generalizations. Besides the non-bipartite matching problem, we have also discussed the Matroid Intersection Problem along these lines. (See Section 1.3.) Is there a common extension of these two generalizations? Such a problem was proposed by Edmonds under the name "Matchoid Problem", and by Lawler in a different — but equivalent — form under the name "Matroid Parity Problem". There are other equivalent versions of the same problem which we shall call collectively the "Matroid Matching Problem". Unfortunately, this problem turns out to have no polynomial-time algorithmic solution. (Interestingly, this negative result does not depend on the "P$\neq$NP" hypothesis.) Also no good characterization of the "matroid matching number" can be found in a precisely defined sense. But in Section 11.2 we shall prove a result about the matching number of matroids which will imply good characterizations of this number, if the matroids in question are presented in certain special ways. In particular, we shall solve the problem for *linear* matroids. Several applications of matroid matching results will be discussed in Section 11.3.

11.1. Formulations of the Matroid Matching Problem

Recall that if S is a finite set and if f is a real-valued function defined on the subsets of S, then f is called **submodular** if the inequality

$$f(X \cap Y) + f(X \cup Y) \leq f(X) + f(Y)$$

holds for every pair of sets $X, Y \subseteq S$. A submodular, integer-valued, monotone increasing set function which is 0 on the empty set is called a **polymatroid function**. A set S endowed with a polymatroid function f is called a **polymatroid** and is denoted by (S, f). The function f will be called the **rank function** of the polymatroid. A k-**polymatroid** is a polymatroid in which every singleton set has rank at most k. Thus 1-polymatroids are just matroids.

Let $r_1, \ldots, r_k$ be the rank functions of k arbitrary matroids on the same set S. Then $(S, r_1 + \cdots + r_k)$ is a k-polymatroid. However, not all k-polymatroids arise in this way. There is another construction to obtain polymatroids from matroids which will prove more general. Given a matroid (S, r), we can construct a polymatroid as follows. Let $\{A_i \mid i \in T\}$ be a set of subsets of S. Define a set function f on the subsets X of T by the formula

$$f(X) = r\Big(\bigcup_{i \in X} A_i\Big). \tag{11.1.1}$$

It is a straightforward matter to verify that this set function is a polymatroid function and so (T, f) is a polymatroid.

It can be shown (Helgason (1974), McDiarmid (1975) and Lovász (1977)) that, conversely, every polymatroid can be represented by this construction. Thus, polymatroids do not really lead to problems fundamentally more general than do matroids. However, often the polymatroid formulation is more compact and therefore easier to handle.

A special case of the above construction occurs when the matroid (S, r) is a linear space. We then obtain the following construction of polymatroids. Let T be a set of subspaces of a linear space. For $X \subseteq T$, let $f(X)$ denote the dimension of the linear span of the union of subspaces in X. Then (T, f) is a polymatroid. Polymatroids arising by this construction are called **linear**. Let us note that if the members of T are 2-dimensional subspaces, then we obtain a 2-polymatroid and it will be these 2-polymatroids for which the Matroid Matching Problem will be solved in this chapter.

When we speak about a *linear* polymatroid we shall tacitly assume that it is given to us as a set of linear subspaces of a linear space over some field, where each subspace is specified by one of its bases. Thus, in fact, what we are given is a matrix whose columns are partitioned into classes. The linear subspaces spanned by the columns in these classes will then define our linear polymatroid.

Another special case of the general construction described above arises when the underlying matroid is *free*. In this case the structure we start with is a hypergraph; that is, a collection $\{A_i \mid i \in T\}$ of subsets of a finite set S, where we define a polymatroid function f on T by

$$f(X) = |\bigcup_{i \in X} A_i|. \tag{11.1.2}$$

In the special case when the given hypergraph is a graph, the polymatroid defined this way will be a 2-polymatroid. This is the most straightforward way to apply results on 2-polymatroids to graphs. For further constructions leading to polymatroids, the reader is referred to Lovász (1977, 1983a).

What does all this have to do with matchings? Let (S, f) be a 2-polymatroid. A set $X \subseteq S$ is called a **matching** if $f(X) = 2|X|$. In the special case of 2-polymatroids defined on the set of lines of a graph as in (11.1.2), the matchings in the 2-polymatroid will be just the matchings of the graph. More generally, let G be a graph and $b : V(G) \to Z_+$ any mapping. Define a function $f : 2^{E(G)} \to Z_+$ by

$$f(X) = \sum_{v \in V(G)} \min\{b(v), \deg_X(v)\}.$$

Then $(E(G), f)$ is a 2-polymatroid and a matching in $(E(G), f)$ is just a b-matching in G. (See Section 6.1.)

As a further example, consider two matroids (S, r_1) and (S, r_2) on the same underlying set. As previously remarked, $(S, r_1 + r_2)$ is a 2-polymatroid. The matchings of $(S, r_1 + r_2)$ are just the common independent sets in (S, r_1) and (S, r_2).

It is easy to see that every subset of a matching is itself a matching. The maximum size of a matching in (S, f) will be denoted by $\nu(S, f)$.

Let X be a subset of the polymatroid (S, f). We define the **span** of X, Span$X$, to be the largest subset of S which includes X and has the same rank as X. (It follows from the submodularity of f that every set has a *unique* span.) We say that X is **spanning** if its span is the whole set S, that is, if $f(X) = f(S)$. We denote by $\rho(S, f)$ the minimum cardinality of a spanning set.

There are a number of simple facts concerning matchings in graphs which can be generalized in a rather straightforward way to matchings in 2-polymatroids. We shall need the following fact, which generalizes Gallai's Identity (1.0.2). (See Lovász (1981).)

11.1.1. LEMMA. *For any 2-polymatroid (S,f), $\nu(S,f)+\rho(S,f)=f(S)$.*

PROOF. Let M be a maximum matching in (S,f). Consider the submodular set function $f_M(X)=f(M\cup X)-f(M)$, defined on the subsets X of $S-M$. It is clear that f_M is a polymatroid function, and for every $x\in S-M$,

$$f_M(x)=f(M\cup\{x\})-f(M)\leq 1,$$

since otherwise $M\cup\{x\}$ would be a matching, contradicting the maximality of M. Hence f_M is a matroid rank function. Now let B be a basis for the matroid $(S-M,f_M)$. Then

$$f(M\cup B)=f(M)+f_M(B)=f(M)+f_M(S-M)=f(S),$$

and hence $M\cup B$ is a spanning set. Furthermore, $|M\cup B|=|M|+|B|=|M|+f_M(S-M)=|M|+f(S)-f(M)=f(S)-|M|=f(S)-\nu(S,f)$. This proves that $\rho(S,f)\leq|M\cup B|=f(S)-\nu(S,f)$. To show the reverse inequality, let T be any spanning set. Delete elements from T as long as each deletion lowers the rank by at most 1. Suppose we are forced to stop after k deletions with a set N. Then, deleting any element from N lowers its rank by at least 2. It is easy to see that this implies that N is a matching. So

$$\begin{aligned}|N|&=|T|-k\\&\leq|T|-(f(T)-f(N))\\&=|T|-f(T)+2|N|,\end{aligned}$$

and hence $|N|\geq f(T)-|T|$. But then

$$\nu(S,f)\geq|N|\geq f(T)-|T|=f(S)-\rho(S,f).$$

This proves the lemma. ∎

Note that since S is spanning, this lemma implies the inequality

$$\nu(S,f)\geq f(S)-|S|. \tag{11.1.3}$$

We have given the proof of this lemma to illustrate the idea that graph-theoretic facts can sometimes be generalized to 2-polymatroids, and to illustrate how we may actually work with polymatroids. Some similar arguments in the material to follow will be left to the reader.

We can now state the main problem of this section.

THE POLYMATROID MATCHING PROBLEM.
Given a 2-polymatroid (S,f), determine the maximum size of a matching.

REMARK 1. Much of the problem is hidden in the slick phrase "given a 2-polymatroid"! How is a 2-polymatroid "given"? It cannot mean that the value of f is given explicitly for each subset separately, since then it would take more than $2^{|S|}$ units of time just to read the data, and to look at each subset to find a maximum matching would double this time! The usual way to circumvent this difficulty is to assume that there is a subroutine (or "oracle") to compute $f(X)$ for any given subset $X \subseteq S$, and that the running time of this subroutine is polynomial in $|S|$. If our polymatroid is linear, or if it is determined by a graph as in equation (11.1.2), then such a subroutine is easily constructed. (See Box 11A for more remarks on such "oracle algorithms".)

REMARK 2. It would be easy to extend the notion of a matching to k-polymatroids, and ask for the maximum size of a matching in a k-polymatroid. It turns out, however, that even for very nice polymatroids, say for those defined by hypergraphs via equation (11.1.2), the matching problem is NP-hard. (See Karp (1972).)

REMARK 3. In view of the approach used mainly in Chapter 7, it would be natural to try to obtain a description of the convex hull of matchings in a (say, linear) 2-polymatroid. Equivalently, we would like to assign weights to the elements of S and then find a matching having maximum weight. This problem, however, is completely unsettled.

BOX 11A. Oracles

How can we handle the Polymatroid Matching Problem from an algorithmic point of view? How is a polymatroid "given"? Answers to such questions are not at all obvious, and the answers we give influence the direction we should take toward improving our algorithms.

These questions already arise in dealing with much simpler problems. Suppose that we are given a matroid with weights on its points and we have to find an independent set having maximum weight. But precisely how is this matroid "given"? If it is given as a list of its independent sets (which is perhaps the most natural way if we look at the definition of a matroid), then the problem is trivial; just scan the list of independent subsets, compute the weight of each and pick a largest one. Essentially, this takes no more time than merely reading the input data. But we saw earlier that a maximum weight independent set can also be found by the Greedy Algorithm. What then was the point to this latter result?

A way to make the maximum weight independent set problem more meaningful by excluding the trivial solution formulated above is to assume that the matroid is not really fully given in advance, but only in the following sense. Suppose we are given the set S of its points (together with their weights), and an "oracle" or "subroutine" to check whether or not a given subset X of S is independent. We shall not be concerned with the details of this subroutine; therefore an oracle is often described as a little black box, the function of which has certain specifications: its input is restricted to subsets of a given set S, its output is "yes" or "no", and it is guaranteed by the manufacturer that the sets to which it answers "yes" form the independent sets of a fixed matroid on S. We shall also agree that one call on the oracle counts only as one step in the algorithm with which we are dealing. Of course in practice, we substitute some realization of the oracle for the black box and then we also have to take into account the running time of this subroutine. But as long as we find polynomial time realizations, the polynomiality of the total running time will not be affected.

In the model for matroid algorithms sketched above, the Greedy Algorithm for computing the maximum weight independent set can be implemented in $n \log n$ time, where n is the number of points in the matroid. More difficult problems, such as that of finding a largest common independent set in two matroids, can also be solved in time polynomial in the number of elements in the two matroids.

The oracle model for matroid algorithms is not only a necessary burden to make certain procedures precise; it also allows us to obtain non-trivial lower bounds on the computational complexity of certain problems. It is surprising that by using this oracle model, we can prove the polynomial time unsolvability of certain problems, whereas previously all such negative results depended on the hypothesis that P$\neq$NP.

To illustrate these remarks, we now show that the 2-polymatroid matching problem cannot be solved in polynomial time. This result was obtained independently by Jensen and Korte (1982) and Lovász (1981). A thorough study of matroid oracles was made by Hausmann and Korte (1981).

Suppose a 2-polymatroid (S, f) is given by an oracle which accepts as input any subset of S and outputs the value of f on this subset. Let us say that the oracle is realized by that famous Delphic seer, Pythia, and our favorite Greek hero, Algorithmos, thinks he has a polynomial-time algorithm to find the matching number of any 2-polymatroid. The difference between Merlin (see Box 1A) and Pythia is that she never lies and so Algorithmos does not have to require a proof of the truth of her answer. (Of course we do know that her favorite pastime is to lure mortal Greek heroes into deep trouble by telling the truth!)

So let us suppose that Algorithmos is busily performing his computations, every now and then asking Pythia about the value of $f(X)$ for certain sets X. Pythia pretends that she has in mind the following 2-polymatroid:

$$f(X) = \begin{cases} 2|X|, & \text{if } |X| < k, \\ 2k-1, & \text{if } |X| = k, \text{ and} \\ 2k, & \text{if } |X| > k. \end{cases}$$

It is clear that this 2-polymatroid has matching number $k-1$. But how fast can Algorithmos arrive at this conclusion? Suppose that he announces that the matching number of this 2-polymatroid is $k-1$, without first asking for all values of $f(X)$ when $|X| = k$. Suppose X_0 is a k-element subset of S for which he has not checked $f(X_0)$. Then Pythia (or fate!) triumphs over him by claiming that the 2-polymatroid she had in mind was in fact the following:

$$f(X) = \begin{cases} 2|X|, & \text{if } |X| < k, \\ 2k-1, & \text{if } |X| = k, X \neq X_0, \text{ and} \\ 2k, & \text{if } X = X_0 \text{ or } |X| > k. \end{cases}$$

The matching number of this 2-polymatroid is k, and Algorithmos will suffer all the consequences known from mythology for misinterpreting Pythia!

This argument shows that to decide whether or not $\nu(S,f) = k$ takes — in the worst case — at least $\binom{|S|}{k}$ calls on the oracle. Choosing $|S| = 2k$, this number of calls is exponentially large in $|S|$, and this proves that no polynomial-time algorithm exists for finding the matching number of a 2-polymatroid given by an oracle which computes $f(X)$.

One might think that it is the oracle model which is responsible for this negative result and whenever we have a "decent" (i.e., polynomial-time) realization of the oracle, we can always design a polynomial-time algorithm to compute the matching number. But this is not so. Let G be a graph, suppose $k \geq 1$ and consider the following 2-polymatroid on $V(G)$:

$$f(X) = \begin{cases} 2|X|, & \text{if } |X| < k \\ 2k-1, & \text{if } |X| = k \text{ and } X \text{ is not a clique in } G, \\ 2k, & \text{if } |X| = k \text{ and } X \text{ is a clique in } G, \text{ or if } |X| > k. \end{cases}$$

Then for every $X \subseteq V(G)$, $f(X)$ is easily computed. But $\nu(V(G),f) = k$ if and only if G contains a clique of size k and so to determine $\nu(V(G),f)$ is NP-hard.

Let us now formulate some equivalent versions of the Polymatroid Matching Problem.

THE MATROID PARITY PROBLEM.

Let (S,r) be matroid and let $S = A_1 \cup \cdots \cup A_m$ be a partition of S into pairs. Find a maximum size collection $\{A_{i_1}, \ldots, A_{i_k}\}$ of these pairs such that $A_{i_1} \cup \cdots \cup A_{i_k}$ is independent in (S,r).

This problem can be reduced to the polymatroid matching problem by simply considering the polymatroid induced on the set $\{1,\ldots,m\}$ by the function (11.1.1). Conversely, the polymatroid matching problem can also be reduced to the matroid parity problem, even though this reduction involves as a subroutine the problem of minimizing a submodular set function (see Box 11B).

THE MATCHOID PROBLEM.
Let G be a graph and assume that for each point $v \in V(G)$, a matroid $\mathcal{M}_v$ is given on the star of v. Call a set M of lines a **matchoid** *if the set of lines in M incident with any point v is independent in $\mathcal{M}_v$. Find a matchoid of maximum size.*

It is rather straightforward to verify that the matchoid and matroid parity problems are equivalent.

A slightly different formulation of this same problem can be stated as follows. A graph together with a matroid structure on its points is called a **matroid graph**. A set of lines in a matroid graph is called a **matching** if it is a matching in the underlying graph and, in addition, the set of points it covers is independent in the matroid.

THE MATCHING PROBLEM FOR MATROID GRAPHS.
Find a maximum matching in a matroid graph.

The equivalence of the next problem to those above is less obvious.

THE MAXIMIZATION OF SMOOTH SUBMODULAR FUNCTIONS.
Given a submodular set function f which satisfies the "smoothness" condition: $|f(X)-f(Y)| \leq |X \oplus Y|$ for any two subsets X and Y, maximize f.

To show that this problem also reduces to polymatroid matching, let f be a smooth submodular set function on S and define f_1 by:

$$f_1(X) = f(X) + |X| - f(\emptyset).$$

It is then easy to verify that f_1 is a 2-polymatroid function (by the smoothness of f), and that

$$\max_{X \subseteq S} f(X) = f(\emptyset) + \nu(S, f_1).$$

We remark that the maximization problem of a general submodular function is NP-hard even for some very nice submodular functions. For example, the problem of maximizing the submodular set function $|X| - |E(G[X])|$ over all $X \subseteq V(G)$, where G is a graph, is equivalent to finding a maximum independent set of points in G!

Finally, in view of Lemma 11.1.1, the polymatroid matching problem is also equivalent to the following.

THE MINIMUM SPANNING SET PROBLEM IN 2-POLYMATROIDS. *Given a 2-polymatroid (S, f), find a minimum spanning subset.*

BOX 11B. Minimizing Submodular Set Functions

Let S be a finite set and f: $2^S \to \Re$, a submodular set function (see Section 1.3). The task of finding a minimum for f is a very general combinatorial optimization problem, which includes quite a few of the combinatorial optimization problems discussed in this book. Let us look at some examples.

11B.1. EXAMPLE. Let (D, c, s, t) be a network. That is, let D be a directed graph, $c : E(G) \to \Re_+$, a capacity function, and s and t specified points called the "source" and "sink" respectively. For each subset $X \subseteq V(G) - s - t$, let $f(X) = c(X \cup s)$ denote the capacity of the $s - t$ cut determined by $X \cup s$. Then f is a submodular set function and the problem of minimizing f is equivalent to finding a minimum $s - t$ cut.

11B.2. EXAMPLE. Let (S, r_1) and (S, r_2) be two matroids on the same underlying set. Then, by Theorem 1.3.19, the maximum size of a common independent set is equal to the minimum of $r_1(X) + r_2(S - X)$ taken over all subsets X of S. Now this sum, as a function of X, is submodular and minimizing it is equivalent to finding the maximum size of a common independent set of two matroids.

11B.3. EXAMPLE. Suppose G is a bipartite graph with bipartition (A, B) and for all $X \subseteq A$, let

$$f(X) = |\Gamma(X)| - |X|.$$

Then $f(X)$ is submodular and minimizing it is equivalent to finding the deficiency of G. In particular, G has a matching from A into B if and only if the minimum of f is 0.

A polynomial-time algorithm to minimize a general submodular set function is known (see Grötschel, Lovász and Schrijver (1981)). We will not go into the details of this algorithm, but let us at least remark that it uses the Ellipsoid Method in a way quite similar to the solution of the matching problem sketched above. This reduces the problem of minimizing a submodular set function f to a linear program of the following type:

$$\begin{array}{lrcll} \text{maximize} & w \cdot x & & & \\ \text{subject to} & x & \geq & \mathbf{0} & \\ & x(T) & \leq & f(T) & \quad (\text{for all } T \subseteq S). \end{array}$$

A polytope of this kind is often called a **polymatroid polytope** or briefiy, a **polymatroid**. To optimize a linear objective function over such a polytope is quite easy using the Greedy Algorithm (see Edmonds (1970)).

Unfortunately, this approach to submodular function minimization is impractical since it uses the Ellipsoid Method. Another disadvantage is that it does not give any combinatorial insight into the structure of submodular set functions. To find an algorithm to minimize a submodular set function based on combinatorial ideas is an important open problem.

Often one would like to minimize a submodular set function over some special family of subsets of S (rather than over *all* subsets). Let us look at some more examples.

11B.4. EXAMPLE. To determine the line connectivity of a graph, we must find the minimum of the submodular set function $|\nabla(X)|$ over the *non-empty proper subsets* of $V(G)$.

11B.5. EXAMPLE. Let G be a graph and suppose $x \in \Re_+^{E(G)}$. To determine whether or not $x \in PM(G)$, we must find the minimum of the submodular set function $x(\nabla(X))$ over the *odd subsets* of $V(G)$.

To find the minimum of a submodular set function over the non-empty proper (or non-empty, or proper) subsets of a set is easily reduced to the unconstrained case, and this reduction is left to the reader as an exercise.

It is also possible to minimize a submodular set function, defined on *all* subsets of a set S, over just the *odd* subsets of S. This is more difficult, and a rather involved extension of the ideas of the Padberg and Rao method for finding a minimum T-cut (see Theorem 6.6.11) is needed. For details we refer the reader to Grötschel, Lovász and Schrijver (1984b).

What does all this have to do with matching algorithms? We saw in Section 9.4 that the Ellipsoid Method reduces the Maximum Weight Matching Problem to the problem of finding the minimum of a submodular set function over the odd subsets of $V(G)$ (cf. Example 11B.5). This problem, in turn, can be reduced to the unconstrained minimization of a submodular set function by a greedy-type algorithm. Then again the Ellipsoid Method can be used to reduce this problem to the problem of optimizing over polymatroid polyhedra. Finally the Greedy Algorithm can be used yet again to solve this last problem. The strength of the combination of these two very general algorithms — the Ellipsoid Method and the Greedy Algorithm (see Box 1C) — is indeed surprising.

Of course the algorithms obtained by such general methods are useless in practice; we dare not even estimate their running time! But such polynomial algorithms, just by virtue of their existence, serve as signals that for certain problems it seems worth-while to look for taylor-made polynomial-time algorithms which, of course, have to make much more use of the peculiarities of the particular problem at hand.

We will now describe a solution of the Matroid Matching Problem in which we attempt to deal with general 2-polymatroids as long as we can, and then use the linearity or other special hypotheses only in one final crucial step. Also, we shall be concerned only with a good characterization of the matching number and not with any algorithmic aspects. Based on these ideas, a polynomial-time algorithm can be developed to determine the matching number of a linear polymatroid (Lovász (1981)). This algorithm is polynomial, but very slow. (A rough estimate of its running time is $O(n^{17})$, but even after much additional work, no implementation faster than $O(n^{10})$ is known at present.)

Before discussing these results, let us take a closer look at the matching problem for linear 2-polymatroids. We are going to show that the approach used in Section 8.2, that is, the use of linear algebra and algebraically independent variables, generalizes easily to the matroid case. We consider the Matroid Parity Problem formulation, where the matroid is given by an $r \times 2n$ matrix A. (For sake of simplicity, assume that we are working over the rational field.) Let the $2n$ columns of A be split into pairs $(\boldsymbol{a}_1, \boldsymbol{b}_1), \ldots, (\boldsymbol{a}_n, \boldsymbol{b}_n)$.

What is the maximum size of an independent set of columns which is the union of some of these pairs? Of course this is a special case of the Matroid Matching Problem. For every $1 \le i \le n$, consider the **wedge product** $\boldsymbol{a}_i \wedge \boldsymbol{b}_i$, that is, the $r \times r$ matrix defined by

$$(\boldsymbol{a}_i \wedge \boldsymbol{b}_i)_{\mu\nu} = (\boldsymbol{a}_i)_\mu (\boldsymbol{b}_i)_\nu - (\boldsymbol{a}_i)_\nu (\boldsymbol{b}_i)_\mu.$$

Clearly $\boldsymbol{a}_i \wedge \boldsymbol{b}_i$ is skew symmetric. Now the following theorem can be proved.

11.1.2. THEOREM. *The maximum number of columns of A which are linearly independent and which consist of some of the pairs $(\boldsymbol{a}_i, \boldsymbol{b}_i)$ is equal to the maximum rank of the matrix $\sum_{i=1}^{n} x_i(\boldsymbol{a}_i \wedge \boldsymbol{b}_i)$, where the maximum is taken over all reals $x_1, \ldots, x_n$. This maximum is attained when $x_1, \ldots, x_n$ are algebraically independent.* ∎

The proof of this result is not difficult and is omitted. (See Lovász (1979b) where a proof for a special case is given. The proof in the general case is the same, however.) Just as in Section 8.2, this theorem can be used to obtain a Monte Carlo algorithm to solve the matching problem for matroids linear over the rational field. This, however, cannot be regarded as a full solution to the problem. First, it does not yield a good characterization of the matching number, and second, this algorithm runs into the same numerical difficulties which we sometimes try to avoid by using matroid matching formulations of certain problems, as we shall see below.

We conclude this section with two examples of models taken from engineering disciplines which lead to combinatorial problems reducible to matroid matching.

Let H be a **planar bar and joint structure**, that is, a set of rigid bars interconnected by flexible joints in which all motions of the structure are assumed to be restricted to the plane. There are many beautiful mathematical problems which arise from the study of rigidity of such structures and quite a few of these concern matroids. Here we only treat one of these, and refer the interested reader to Crapo (1979) and Lovász and Yemini (1982). A forthcoming book by A. Recski on matroids will treat several of these in detail.

A planar bar and joint structure can be represented by a graph G whose points are points of the euclidean plane. An **infinitesimal motion** of such a structure is an assignment of a velocity $\boldsymbol{v}(x) \in \Re^2$ to each point $x \in V(G)$ so that no line is "stretched" or "compressed", that is, for every line $xy \in E(G)$, the vector $\boldsymbol{x}-\boldsymbol{y}$ is perpendicular to the vector $\boldsymbol{v}(x)-\boldsymbol{v}(y)$, or equivalently, $(\boldsymbol{x}-\boldsymbol{y})\cdot(\boldsymbol{v}(x)-\boldsymbol{v}(y))=0$.

A **finite** or **physical** motion of G consists of an "orbit" $x(t)$, $(t_1 \le t \le t_2)$, of each point $x \in V(G)$ such that the distance $|y(t)-x(t)|$ remains constant for each bar $xy \in E(G)$. Under appropriate differentiability conditions, the velocities $D_t x(t_0) = v(x)$ form an infinitesimal motion of the structure at time t_0. The converse, however, is not true and a structure may have non-trivial infinitesimal motions which do not correspond to any physical motion. In what follows we shall be concerned with infinitesimal motions only.

We shall also consider structures in which some of the points are "pinned down". If a set $P \subseteq V(G)$ is pinned down then the infinitesimal motions of the resulting pinned planar structure (G,P) are those infinitesimal motions $\boldsymbol{v}$ of G for which $\boldsymbol{v}(x)=0$ for each $x \in P$. We say that a pinned structure is **rigid** if it has no infinitesimal motion other than **0**.

We now ask: given a planar structure G, what is the minimum number of pins needed to "fix" it, that is, what is the minimum cardinality of a set $P \subseteq V(G)$ such that if (G,P) denotes a graph G pinned at the points of P, then (G,P) is rigid? This question was suggested by Crapo and discussed in Lovász (1980). Here we only show how to reduce it to the Matroid Matching Problem.

The infinitesimal motions of a (pinned) planar structure (G,P) form a linear space. We denote the dimension of this space by $\phi(G,P)$ and call it the **degree of freedom** of the structure (G,P). Given an (unpinned) planar structure G, let us define a set function on $V(G)$ by

$$f(X) = \phi(G,\emptyset) - \phi(G,X). \tag{11.1.4}$$

Then $f(X)$ reflects how many degrees of freedom are removed by pinning down the set X. It is not difficult to show, using the definition of infinitesimal motions and linear algebra, that f is submodular, monotone, $f(\emptyset) = 0$ and $f(\{x\}) = 2$ for every $x \in V(G)$. So (S,f) is a 2-polymatroid. (In fact, this 2-polymatroid is even *linear*!) Furthermore, pinning down X makes the structure rigid if and only if X is a spanning set. So the pinning problem can be reduced to the Minimum Spanning Set Problem for a 2-polymatroid.

Our second example comes from electrical engineering. Consider an electrical network consisting of resistors as well as voltage and current generators. This may be described by a graph G, having three kinds of lines: "resistors", "voltage sources" and "current sources". Every line e of G has a certain current, i_e, and a certain voltage, u_e, assigned to it. To make this precise, we must specify an arbitrary orientation of G and define the sign of the current i_e to be positive if it goes from the tail to the head of e and negative, otherwise. Similarly, the voltage u_e has a positive sign if the head of e has higher potential than its tail, and a negative sign, otherwise. We must emphasize that this orientation of the lines is purely for reasons of reference and has no physical significance. We shall assume also that G is connected throughout.

The voltages and currents satisfy the so-called Kirchhoff Laws. The first set of these, called "node laws", one for each point ("node") v of G, says that the currents i_e form a "circulation". That is

$$\sum_\alpha i_\alpha - \sum_\beta i_\beta = 0. \tag{11.1.5}$$

Here α ranges over all lines having their tail at v and β ranges over those lines having their head at v. The second set of laws, one for each routed cycle ("loop") C in G, are called "loop laws" and read:

$$\sum_{\gamma} u_{\gamma} - \sum_{\delta} u_{\delta} = 0. \qquad (11.1.6)$$

Here γ ranges over all forward lines of C and δ over all backward lines of C. These laws reflect the fact that u_e is the difference of potentials of the endpoints of e.

The Kirchhoff Laws do not determine the currents and voltages uniquely. In fact, it is well known (see Bollobás (1979) or Seshu and Reed (1961)) that they provide exactly q linearly independent linear equations among the $2q$ variables i_e and u_e. Further relations are provided by Ohm's Law:

$$u_e = R_e i_e \qquad (11.1.7)$$

for every line e containing a resistor, and by lines e containing given voltage sources:

$$u_e = \bar{u}_e \qquad (11.1.8)$$

and containing given current sources:

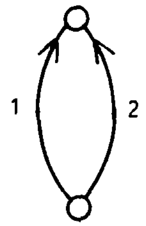
$$i_e = \bar{i}_e, \qquad (11.1.9)$$

where $\bar{u}_e$ and $\bar{i}_e$ are given values. These relations yield q further equations for the variables i_e and u_e. Thus the correct total number of these equations is obtained and we ask the question: do these equations determine the currents and voltages uniquely? There is, of course, a general answer to this question. Pick q independent Kirchhoff equations together with q further equations of the form (11.1.7) - (11.1.9) and evaluate the determinant of the resulting system. If this determinant is non-zero then there is a unique solution; otherwise, the solution is not unique. But there is a simpler approach to this question. It is not difficult to prove that *the system* (11.1.5) - (11.1.9) *has a unique solution if and only if* G *has a spanning tree which contains all voltage sources, but none of the current sources*. Not only is this condition more appealing to a graph theorist, but it is easier to verify and we do not run into difficulties with numerical calculations (rounding errors, etc.) which may arise if we must evaluate a determinant.

In the spanning tree condition for unique solvability of (11.1.5) - (11.1.9) formulated above, we have tacitly assumed that the resistances R_e are positive. This is too restrictive, however, since in modelling some of the more advanced electronic devices, we must use negative resistances!

If R_e is allowed to be negative then the condition given above for the unique solvability of the system of equations (11.1.5) – (11.1.9) is no longer sufficient.

Consider the very simple network in Figure 11.1.1. Here we have the equations $u_1 - u_2 = 0$, $i_1 + i_2 = 0$, $u_1 = Ri_1$ and $u_2 = -Ri_2$. This system has no unique solution, in spite of the obvious fact that either one of the



$R_1 = R$
$R_2 = -R$

FIGURE 11.1.1.

two lines forms a spanning tree which contains all voltage sources, but no current source.

There is, however, a way to save something of this nice combinatorial condition involving a spanning tree. One may argue that the situation in Figure 11.1.1 will never occur in a real-life situation since the two lines of the graph represent two different devices and if R and $-R$ are physical parameters associated with these devices, the two will never be precisely equal! If they are even infinitesimally different, the system will indeed have a unique solution.

We can generalize this remark by proving the following: if an electrical network consists of voltage and current generators and (positive and negative) resistances, and the resistances are algebraically independent over the rationals, then the system (11.1.5)–(11.1.9) has a unique solution if and only if the graph G has a spanning tree which contains all voltage generators, but none of the current generators.

When modelling even more complicated electrical networks, we must sometimes consider devices having more than two "outlets". A traditional way to do this is to consider certain pairs of lines of the graph G and assume that the two currents and the two voltages of these lines satisfy two independent linear relations. Such a pair of lines is called a **2-port**.

The most common example of a 2-port is an (ideal) **transformer**, where the two voltages u_1, u_2 and the two currents i_1, i_2 of the two lines satisfy $u_2 = ku_1$ and $i_2 = -i_1/k$. Here we shall treat in detail networks containing a less well-known, but still fairly simple device, a so-called gyrator. (The case of general 2-ports can be handled by similar means; see Recski (1980).) A **gyrator** is a pair of lines e and f whose currents and voltages are required to satisfy the following two equations:

$$u_e = R_f i_f, \qquad u_f = R_e i_e, \tag{11.1.10}$$

where $R_e + R_f = 0$. (Note the difference between a gyrator and the pair of lines in Figure 11.1.1. In a gyrator, the two R's are physical parameters of the *same* device and so may be precisely equal in absolute value for physical reasons.)

We now sketch the proof of the following theorem (Milić (1974)).

11.1.3. THEOREM. *Let G be a network consisting of voltage and current sources, resistances, and gyrators. Suppose that the resistances and the parameters of the gyrators are algebraically independent over the rationals. Then the system* (11.1.5)–(11.1.10) *has a unique solution if and only if G has a spanning tree which contains all voltage sources, no current source, and either both or none of the lines of any gyrator.*

PROOF. (sketch): For the sake of simplicity, suppose that the network contains only gyrators. Then the lines of the network may be divided into pairs $(x_i y_i, u_i v_i)$, for $i = 1, \ldots, n$, in such a way that the voltages and currents of each pair satisfy the equations

$$u_{x_k y_k} = R_k i_{u_k v_k}, \qquad u_{u_k v_k} = -R_k i_{x_k y_k}. \tag{11.1.11}$$

From the Kirchhoff voltage equations we know that the voltages can be expressed in the form:

$$u_{xy} = \Phi_y - \Phi_x,$$

where Φ is the potential. From (11.1.11), the currents i_{xy} can also be expressed in terms of the potential Φ_x, and then, upon substituting these values into the second set of Kirchhoff voltage equations, we get p equations for the p variables Φ_x. Of course these equations are not independent, since the translation $\Phi_x \to \Phi_x + t$ leaves them invariant. The original system then has a unique solution if and only if this new system has rank $p-1$.

Now the crucial point to be made is that working through these computations, we see that this new system has the form $C\Phi = 0$, where

$$C = \sum_{i=1}^{n} \frac{1}{R_i}(\boldsymbol{a}_i \wedge \boldsymbol{b}_i),$$

$\boldsymbol{a}_i$ is the incidence vector of the directed line $x_i y_i$ and $\boldsymbol{b}_i$ is the incidence vector of the directed line $u_i v_i$. Theorem 11.1.3 now follows from Theorem 11.1.2. ∎

Of course, (11.1.5)–(11.1.10) is a system of linear equations and thus its unique solvability can be checked by evaluating its determinant. However, for the same reasons as before, a combinatorial solution is often preferable.

Theorem 11.1.3 can actually be interpreted as saying that under the hypothesis that the physical parameters are "independent", the problem is equivalent to a parity problem for a *graphic matroid.*

Let G be a graph and suppose $A \subseteq V(G)$. A path in G is called an **A-path** if its endpoints belong to A, but no inner point of it belongs to A. Gallai (1961) solved the problem of finding the maximum number of *point-disjoint* A-paths by reducing it to a matching problem. He also raised the problem for *openly disjoint* A-paths. This latter problem was solved (in the sense of a minimax result) by Mader (1978b). This problem can be reduced to a matroid matching problem as described in the exercises below. For a derivation of the minimax formula via this reduction, see Lovász (1980).

11.1.4. EXERCISE. Show that the problem of finding the maximum number of point-disjoint A-paths can be reduced to a maximum matching problem.

11.1.5. EXERCISE. Let G be a graph and suppose $A \subseteq V(G)$. For each $X \subseteq V(G)$, define

$$r(X) = |X - A| + \min\{2, |X \cap A|\}.$$

Now let $Y \subseteq E(G)$ and suppose $Y_1, \dots, Y_k$ are the connected components of the graph formed by the lines in Y. Let

$$f(Y) = \sum_{i=1}^{k} (r(V(Y_i)) + |Y_i| - 2).$$

(a) Verify that $(E(G), f)$ is a 2-polymatroid.

(b) Show that the maximum number of openly disjoint A-paths is given by $\nu(E(G), f) + |A| - p$.

(c) Determine the double circuits in the 2-polymatroid $(E(G), f)$.

(d) Find a minimax formula for the maximum number of openly disjoint A-paths.

11.2. The Main Theorem of Polymatroid Matching

We now try in some sense to determine the number $\nu(S, f)$, that is, the size of a maximum matching for a 2-polymatroid (S, f). We shall need some preliminaries which may be viewed as a basic study of 2-polymatroids. Analogous results are commonplace for graphs and well-known for matroids, but we must include some of them here since they are not formulated for 2-polymatroids elsewhere.

Let (S, f) be a polymatroid, and $S = S_1 \cup S_2$ a partition of S. We say that (S, f) is a **direct sum** of (S_1, f) and (S_2, f) if $f(S_1) + f(S_2) = f(S)$. It then follows that for any $X_1 \subseteq S_1$ and $X_2 \subseteq S_2$,

$$f(X_1) + f(X_2) = f(X_1 \cup X_2).$$

(This fact is well-known for matroids and it follows for polymatroids in precisely the same way).

Let (S, f) and (S, g) be two polymatroids on the same set. We say that (S, g) is a **projection** of (S, f) if the set function $f - g$ is monotone increasing. Since $f(\emptyset) = g(\emptyset) = 0$, it follows that $f(X) \geq g(X)$ for every $X \subseteq S$. The name "projection" comes from the most important example thereof in which (S, f) is linear and (S, g) is represented by the orthogonal projections of the subspaces representing (S, f) into some subspace. Another example of a projection is the following. Let (S, f) be an arbitrary polymatroid and suppose $a \in S$. Define a set function f_a on S by

$$f_a(X) = \begin{cases} f(X) - 1, & \text{if } f(X \cup \{a\}) = f(X) \\ f(X), & \text{otherwise.} \end{cases}$$

It is easy to check that (S, f_a) is a polymatroid and, in fact, a projection of (S, f). In the special case when (S, f) is linear, (S, f_a) may be viewed as the projection of (S, f) to a hyperplane perpendicular to a "general" vector in the subspace representing a. Such a vector v is contained in the linear span of $X \subseteq S$ if and only if the whole subspace representing a is contained in the same span; this fact is reflected in the definition of f_a. But let us emphasize that this construction works for non-linear polymatroids as well.

11.2.1. EXERCISE. In Section 1.2, a similar set function was derived from a submodular set function (S, f) by the formula

$$f_1(X) = \begin{cases} f(X) - 1, & \text{if } f(X - a) + f(a) = f(X), \\ f(X), & \text{otherwise.} \end{cases}$$

Show that if (S, f) is a polymatroid and if $f(a) > 0$, then (S, f_1) is also a polymatroid, but not necessarily a projection of (S, f).

We shall also need to identify some special configurations in 2-polymatroids. A 2-polymatroid (Z, f) is called a **circuit** if $f(Z) = 2|Z| - 1$, but $f(X) = 2|X|$ for every proper subset X of Z. Thus, circuits are "almost" matchings in the sense that deleting any element from them leaves a matching, and the rank of the whole set is only one less than it would be for a matching. (A set Z such that Z is not a matching, but every proper subset of Z is a matching, is not necessarily a circuit. As a counterexample consider a linear space of dimension 4 and three subspaces of dimension 2 such that the intersection of any two is the 0 subspace.)

A 2-polymatroid (Z, f) is called a **flower** if $f(Z) = 2|Z| - 1$. A **k-flower** is a flower with $|Z| = k + 1$. Thus, every circuit is a flower, but not conversely. In fact, flowers are characterized by the following simple lemma.

11.2.2. LEMMA. *A 2-polymatroid is a flower if and only if it is the direct sum of a circuit and a matching.*

PROOF. It is trivial to see that the direct sum of a circuit and a matching is a flower. To prove the converse, assume that (Z, f) is a flower. Let T be a minimal subset of Z such that $f(T) < 2|T|$. Then

$$f(Z) \leq f(T) + f(Z - T) \leq 2|T| - 1 + 2|Z - T| \leq 2|Z| - 1 = f(Z).$$

Thus we must have equality throughout; in particular we have that $f(T) = 2|T| - 1$ and $f(Z - T) = 2|Z - T|$. Since every proper subset of T is a matching by definition, we see that T is a circuit and $Z - T$ is a matching. Furthermore, the fact that the first inequality holds with equality means that (Z, f) is the direct sum of $Z - T$ and T. ∎

11.2.3. LEMMA. *Every flower contains a unique circuit.*

PROOF. We know by the preceding lemma that every flower contains at least one circuit. Suppose that a flower (Z, f) contains two circuits, say C_1 and C_2. Then, $C_1 \cap C_2$ is a matching, but $C_2 \cup (Z - C_1)$ is not. So, applying submodularity, we obtain

$$\begin{aligned} f(Z) &\le f(C_1) + f(C_2 \cup (Z - C_1)) - f(C_1 \cap C_2) \\ &\le 2|C_1| - 1 + 2|C_2 \cup (Z - C_1)| - 1 - 2|C_1 \cap C_2| \\ &= 2|Z| - 2, \end{aligned}$$

a contradiction. ■

A 2-polymatroid (Z, f) is called a **double circuit** if $f(Z) = 2|Z| - 2$, and $f(Z - x) = f(Z) - 1 = 2|Z - x| - 1$ for every $x \in Z$. In a sense similar to the case of circuits we can point out that double circuits have the property that they are neither matchings nor flowers, but every proper subset of them is either a matching or a flower. Again this minimality property does not characterize double circuits. The structure of double circuits is described by the following lemma, the proof of which is left to the reader.

11.2.4. LEMMA. *A 2-polymatroid (Z, f) is a double circuit if and only if Z has a partition $Z = Z_1 \cup \cdots \cup Z_k$ $(k \ge 2)$ such that*

$$f(X) = \begin{cases} 2|X| - 2, & \text{if } X = Z, \\ 2|X| - 1, & \text{if } X \text{ contains exactly } k-1 \text{ of the sets } Z_1, \ldots, Z_k, \\ 2|X|, & \text{otherwise.} \end{cases}$$

The sets $Z - Z_i$ for $i = 1, \ldots, k$ are circuits and these are all the circuits contained in the double flower. Consequently, the above partition is uniquely determined. ■

In the case when $k = 2$, the double circuit is simply the direct sum of two circuits. Such a double circuit will be called **trivial**.

The direct sum of a double circuit and a matching is called a **double flower**. A **k-double flower** is a double flower with $|Z| = k + 2$. It follows that a double flower (Z, f) has rank $f(Z) = 2|Z| - 2$. The description of double flowers is slightly more complicated than that given in Lemma 11.2.1, but it can be derived in an analogous way.

11.2.5. LEMMA. *If a 2-polymatroid (Z, f) has rank $f(Z) = 2|Z| - 2$, then it is either a double flower or Z has an element x such that $Z - x$ is a matching.* ■

With these definitions and lemmas in hand, we can turn our attention once again to matchings. It is clear that for any partition $S = S_1 \cup S_2$,

$$\nu(S, f) \le \nu(S_1, f) + \nu(S_2, f). \tag{11.2.1}$$

Of course, we do not have equality here in general. One instance of equality certainly occurs if (S, f) is the direct sum of (S_1, f) and (S_2, f), but equality may hold in other cases as well.

Let (S, f) be a 2-polymatroid and let (S, g) be a projection of it. Then we can prove that

$$\nu(S, f) \leq \nu(S, g) + f(S) - g(S). \tag{11.2.2}$$

In fact, let X be a maximum matching in (S, f). Then

$$g(X) \geq f(X) - f(S) + g(S)$$

by the definition of projection, and so by (11.1.3)

$$\begin{aligned} \nu(S, g) &\geq \nu(X, g) \\ &\geq g(X) - |X| \\ &\geq f(X) - f(S) + g(S) - |X| \\ &= |X| - f(S) + g(S) \\ &= \nu(S, f) - f(S) + g(S). \end{aligned}$$

Also note that trivially, $\nu(S, f) \leq \lfloor f(S)/2 \rfloor$. Combining these three observations, we obtain the following.

11.2.6. LEMMA. *Let (S, f) be a 2-polymatroid, let (S, g) be a projection of it and let $S = \{S_1, \ldots, S_k\}$ be a partition of S. Then*

$$\nu(S, f) \leq f(S) - g(S) + \sum_{i=1}^{k} \left\lfloor \frac{g(S_i)}{2} \right\rfloor. \tag{11.2.3}$$

∎

The authors do not know of any counterexample to the conjecture that the minimum of the right hand side of inequality (11.2.3), taken over all partitions of S and all projections of (S, f), is equal to $\nu(S, f)$. As we shall see below, this is certainly true for *linear* 2-polymatroids. The reason why an appropriate projection is easier to construct for linear 2-polymatroids is that these polymatroids have linear projections which can be easily described. In the general case, even if the above-mentioned conjecture were true, the projection (S, g) giving equality in (11.2.3) could not be constructively described. More precisely, even if the existence of g and $\{S_1, \ldots, S_k\}$ could be proved, the numbers $g(S)$ and $g(S_i)$ could not be computed in polynomial time from an oracle evaluating f on any given subset (see Box 11A).

However, it turns out that the following general theorem can be proved for *every* 2-polymatroid. While it does not provide a good characterization of the matching number in general, it does circumvent most of the difficulties in the sense that in special cases only relatively simple arguments need be added to obtain a minimax formula. The proof is constructive in the sense that it can be turned into a polynomial-time algorithm. We shall not describe the algorithm here, but only refer the interested reader to Lovász (1981).

11.2.7. THEOREM. *Let (S,f) be a 2-polymatroid and set $\nu(S,f)=\nu$. Then at least one of the following assertions holds:*

(i) $f(S)=2\nu+1$.

(ii) *S has a partition $\{S_1,S_2\}$ into non-empty subsets so that $\nu(S_1,f)+\nu(S_2,f)=\nu$.*

(iii) *S has an element a which is contained in the span of every maximum matching.*

(iv) *(S,f) contains a non-trivial ν-double flower.*

Before proceeding to the proof of this theorem let us make a few observations. In fact, let us discuss the theorem from the point of view of our old friends King Arthur and Merlin (see Box 1A). The King sets Merlin the task of finding $\nu(S,f)$ for a 2-polymatroid (S,f). The 2-polymatroid is supplied by Pythia (see Box 11A), if we may be forgiven for mixing our historical metaphors. Again, Merlin of course knows the correct answer that $\nu(S,f)=100$ immediately. (In fact, he can learn *all* values of f from Pythia; transferring data in such quantities is no problem between supernatural beings!) But how does Merlin convince King Arthur of this?

First he picks a set $X\subseteq S$ with $|X|=100$ and $f(X)=200$. (Pythia readily testifies to this latter value.) So the King is convinced that $\nu(S,f)\geq 100$, but it remains to convince him that $\nu(S,f)\leq 100$.

Let us consider the four cases in Theorem 11.2.7. If (i) holds, then Merlin asks Pythia (in the presence of Arthur) for the value of $f(S)$. As $f(S)=201$, even King Arthur has to agree that no matching larger than 100 occurs in (S,f).

Now suppose (ii) holds. That is, suppose there exists a partition $S=S_1\cup S_2$ with $\nu(S_1,f)=72$ and $\nu(S_2,f)=28$. Then Merlin can break up his task into the two smaller tasks of exhibiting that $\nu(S_1,f)\leq 72$ and $\nu(S_2,f)\leq 28$. This will of course convince King Arthur that $\nu(S,f)\leq 72+28=100$.

If (iii) holds then he considers the projection $\nu(S,f_a)$ as defined at the beginning of this section. Then $\nu(S,f_a)=99$. In fact, $\nu(S,f_a)\geq\nu(S,f)-1$ by inequality (11.2.2). However, if M is any maximum matching in (S,f), then $f(M\cup\{a\})=f(M)$ by the hypothesis on element a and so $f_a(M) = f(M) - 1 = 2|M| - 1$. But then M is not a matching in (S, f_a). Hence $\nu(S, f_a) = \nu(S, f) - 1 = 99$. So it suffices to convince King Arthur that $\nu(S, f_a) \leq 99$, for then he will have to accept the fact that $\nu(S, f) \leq \nu(S, f_a) + 1 \leq 100$.

No general advice can be given to Merlin if case (iv) occurs, but in the case of various special classes of 2-polymatroids simple reductions to smaller cases can be found using the existence of a ν-double flower. Some examples will be discussed in the next section.

PROOF (of Theorem 11.2.7). Let us suppose that (S, f) does not satisfy (i), (ii) or (iii). We shall prove that it satisfies (iv).

Observe first that if (S, f) contains an element with rank 0 or 1 then (ii) is satisfied by the partition of S into this element and its complement in S. Hence every element has rank 2.

Let us construct a hypergraph H whose points are the elements of S and whose lines are those circuits of (S, f) which are contained in some ν-flower. We proceed to derive some properties of this hypergraph.

Claim 1. H contains no isolated points.

In fact, suppose $a \in S$. Since (iii) does not hold, there exists a maximum matching B in (S, f) such that a is not contained in the span of B, that is,

$$f(B \cup a) > f(B) = 2|B|.$$

But $B \cup a$ cannot be a matching by the maximality of B, and so we must have

$$f(B \cup a) = 2|B| + 1.$$

Thus $B \cup a$ is a ν-flower. Let C be the unique circuit in $B \cup a$ (see Lemma 11.2.3). Since B is a matching, it cannot contain C and so $a \in C$. Since C is a line of the hypergraph H, this proves the claim.

Claim 2. H is connected.

Suppose not. Then let $S = S_1 \cup S_2$ be a partition of S into two non-empty parts so that every line of H is contained in either S_1 or S_2. Let M_i be a maximum matching in (S_i, f) and let M be a maximum matching in (S, f). Furthermore, choose M_1, M_2 and M in such a way that

$$|M \cap M_1| + |M \cap M_2| \tag{11.2.4}$$

is maximum. We will show that $M_1 \subseteq M$ and $M_2 \subseteq M$. This will clearly imply that $M = M_1 \cup M_2$ and so $\nu(S, f) = \nu(S_1, f) + \nu(S_2, f)$, which will in turn contradict the hypothesis that (ii) does not hold.

So suppose, by way of contradiction, that $M_1 \not\subseteq M$ (say). Suppose first now that $M_1 \not\subseteq \operatorname{Span} M$. Then we can choose an element $a \in M_1 -$

$\operatorname{Span} M$. As in the proof of Claim 1, we see that $M \cup a$ is a ν-flower. Let C be the unique circuit contained in $M \cup a$. Then $C \subseteq S_1$ by the definition of S_1. Since M_1 is a matching, $C \not\subseteq M_1$ and so we can choose an element $a' \in C - M_1$. But then $M' = M - a' \cup a$ is a maximum matching of (S, f), $|M' \cap M_1| > |M \cap M_1|$ and $M' \cap M_2 = M \cap M_2$. This contradicts the maximality of (11.2.4).

Thus we may assume that $M_1 \subseteq \operatorname{Span} M$. By our hypothesis, $M_1 \not\subseteq M$. Let $c \in M_1 - M$. Then by Claim 1, we can find a ν-flower Z whose circuit contains c. Then $N = Z - c$ is a maximum matching of (S, f) such that $M_1 - M \not\subseteq \operatorname{Span} N$. Among all choices of M, M_1, M_2 and N satisfying these hypotheses, select a set for which

$$|M \cap N| \tag{11.2.5}$$

is maximal. Note that $N \not\subseteq \operatorname{Span} M$, since otherwise we would have $\operatorname{Span} N = \operatorname{Span} M$ and this would contradict the facts that $M_1 \subseteq \operatorname{Span} M$, but $M_1 \not\subseteq \operatorname{Span} N$. Thus we may choose an element $b \in N - \operatorname{Span} M$. Let C be the unique circuit in $M \cup b$. Again by the definition of S_1, we have that $C \subseteq S_1$.

Let us again distinguish two cases. If $C - b \not\subseteq M_1$, then let $b' \in C - b - M_1$. Then $M' = M \cup b - b'$ is a maximum matching of (S, f). Furthermore, $M' \cap M_1 = M \cap M_1$, $M' \cap M_2 = M \cap M_2$, but $|M' \cap N| > |M \cap N|$. This contradicts property (11.2.5).

If, on the other hand, $C - b \subseteq M_1$ then let $b' \in C - b$, $M' = M \cup b - b'$ and $M_1' = M_1 \cup b - b'$. Then clearly M' is a maximum matching of (S, f), M_1' is a maximum matching of (S_1, f) and furthermore, $|M' \cap M_1'| = |M \cap M_1|$, $|M' \cap M_2| = |M \cap M_2|$, $M_1' - M' = M_1 - M \not\subseteq \operatorname{Span} N$, but $|M' \cap N| > |M \cap N|$, which again contradicts the maximality of quantity (11.2.5). This proves Claim 2.

Claim 3. (S, f) contains a non-trivial ν-double flower.

To show this, let F_0 be any ν-flower in (S, f). If $\operatorname{Span} F_0 = S$ then (i) is trivially fulfilled, so we may assume that $\operatorname{Span} F_0 \neq S$. Define two subsets of $E(H)$ as follows: $\mathcal{A}$ is the set of those circuits of (S, f) which are contained in a ν-flower F with $\operatorname{Span} F = \operatorname{Span} F_0$, and $\mathcal{B} = E(H) - \mathcal{A}$. Then neither $\mathcal{A}$ nor $\mathcal{B}$ is empty for $\mathcal{A}$ contains the circuit of F_0, while $\mathcal{B}$ contains any line of H containing any element of $S - \operatorname{Span} F_0$, and such a line exists by Claim 1. Thus by the connectivity of H, there exist two circuits $C \in \mathcal{A}$ and $D \in \mathcal{B}$ such that $C \cap D \neq \emptyset$. Choose ν-flowers F and G containing C and D, respectively, such that $\operatorname{Span} F = \operatorname{Span} F_0$ and $|F \cap G|$ is maximum. By the definition of $\mathcal{B}$, we know that $\operatorname{Span} G \neq \operatorname{Span} F$, and so there exists an element $b \in F - \operatorname{Span} G$.

We claim that $G \cup b$ is a non-trivial ν-double flower. Since $b \notin \operatorname{Span} G$, we have that $f(G \cup b) > f(G) = 2\nu + 1$. If $f(G \cup b) = 2\nu + 4$ then $G \cup b$ is a $(\nu+2)$-element matching; if $f(G \cup b) = 2\nu + 3$ then $G \cup b$ is a $(\nu+1)$-flower. But both of these situations are impossible since the maximum size of a matching is ν. Thus $f(G \cup b) = 2\nu + 2$. By Lemma 11.2.5, this implies that $G \cup b$ is either a ν-double flower or it has an element x such that $G \cup b - x$ is a $(\nu + 1)$-element matching. But this second situation is again ruled out by the same argument as above. Thus $G \cup b$ is a ν-double flower.

It remains to show that this ν-double flower is non-trivial. Suppose that it is trivial. Then by Lemma 11.2.4, it contains exactly two circuits, one of which is clearly D. Let D' be the other circuit in $G \cup b$. Since $D \cap D' = \emptyset$, it must be the case that $D' \neq C$ and since C is the unique circuit in F, we have $D \not\subseteq F$. Suppose $d \in D - F$. Then $G \cup b - d = G'$ is a ν-flower containing D such that $|G' \cap F| > |G \cap F|$. This contradiction proves Claim 3 and the proof of the theorem is complete. ■

11.3. Matching in Special Polymatroids

In this section we show how Theorem 11.2.7 can be applied if the polymatroids in question have some special structure. As a first example we show that (with a few additional tricks) the Matroid Intersection Theorem can be obtained from Theorem 11.2.7. (Recall that the common independent sets of two matroids are the matchings of the 2-polymatroid obtained by adding their rank functions.) This will depend on the following lemma, which will rule out the unpleasant alternative (iv) of Theorem 11.2.7.

11.3.1. LEMMA. *Let (S, r_1) and (S, r_2) be two matroids on the same set with both $r_1(S)$ and $r_2(S) \leq k+1$, where k is any non-negative integer. Then the 2-polymatroid (S, r_1+r_2) contains no non-trivial k-double flower.*

PROOF. Suppose that D is a k-double flower in this 2-polymatroid. Let D be the direct sum of the non-trivial double circuit C and a matching M. Then $r_1(D) + r_2(D) = 2k + 2$. Since $r_i(D) \leq k + 1$ by hypothesis, this implies that

$$r_1(D) = r_2(D) = k + 1. \tag{11.3.1}$$

Furthermore, for every $x \in M$ we have $r_1(D - x) + r_2(D - x) = 2k$. Since $r_i(D - x) \geq r_i(D) - 1 = k$, this implies that $r_1(D - x) = r_2(D - x) = k$.

Finally, for every $x \in C$ we have $r_1(D-x)+r_2(D-x) = 2k+1$, and hence it follows that either $r_1(D-x) = k$ and $r_2(D-x) = k+1$ or vice versa. Thus we have a partition $C = C_1 \cup C_2$, such that $r_i(D-x) = k$ and $r_{3-i}(D-x) = k+1$ if $x \in C_i$. Neither C_1 nor C_2 can be empty, for if $C_1 = \varnothing$, for example, then $r_2(D-x) = k = r_2(D)-1$ for every $x \in D$, which implies that D is independent in the matroid (S, r_2) and hence $r_2(D) = |D| = k+2$, contradicting equation (11.3.1). It follows by the same argument that C_i is a circuit in the matroid (S, r_{3-i}). But then C_i is a circuit in the 2-polymatroid (S, r_1+r_2) as well, and thus D is a trivial double flower. ■

Now recall the Matroid Intersection Theorem from Section 1.3.

1.3.19. THEOREM. *Let (S, r_1) and (S, r_2) be two matroids with a common underlying set. Then the maximum size of a common independent set equals the minimum of $r_1(X)+r_2(S-X)$ over all subsets X of S.*

PROOF. We will show how to derive the non-trivial part of this result from Theorem 11.2.7 and Lemma 11.3.1.

Let k denote the maximum size of a common independent set of the two matroids. We want to show that S has a subset X such that $r_1(X)+r_2(S-X) = k$. We may assume without loss of generality that $r_1(S) \le k+1$ and $r_2(S) \le k+1$, since otherwise we could "truncate" the two matroids (that is, consider the rank function $r_i' = \min\{r_i, k+1\}$ instead of r_i). Set $f = r_1+r_2$ and note that $\nu(S,f) = k$. Let us now apply Theorem 11.2.7. Alternative (iv) is ruled out by Lemma 11.3.1, so one of (i), (ii) and (iii) must hold. If (i) occurs then $f(S) = r_1(S)+r_2(S) = 2k+1$, and hence one of $r_1(S)$ and $r_2(S)$, say $r_1(S)$, is equal to k. But then we can take $X = S$.

Suppose that (ii) holds; that is, suppose S has a partition $S_1 \cup S_2$ such that $\nu(S_1,f)+\nu(S_2,f) = k$. Using induction on S, we may assume that each S_i has a subset X_i such that $r_1(X_i)+r_2(S_i-X_i) = \nu(S_i,f)$. If we set $X = X_1 \cup X_2$, then

$$\begin{aligned} r_1(X)+r_2(S-X) &\le r_1(X_1)+r_1(X_2)+r_2(S_1-X_1)+r_2(S_2-X_2) \\ &= \nu(S_1,f)+\nu(S_2,f) = k. \end{aligned}$$

This proves the theorem in this case.

Finally, if (iii) holds, then S has an element a which is contained in the span of every maximum matching. We may assume, for example, that $r_1(a) = 1$, since if $r_1(a) = r_2(a) = 0$, then a could be deleted from both matroids without changing anything. Now let $(S-a, r_1')$ be the matroid obtained from (S, r_1) by contracting the element a. But then the 2-polymatroid $(S-a, r_1'+r_2)$ contains no k-element matching, for such a matching M would be a maximum matching in (S, f) as well, and a would not be contained in its span. So we may assume (by induction on k), that $S-a$ has a subset X such that

$$r_1'(S-a-X)+r_2(X)\le k-1.$$

But then

$$r_1(S-X)+r_2(X)=r_1'(S-a-X)+1+r_2(X)\le k,$$

and the theorem is proved again. ∎

This proof of the Matroid Intersection Theorem is not very elegant, especially if we consider that we have used Theorem 11.2.7, the proof of which was quite involved. But it is important to see that at least in a sense, general polymatroid matching results can be used to solve the Matroid Intersection Problem. (The reader who has solved Exercise 3.1.4 will remember that to obtain the Marriage Theorem from Tutte's Theorem is not completely straightforward either; it is instructive to compare the arguments needed there with our proof.) On the other hand, the next two applications of the polymatroid matching theorem are such that no other proofs of the results are known (although in the case of the second, a direct proof would be desirable.)

11.3.2. THEOREM. *Let (S, f) be a linear 2-polymatroid. Then*

$$\nu(S,f)=\min\{f(S)-g(S)+\sum_{i=1}^{k}\lfloor g(S_i)/2\rfloor\},$$

where the minimum is taken over all partitions $\{S_1,\dots,S_k\}$ of S and over all linear projections (S, g) of (S, f).

Before proving this theorem it is worth-while to consider some special cases. For $\nu(S,f)=1$, the theorem says that if we have a set of lines in a projective space such that any two such lines have a point in common, then either the lines all lie in a 2-dimensional subspace or they all pass through the same point. In the case when $\nu(S,f)=2$, the above theorem says that if a set of lines in a projective space contains no three lines such that each of them is disjoint from the span of the others (i.e., no 3 lines span a 5-dimensional subspace), then one of the following possibilities occurs: (a) all the lines lie in a 4-dimensional subspace; (b) all the

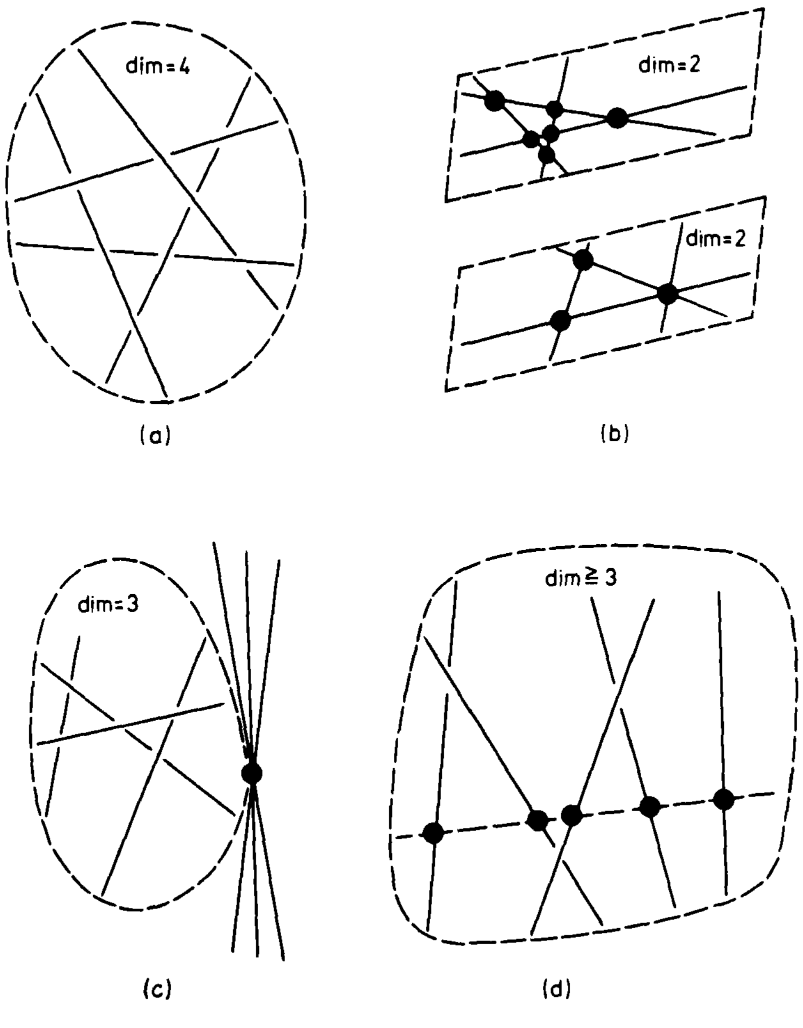


FIGURE 11.3.1. Linear polymatroids with $\nu = 1$

lines lie in two 2-dimensional subspaces; (c) there are two 3-dimensional subspaces, intersecting in a point p, such that every given line is either contained in one of the two subspaces or goes through p; (d) all the given lines meet a fixed line (see Figure 11.3.1). It is quite instructive to derive these configurations from Theorem 11.3.2.

11.3.3. LEMMA. *Let (D, f) be a non-trivial double circuit and let $\{D_1, \ldots, D_k\}$ be the partition of D such that $D - D_i$ is a circuit. Suppose* *that (D, f) is represented by 2-dimensional subspaces of a linear space and let K_i denote the linear subspace spanned by $D - D_i$. Then $K_1 \cap \cdots \cap K_k$ is not the null subspace.*

PROOF. First we prove the inequality

$$\dim(K_1 \cap \cdots \cap K_i) \geq \sum_{j=1}^{i} \dim K_i - (i-1)\dim(K_1 \cup \cdots \cup K_i). \quad (11.3.2)$$

This follows easily by induction on i. For $i = 1$ it is trivial. Suppose, then, that $i > 1$. Then

$$\begin{aligned}&\dim(K_1 \cap \cdots \cap K_i)\\ &= \dim(K_1 \cap \cdots \cap K_{i-1}) + \dim K_i - \dim((K_1 \cap \cdots \cap K_{i-1}) \cup K_i).\end{aligned}$$

Using the induction hypothesis, we obtain

$$\begin{aligned}&\dim(K_1 \cap \cdots \cap K_i)\\ &\geq \sum_{j=1}^{i} \dim K_j - (i-2)\dim(K_1 \cup \cdots \cup K_{i-1}) - \dim((K_1 \cap \cdots \cap K_{i-1}) \cup K_i)\\ &\geq \sum_{j=1}^{i} \dim K_i - (i-1)\dim(K_1 \cup \cdots \cup K_i).\end{aligned}$$

This proves inequality (11.3.2).

Since $D - D_i$ is a circuit, we have $\dim K_i = f(D - D_i) = 2|D - D_i| - 1$, and since D is a double circuit, we have $\dim(K_1 \cup \cdots \cup K_k) = f(D) = 2|D| - 2$. Substituting these values into (11.3.2) with $i = k$, we obtain

$$\begin{aligned}\dim(K_1 \cap \cdots \cap K_k) &\geq \sum_{j=1}^{k}(2|D - D_i| - 1) - (k-1)(2|D| - 2)\\ &= 2|D| - 2\sum_{j=1}^{k}|D_i| + k - 2\\ &= k - 2.\end{aligned}$$

Now since $k \geq 3$ by the hypothesis that (D, f) is a non-trivial double circuit, the proof is complete. ∎

Our next lemma describes a way to eliminate ν-double flowers in the case of linear 2-polymatroids. For a double flower $D \subseteq S$, let $K(D)$ denote the intersection of linear spans of all circuits contained in D. By Lemmas 11.2.4 and 11.3.3, $K(D)$ is not the null space.

11.3.4. LEMMA. *Let (S, f) be a linear 2-polymatroid with $\nu(S, f) = \nu$, and let D be a non-trivial ν-double flower in (S, f). Then $K(D)$ is contained in the linear span of every maximum matching of (S, f).*

PROOF. Assume, to the contrary, that (S, f) has a maximum matching M such that $K(D) \not\subseteq \overline{M}$, where $\overline{M}$ is the linear span of M. Choose M so that $M \cap D$ is maximum. Let p be any non-zero vector in $K(D) - \overline{M}$. Then $\dim(\overline{M \cup p}) = 2\nu + 2$. Since $\dim \overline{D} = f(D) = 2\nu + 1$, there exists a line $d \in D$ not contained in $\overline{M \cup p}$. But then $\dim(M \cup p \cup d) \geq 2\nu + 2$. On the other hand, $M \cup d$ is not a matching (since M is already a maximum matching), and so $\dim(M \cup d) \leq 2\nu + 1$. Hence it follows that $p \notin \overline{M \cup d}$, and also that $\dim \overline{M \cup d} = 2\nu + 1$; that is, $M \cup d$ is a flower in (S, f). Let C be the circuit in $M \cup d$. Then $p \notin \overline{C}$ and so $K(D) \not\subseteq \overline{C}$. But this implies that $C \not\subseteq D$, since $K(D)$ is the intersection of the linear spans of all circuits in D by definition. Thus there exists a $c \in C - D$. But then $M' = M \cup d - c$ is another maximum matching in (S, f) such that $K(D) \not\subseteq \overline{M'}$ (since $\overline{M'} \subseteq \overline{M \cup d}$, and $p \notin \overline{M \cup d}$), but $|M' \cap D| > |M \cap D|$. This contradiction proves the lemma. ■

We are prepared now to prove Theorem 11.3.2.

PROOF (of Theorem 11.3.2). By Lemma 11.2.6, it suffices to prove there exists a linear projection (S, g) of (S, f) and a partition $\{S_1, \ldots, S_k\}$ of S such that equality holds in the inequality (11.2.3). To this end, let (S, g) be a linear projection of (S, f), such that $\nu(S, g) = \nu(S, f) - f(S) + g(S)$ and $g(S)$ is minimum. Furthermore, let $\{S_1, \ldots, S_k\}$ be a finest partition of S such that $\nu(S, g) = \sum_{i=1}^{k} \nu(S_i, g)$. We claim that

$$\nu(S_i, g) = (g(S_i) - 1)/2, \tag{11.3.3}$$

which will prove the theorem.

Consider any S_i and apply Theorem 11.2.7 to the 2-polymatroid (S_i, g). Then (ii) is ruled out, for if S_i could be split into two non-empty sets S'_i and S''_i such that $\nu(S_i, g) = \nu(S'_i, g) + \nu(S''_i, g)$, then replacing S_i by $S'_i \cup S''_i$ in the partition we would get a contradiction of the hypothesis that $\{S_1, \ldots, S_k\}$ is non-refinable. Furthermore, (iii) is ruled out since if S_i has an element a contained in the span of every maximum matching of (S_i, g), then a is also contained in the span of every maximum matching of (S, g) (because for any maximum matching M of (S, g), $M \cap S_i$ is a maximum matching of (S_i, g)). Choose any non-zero vector v from a (which is a non-zero subspace in the linear representation of (S, g)) and project (S, g) onto the hyperplane orthogonal to v. Then the resulting

2-polymatroid (S, h) has $\nu(S, h) = \nu(S, g) - 1$ (since by the choice of v, the projection destroys every maximum matching of (S, g)). But this contradicts the choice of (S, g).

Finally, we can rule out (iv) as well. For let (D, g) be any $\nu(S_i, g)$-double flower in (S_i, g). Then $K(D)$ is not the null space by Lemma

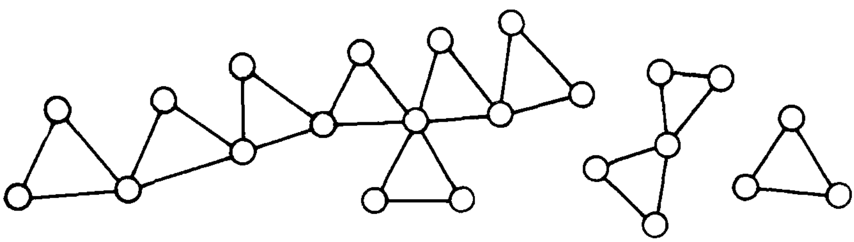


FIGURE 11.3.2. A triangular cactus

11.3.3. If we choose any $v \in K(D)$, $v \neq 0$, then v is contained in the linear span of every maximum matching of (S_i, g), by Lemma 11.3.4. From this point we proceed to get a contradiction just as in the preceding case.

Thus (i) must occur, which is just the content of equation (11.3.3). ∎

The last application of Theorem 11.2.7 we give is purely graph-theoretical. A **triangular cactus** is a graph whose blocks are triangles (see Figure 11.3.2). We study the following question: When does a graph contain a spanning connected triangular cactus? More generally, what is the maximum number of blocks (triangles) in any triangular cactus contained in G as a subgraph? We denote this maximum number of blocks by $\beta(G)$, and shall derive a minimax formula for $\beta(G)$.

Let us point out that this problem includes the matching problem in the following sense. Let G be any graph. Construct a new graph G' by adding a new point v to G and connecting v to all points of G. Then it is easy to verify that G' has a spanning triangular cactus if and only if G has a perfect matching. In fact, if G has a perfect matching then this yields a spanning triangular cactus of G' in a trivial way: just add all lines incident with the new point to this perfect matching. Conversely, if G' has a spanning triangular cactus H then deleting the new point from this triangular cactus, the remainder contains a perfect matching (since a triangular cactus is clearly a factor-critical graph). In sharp contrast, however, we do not know any way to reduce the Triangular Cactus Problem to the Matching Problem, although such a reduction is not inconceivable.

We can, however, formulate the Triangular Cactus Problem as a Matroid Matching Problem. Let G be a graph and let $(E(G), r)$ be its polygon matroid. Every triangle in G is a line (i.e., rank 2 subset) of this matroid. Thus the triangles in G can be viewed as the elements of a 2-polymatroid (S, f), in which $f(\{T_1, \ldots, T_k\}) = r(T_1 \cup \cdots \cup T_k)$ for any subset $\{T_1, \ldots, T_k\} \subseteq S$ and S is the set of all triangles. The connection between matchings in (S, f) and cacti in G is described in the following lemma.

11.3.5. LEMMA. *The triangles of a triangular cactus of G form a matching in (S, f) and vice versa.*

PROOF. Let H be a triangular cactus with c connected components and let H' be the set of its triangles. Then $r(E(H)) = |V(H)| - c$, and hence $f(H') = |V(H)| - c$. Since $H' = (|V(H)| - c)/2$ by an obvious calculation, this implies that H' is a matching in (S, f). The reverse implication follows similarly. ∎

The maximum number of blocks in a triangular cactus of a graph G is characterized in the following minimax theorem. Let $\mathcal{P} = \{V_1, \ldots, V_k\}$ be a partition of $V(G)$ and $\mathcal{Q} = \{E_1, \ldots, E_m\}$, a partition of $E(G)$. For $1 \le i \le m$, let u_i denote the number of classes V_j met by E_i and set

$$\phi(\mathcal{P}, \mathcal{Q}) = \sum_{i=1}^{m} (u_i - 1)/2 + p - k.$$

11.3.6. THEOREM. *The maximum number of blocks in a triangular cactus in a graph G is equal to the minimum of $\phi(\mathcal{P}, \mathcal{Q})$, taken over all partitions $\mathcal{P}$ of $V(G)$ and $\mathcal{Q}$ of $E(G)$ with the property that every triangle of G has either at least two points in the same class of $\mathcal{P}$ or all three lines in the same class of $\mathcal{Q}$.* ∎

As suggested by Lemma 11.3.5, this theorem can be proved by using the theorem on matroid matching (Theorem 11.2.7) in a fashion analogous to the proof of Theorem 11.3.2. One has to work out what the circuits and double circuits of (S, f) are. The main observation to be made is that for every non-trivial double circuit in (S, f) there are two points in G such that the identification of these two points destroys all circuits contained in this double circuit. The details are left to the reader as a series of Exercises (11.3.7 - 11.3.9).

11.3.7. EXERCISE. Let G be a simple graph.

(a) Show that two triangles of G with a line in common form a circuit in (S, f).

(b) Let C be a cycle in G and let C' be a set of triangles in G such that each triangle in C' contains exactly one line of C, each line of C is contained in exactly one triangle of C' and the points of the triangles in C' opposite to their lines in C are distinct points in $V(G)-V(C)$. Then C' is a circuit in (S, f).

(c) All circuits in (S, f) are as described in (a) or (b).

11.3.8. EXERCISE. (a) Show that three triangles of G with a line in common form a non-trivial double circuit in (S, f).

(b) Let C' arise from a cycle C of G as in part (b) of the previous exercise. Add to C' a triangle spanned by a line of C and a point not contained in any triangle of C'. Then the resulting set C'' of triangles is a non-trivial double circuit.

(c) Let D be a subgraph of G consisting of three openly disjoint paths joining the same pair of points in G (i.e., a so-called "θ-subgraph" of G) and let D' be a set of triangles in G such that each triangle in D' contains exactly one line of D, each line of D is contained in exactly one triangle of D' and the points of the triangles in D' opposite to their lines in D are distinct points in $V(G)-V(D)$. Then D' is a non-trivial double circuit in (S, f).

(d) All non-trivial double circuits in (S, f) are as described in (a), (b) and (c).

11.3.9. EXERCISE. Using Theorem 11.2.7, Lemma 11.3.5, and the previous exercises, prove Theorem 11.3.6.

This Page intentionally left blank

12

Vertex Packing and Covering

12.0. Introduction

The notion of independent points in a graph is formally dual to the notion of independent lines and, in this sense, the maximum number $\alpha(G)$ of independent points in G (called variously the **independence number** of G, the **stability number** of G or the vertex **packing number** of G) is dual to the notion $\nu(G)$ for independent lines. But there is a deeper relation between these two graphical invariants.

First, let us recall Gallai's Identity (Lemma 1.0.2):

$$\alpha(G) + \tau(G) = |V(G)|.$$

This tells us that the problem of determining a maximum set of independent points in a graph is equivalent to the problem of determining the minimum number of points covering all lines. The problem of point covers has already occurred in this book as the linear programming dual of the matching problem for bipartite graphs as well as for certain other classes of graphs (see Sections 6.3 and 7.1). Moreover, we saw that the bipartite matching algorithm presented in Section 1.2 produces, as a bonus for us, a minimum cover as well. So we know that vertex packing is polynomial for bipartite graphs.

But we also know that the problem of minimum point covers for non-bipartite graphs is NP-complete. (See Karp (1972) or Garey and Johnson (1979).) In fact, the problem remains NP-complete even in the special case of *triangle-free* graphs (Poljak (1974)) and in the case of *cubic planar* graphs (Garey, Johnson and Stockmeyer (1976)). Should this mean that it is useless to concern ourselves with point covers (or, equivalently, with independent sets of points) in the general case? We believe that NP-completeness of a problem only means that no answers of the strength and generality of, say, Tutte's Theorem can be expected, but the study of such problems may still prove quite fruitful.

Since deletion of a line does not decrease the maximum number of independent points, it is natural to ask which are the minimal graphs (with respect to deleting lines) with a given value of $\alpha(G)$ (or equivalently, $\tau(G)$). Such graphs are called **α-critical** (respectively, **τ-critical**) and turn out to have quite interesting properties. Of course, a complete description of τ-critical graphs would yield a good characterization of $\tau(G)$. More precisely, if the class of τ-critical graphs were in NP (that is, if there were a polynomial length way to exhibit that a given graph is τ-critical), then a good characterization of $\tau(G)$ could be easily obtained. (Can you see how to do this?)

Because of the NP-completeness of $\tau(G)$ we cannot hope that τ-critical graphs have a really simple structure, but several interesting and deep structural properties can be verified and a certain classification theorem proved. We shall present these results while sketching the theory of τ-critical graphs in Section 12.1.

We next turn to the problem of describing the vertex packing polytope, that is, the convex hull of incidence vectors of independent points in a graph. Again, we cannot hope to obtain a complete description of the facets or even a list of the inequalities describing the vertex packing polytope (as, for example, Edmonds' Theorem 7.3.1 provides in the case of line graphs). But we shall be able to describe several nice classes of facets. Given a class of facets, we may ask: for what graphs is this subclass already sufficient to describe the vertex packing polytope? Interesting classes of graphs can be obtained in this way, most notably the class of *perfect* graphs. (See Section 12.2.) In view of the remarks above, this also yields information about $PC(G)$ as promised in Section 7.2.

Yet another way to view the vertex packing problem is to observe that it is equivalent to the matching problem for hypergraphs. In fact, given a **hypergraph**, a matching (that is, a set of disjoint lines) in this hypergraph corresponds to a set of independent points in its intersection graph. Since every graph is the intersection graph of a hypergraph (more specifically, of the hypergraph formed by using the lines of the graph as points and the stars of the points of the graph as lines), in this sense the matching problem for hypergraphs is equivalent to the vertex packing problem. But viewing hypergraphs as generalizations of graphs suggests not only some new approaches, but some new difficulties as well, and these are surveyed in Section 12.3.

Independent lines of a graph G correspond to independent points in its line graph $L(G)$ and vice versa. Hence the matching problem for G is equivalent to the vertex packing problem in $L(G)$. In this sense, all results on matchings discussed in this book could be rephrased as results concerning vertex packing in a special class of graphs, namely line graphs. This suggests the following method of attack on the vertex packing problem: find classes of graphs more general than line graphs for which the vertex packing problem is still polynomially solvable. It turns out that there is a very natural superclass of line graphs of this kind, namely the so-called claw-free graphs, for which the vertex packing problem can be solved in polynomial time. (A graph G is said to be **claw-free** if it contains no induced subgraph isomorphic to $K_{1,3}$.) This important result of Minty (1980) and Sbihi (1980) will be the main topic of Section 12.4.

12.1. Critical Graphs

An often-used approach to graph-theoretic problems is to study those graphs which have a given property, but for which no proper subgraph has the property. Often these subgraphs can be characterized and this provides a powerful tool for the study of the property at hand. For example, the graphs minimal with respect to having chromatic number 3 are just the odd cycles.

Let us call a graph G **τ-critical** if $\tau(G') < \tau(G)$ for every proper subgraph G' of G. Note that this implies that G has no isolated points. (Most results in this section are somewhat easier to state if they are formulated in terms of the point covering number $\tau(G)$. By the discussion in the introduction to this chapter, this is clearly equivalent to stating these results in terms of the vertex packing number (i.e., the independence number) $\alpha(G)$.) We shall also say that a line e or a point v of the graph G is **critical** if $\tau(G-e) < \tau(G)$ or $\tau(G-v) < \tau(G)$, respectively. It is clear that a graph G is τ-critical if and only if every point and line of it is critical.

As discussed in the introduction to this chapter, there is no complete description of the structure of τ-critical graphs. However, these graphs do have a number of important structural properties, which have interesting consequences for the point covering number in general.

A few small τ-critical graphs are shown in Figure 12.1.1.

Let us start with some trivial observations.

12.1.1. LEMMA. *A graph is τ-critical if and only if it contains no isolated points and all lines are critical.* ∎

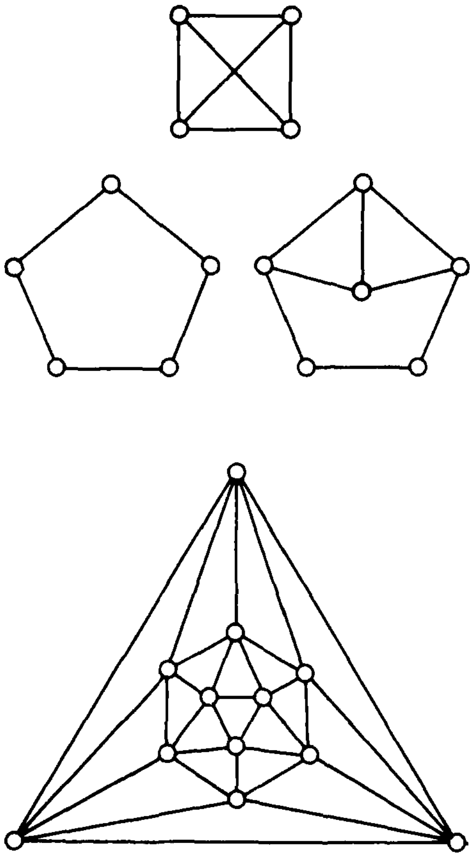


Figure 12.1.1. Some τ-critical graphs

12.1.2. LEMMA. *Let xy be a line of a τ-critical graph G. Then G has a minimum point cover containing x, but not y, and also one containing y, but not x. If, in addition, xy is not a connected component of G, then G has a minimum point cover containing both x and y.* ∎

The proofs of these results are straightforward and left to the reader. Note that Lemma 12.1.2 implies that the intersection of all minimum point covers and the intersection of all maximum independent sets in a τ-critical graph are each empty. The first lemma shows that a τ-critical graph has no points of degree 0. Our next statement describes points of degree 1.

12.1.3. LEMMA. *If a point x of a τ-critical graph G has degree 1, then x and its neighbor y form a connected component of G.*

PROOF. Suppose not. Then by Lemma 12.1.2, there is a minimum point cover T containing both x and y. But then $T - x$ is still a point cover, which is a contradiction. ∎

Next we discuss points of degree 2. This is more complicated, but in a sense even these points can be eliminated from consideration.

12.1.4. LEMMA. *Let G be a τ-critical graph and x, a point of degree 2 in G. Let y and z be the neighbors of x. If y and z are adjacent, then $\{x,y,z\}$ induces a connected component of G. If y and z are not adjacent, then no point different from x is adjacent to both of them and furthermore, if we contract the lines xy and xz, the resulting graph G' is τ-critical.*

Conversely, suppose G' is a τ-critical graph and w any point of G'. Split w into two points y and z, each of degree at least 1, create a new point x and connect it to both y and z. Then the resulting graph is τ-critical.

PROOF. I. Suppose that G is τ-critical, x is the point of degree 2 and y and z are the two neighbors of x. Also suppose that $yz \in E(G)$, and that $\{x,y,z\}$ does not induce a connected component; that is, suppose there is a fourth point u adjacent to, say, y. Then $G - uy$ has a $(\tau(G)-1)$-element point cover T. Clearly $y, u \notin T$ and therefore $x, z \in T$. But then $(T - x) \cup y$ is a $(\tau(G)-1)$-element point cover of G, a contradiction.

A similar argument shows that if y and z are non-adjacent, then no point other than x is adjacent to both of them.

We show next that $\tau(G') \geq \tau(G) - 1$. Let T be any point cover of G'. Let w denote the image of x (and y and z) in G'. If $w \notin T$, then $T \cup x$ is a point cover of G. If $w \in T$, then $(T - w) \cup y \cup z$ is a point cover in G. In both cases we see that $|T| + 1 \geq \tau(G)$, and hence $\tau(G') \geq \tau(G) - 1$.

Consider now any line e of G'. Then e may be viewed also as a line of G. By definition, $G - e$ has a point cover S with $|S| < \tau(G)$. If $x \in S$ then $S - x$ is a point cover of G'. If $x \notin S$ we must have $y, z \in S$, and then $S - y - z \cup w$ is a point cover of G'. In both cases we see that $\tau(G' - e) \leq |S| - 1 < \tau(G) - 1 \leq \tau(G')$. Thus G' is τ-critical as claimed.

II. The converse half of the proof follows by essentially the same arguments, which we omit. ∎

It is worth-while to consider the special case when x and y are both of degree 2 in G. Then the above lemma implies that either they belong to a component which is a triangle, or we can replace this path of length 3 by a single line and preserve τ-criticality. Conversely, the second half of the above lemma implies, in particular, that subdivision of a line by two points preserves τ-criticality.

This construction is a special case of a more general one, which splices two τ-critical graphs together or, conversely, describes those τ-critical graphs which are not 3-connected. This construction was discovered independently by Gallai, Plummer and Wessel (see Wessel (1970)). We shall state this result without proof, because we shall not need it in the sequel, and the proof is somewhat tedious.

12.1.5. LEMMA. *Let G_1 and G_2 be point-disjoint τ-critical graphs, suppose $y_1z_1 \in E(G_1)$, and suppose $w \in V(G_2)$. Let us split w into two points y_2 and z_2 of degree at least 1. Identify y_1 with y_2 and z_1 with z_2. Then the resulting graph is τ-critical. Conversely, let G be a τ-critical graph and suppose $\{x, y\} \subseteq V(G)$ is a cutset of G. Then G can be obtained by the above construction from two smaller τ-critical graphs so that x and y are the two points arising by identification.* ∎

Now we come to another very useful result on the structure of τ-critical graphs. Andrásfai (1967) proved that any two adjacent lines of a τ-critical graph are contained in an odd cycle and Beineke, Harary and Plummer (1967) proved that every two adjacent lines of a τ-critical graph are contained in a chordless cycle. The following theorem of Berge (1972) gives the natural common generalization of these two results.

12.1.6. THEOREM. *Every two adjacent lines of a τ-critical graph are contained in a chordless odd cycle.*

PROOF. Let xy_1 and xy_2 be two adjacent lines. Let T_i be a minimum point cover of the graph $G - xy_i$ $(i = 1, 2)$. Clearly $|T_i| = \tau(G) - 1$, $x \notin T_i$ and $y_i \in T_{3-i} - T_i$. Let G' denote the subgraph induced by the set $(T_1 - T_2) \cup (T_2 - T_1)$. Evidently G' is bipartite with color-classes $T_1 - T_2$ and $T_2 - T_1$. It is also clear that the only neighbors of x in G' are y_1 and y_2.

We claim that y_1 and y_2 must belong to the same connected component of G'. If they did not, then $V(G') \cup x$ would induce a bipartite subgraph G'' with $|T_1 - T_2| + |T_2 - T_1| + 1 = 2(\tau(G) - |T_1 \cap T_2|) - 1$ points. So the smaller color-class of G'', say A, would contain at most $\tau(G) - |T_1 \cap T_2| - 1$ points. But then $A \cup (T_1 \cap T_2)$ would be a point cover of G of size at most $\tau(G) - 1$, a contradiction.

So y_1 and y_2 belong to the same connected component of G'. Let P be a shortest path in G' connecting y_1 and y_2. Then $P \cup xy_1 \cup xy_2$ is a chordless odd cycle. ∎

REMARK. The reader may have noticed the similarity between the above proof and the proof of König's Minimax Theorem 1.1.1. In fact, Theorem 12.1.6 easily implies König's Theorem (see Corollary 12.1.7 below).

Also note that in the proof we have used the criticality of lines xy_1 and xy_2 only. So the conclusion of Theorem 12.1.6 holds true for any two adjacent critical lines in any graph.

12.1.7. COROLLARY. *Every bipartite r-critical graph consists of a set of independent lines.* ∎

12.1.8. COROLLARY. *A r-critical graph has no cutpoint.* ∎

12.1.9. COROLLARY. *Let G be a connected r-critical graph. Then no (inclusionwise) minimal cutset of points in G induces a complete graph.*

PROOF. Suppose that a minimal cutset $A \subseteq V(G)$ induces a complete graph. If $x \in A$, then by the minimality of A, x has two neighbors y_1 and y_2 separated by A. Consider a chordless cycle C through xy_1 and xy_2. (Such exists by Theorem 12.1.6, although the parity of this cycle is irrelevant here.) The cycle C must meet A in a point $z \neq x$, since A separates y_1 and y_2. But then $xz \in E(G)$ since A induces a complete graph, and this contradicts the fact that C is chordless. ∎

In the next part of this section the following simple — but powerful — lemma will play an important role. This lemma is a slight extension of a result of Hajnal (1965). Note that it concerns arbitrary graphs (not only r-critical ones). Recall that $\Gamma(X) = \{y \in V(G) \mid y \text{ is adjacent to } x\}$ and that in general, $\Gamma(X) \cap X \neq \emptyset$. We defined and investigated the **surplus**, $\sigma(X)$, of a set X of points in a bipartite graph in Chapter 1. We now extend this definition to general graphs by letting $\sigma(X) = |\Gamma(X)| - |X|$, for all $X \subseteq V(G)$. The set function $\sigma(X)$ is then submodular; this follows just as in the case of bipartite graphs.

12.1.10. LEMMA. *Let A be a maximum independent set of points and X any independent set of points in the graph G. Then*

$$\sigma(X \cup A) \geq \sigma(A) \quad \text{and} \quad \sigma(X \cap A) \leq \sigma(X).$$

PROOF. Substituting for σ and rearranging terms, the first inequality can be written equivalently as follows:

$$|\Gamma(X \cup A) - \Gamma(A)| \geq |X - A|. \tag{12.1.1}$$

To prove this, note that $\Gamma(X \cup A) - \Gamma(A) \supseteq A \cap \Gamma(X)$. The set $A-(A\cap \Gamma(X))\cup(X-A)$ is independent, and so by the maximality of A, we must have $|A-(A\cap\Gamma(X))\cup(X-A)| \leq |A|$ and so $|A\cap\Gamma(X)| \geq |X-A|$. Thus $|\Gamma(X\cup A)-\Gamma(A)| \geq |A\cap\Gamma(X)| \geq |X-A|$, which proves (12.1.1) and thereby the first inequality of the lemma. The second inequality follows from the first and from the submodular inequality:

$$\sigma(X\cup A)+\sigma(X\cap A) \leq \sigma(X)+\sigma(A).$$

■

As a first application of this inequality, let us prove the following result of Hajnal (1965).

12.1.11. THEOREM. *Each r-critical graph has a perfect 2-matching.*

PROOF. By Corollary 6.1.5, it suffices to show that $\sigma(X) \geq 0$ for every independent set X. By applying Lemma 12.1.10 repeatedly for all maximum independent sets $A_1, \dots, A_N$, we obtain

$$\sigma(X) \geq \sigma(X\cap A_1\cap\cdots\cap A_N) = \sigma(\emptyset) = 0$$

by Lemma 12.1.2. ■

12.1.12. COROLLARY. *If G is r-critical then $r(G) \geq p/2$.*

PROOF. In fact, we need only observe that if X is a maximum independent set of points in G then $0 \leq \sigma(X) = r(G)-(p-r(G))$. ■

This corollary was proved first by Erdős and Gallai (1961). Its significance is that it suggests a measure of "complexity" for r-critical graphs. More precisely, let us define the **Gallai class number** of G, $\delta(G)$, by

$$\delta(G) = 2r(G)-p = r(G)-\alpha(G) = p-2\alpha(G).$$

Then $\delta(G) \geq 0$ and by Theorem 12.1.11, we also have $\delta(G) = 2r(G)-r_2(G)$.

The results that follow will show that the number $\delta(G)$ has a very close connection with the structure of a r-critical graph G. In fact, it turns out that for every fixed value of δ, the class of all r-critical graphs with $\delta(G)=\delta$ has, in a sense, a "finite basis". But before dealing with that, we obtain a number of other applications of the parameter $\delta(G)$. The following inequality is due to Surányi (1975) and to Lovász (1977).

12.1.13. THEOREM. *Let G be a τ-critical graph, X an independent set of points in G and suppose $x \in X$. Then*

$$\deg_G(x) \leq 1 + \sigma(X).$$

PROOF. Let $A_1, \ldots, A_M$ be all maximum independent sets in G containing x, and set $X' = X \cap A_1 \cap \cdots \cap A_M$. Then by Lemma 12.1.10, $\sigma(X') \leq \sigma(X)$. We now claim that

$$\Gamma(X' - x) \cap \Gamma(x) = \emptyset. \tag{12.1.2}$$

For suppose there is a point u such that $xu \in E(G)$ and also $yu \in E(G)$ for some $y \in X' - x$. Let T be a point cover of $G - xu$ of size $\tau(G) - 1$. Then $T \cup u$ is a minimum point cover of G and so $A = V(G) - T - u$ is a maximum independent set of G. But clearly, $x \in A$, whereas $y \notin A$, and this contradicts the definition of X'. This verifies equation (12.1.2).

Now (12.1.2) implies that $\sigma(X') = \sigma(x) + \sigma(X' - x)$, and so by Theorem 12.1.11, $\sigma(X') \geq \sigma(x) = \deg_G(x) - 1$. This proves the theorem. ∎

We obtain another important result of Hajnal (1965) as an easy corollary.

12.1.14. COROLLARY. *In a τ-critical graph G, every point has degree at most $\delta(G) + 1$.*

PROOF. Let x be any point of G. By Lemma 12.1.2, we can find a maximum independent set A containing x. But then by Theorem 12.1.13,

$$\deg_G(x) \leq 1 + \sigma(A) = 1 + \tau(G) - \alpha(G) = 1 + \delta(G).$$ ∎

12.1.15. EXERCISE. Determine all τ-critical graphs G which are $(\delta(G) + 1)$-regular.

12.1.16. EXERCISE. Prove that every τ-critical graph G in which $\delta(G) = \delta \geq 2$ contains at most $\delta + 2$ points with degree $\delta + 1$. (Surányi (1975)).

12.1.17. EXERCISE. Prove that every connected τ-critical graph G with $\delta(G) \geq 2$ has a spanning τ-critical subgraph H with $\delta(H) = \delta(G) - 2$. (Surányi (1975)).

12.1.18. EXERCISE. Let G be a τ-critical graph and $X \subsetneq V(G)$, an independent set of points in G which contains a point of degree $\sigma(X)+1$. Then X is contained in some maximum independent set.

As an application of Corollary 12.1.14, we derive an upper bound on the number of lines in a τ-critical graph. This bound was first obtained by Erdős, Hajnal and Moon (1964).

12.1.19. COROLLARY. *A τ-critical graph G has at most $\binom{\tau(G)+1}{2}$ lines.*

PROOF. By Corollary 12.1.14,

$$|E(G)| \leq p(\delta(G)+1)/2 = (2\tau(G) - \delta(G))(\delta(G)+1)/2.$$

The sum of the two factors in this last product is constant if $\tau(G)$ is fixed and $\delta(G)$ varies, so the product is largest if they are as nearly equal in size as possible. Hence

$$(2\tau(G) - \delta(G))(\delta(G)+1)/2 \leq (\tau(G)+1)\tau(G)/2 = \binom{\tau(G)+1}{2}. \quad \blacksquare$$

Note that equality holds in the above corollary if and only if G is complete.

The following extension of this corollary, due to Lovász (1977), will play an important role in what is to follow. The proof is substantially more involved, however; in particular it uses multilinear algebra. We do not give the proof here; the idea of the proof is illustrated by Exercises 12.1.22 and 12.1.23. For the complete proof and for the matroid-theoretic background, see Lovász (1977).

12.1.20. THEOREM. *Let G be a τ-critical graph and T any set of points covering all lines in G. Then T induces at most $\binom{|T|-\alpha(G)+1}{2}$ lines.* $\blacksquare$

Let G be any graph with $\tau(G) = t$. Then Corollary 12.1.12 implies that G has a subgraph G' with $\tau(G') = t$ and $|V(G')| \leq 2t$. Theorem 12.1.20 implies that G has a subgraph G'' with $\tau(G'') = t$ and $|E(G'')| \leq \binom{t+1}{2}$. But how can we find such subgraphs algorithmically? The series of exercises that follow will yield polynomial-time algorithms to accomplish these aims.

12.1.21. EXERCISE. Let G be a graph and let $A \subsetneq V(G)$ be an independent set of points which has negative surplus, but every proper subset of A has non-negative surplus. Suppose $x \in A$. Then show that $\tau(G-x) = \tau(G)$.

12.1.22. EXERCISE. Let G be a graph with $\tau(G) = t$, and let $\boldsymbol{a}_x \in \Re^t$ ($x \in V(G)$) be vectors such that any t of them are linearly independent. Let $\boldsymbol{a}_e = \boldsymbol{a}_u \boldsymbol{a}_v^{\mathrm{T}} + \boldsymbol{a}_v \boldsymbol{a}_u^{\mathrm{T}}$ for every line $e = uv$. Suppose that A_{e_0} is a linear combination of the matrices A_e ($e \neq e_0$). Show that $\tau(G - e_0) = t$.

12.1.23. EXERCISE. Use the results of the previous exercises to design a polynomial-time algorithm which accomplishes the following: Given a graph G with $\tau(G) = t$, find a subgraph G' such that $\tau(G') = t$, G' has a perfect 2-matching (and so $|V(G')| \leq 2t$), and G' has at most $\binom{t+1}{2}$ lines.

In the rest of this section we prove a theorem which will enable us to classify all τ-critical graphs with a given $\delta(G)$. This classification was conjectured by Gallai, obtained for $\delta = 1$ by Hajnal (1965), for $\delta = 2$ by Andrásfai (1967), for $\delta = 3$ by Surányi (1975) and later for all δ by Lovász (1978).

Let us begin with some preliminary remarks to simplify the description of the results. First, it is trivial that if G is a τ-critical graph with connected components $G_1, G_2, \ldots, G_k$, then these components are also τ-critical and $\tau(G) = \tau(G_1) + \cdots + \tau(G_k)$. This remark enables us to restrict our attention to connected τ-critical graphs. Moreover, we shall not allow loops or points of degree 1.

Second, suppose that a connected graph $G \neq K_3$ has a point of degree 2, and let G' be the graph obtained by contracting the two lines of G incident with this point. Then by Lemma 12.1.4, G is τ-critical if and only if G' is τ-critical, and a trivial calculation shows that $\tau(G') = \tau(G) - 1$. By this remark, we may restrict our attention only to τ-critical graphs in which all degrees are at least 3. Such a τ-critical graph will be called **basic**.

Next, let us settle the case of τ-critical graphs with $\delta \leq 2$. If $\delta = 0$, then by Corollary 12.1.14, every point has degree at most 1. The only connected graph with this property and with no point of degree 0 is K_2. If $\delta = 1$ then again by Corollary 12.1.14, every point has degree at most 2. By the remark above, the graph must be K_3. If $\delta = 2$, then yet again by Corollary 12.1.14 and the remark above, G must be a connected 3-regular graph. But then by Exercise 12.1.15, G has at most 4 points and so $G = K_4$.

At this point we might suspect for a moment that we will always have $G = K_{\delta+2}$ when $\delta(G) = \delta$, but as soon as $\delta > 2$, life becomes much more complicated! There are precisely four connected basic τ-critical graphs having $\delta(G) = 3$; they are shown in Figure 12.1.2.

However, the following theorem is true.

12.1.24. THEOREM. *For every $\delta \geq 2$, there are finitely many connected basic r-critical graphs with $\delta(G) = \delta$.*

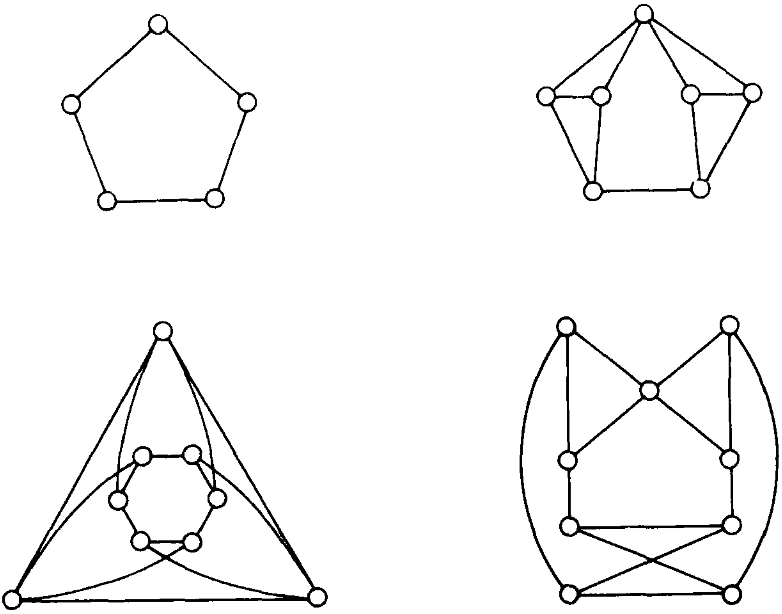


FIGURE 12.1.2. The basic r-critical graphs with $\delta = 3$

PROOF. Let G be a connected basic r-critical graph with $\delta(G) = \delta$. Let A be a maximum independent set in G and let $T = V(G) - A$. Let $T' = T \cap \Gamma(T)$. Then by Theorem 12.1.20, $|T'| \leq \delta(\delta+1)$.

Let $Z_1, \ldots, Z_N$ be all the subsets of T' meeting all lines induced by T. Then $r(G - Z_i) \geq r(G) - |Z_i|$ and hence $G - Z_i$ (which is a bipartite graph) contains a matching M_i of size $r(G) - |Z_i|$ for each i. Set $U_i = V(G) - V(M_i)$. Furthermore, let G' denote the bipartite graph obtained from G by deleting all lines spanned by T.

Thus we have a bipartite graph G' and subsets $U_1, \ldots, U_N \subseteq V(G')$ such that $G' - U_i$ contains a perfect matching for all $1 \leq i \leq N$. Trivially, $|U_i| = p - 2|M_i| = p - 2(r(G) - |Z_i|) = -\delta + 2|Z_i| \leq \delta(2\delta+1)$ and $N \leq 2^{\delta(\delta+1)}$.

We claim that if we delete any line e from G', there will be an i, $1 \leq i \leq N$, such that $G' - e - U_i$ has no perfect matching. In fact,

$r(G-e) < r(G)$ since G is r-critical and hence $G-e$ has a point cover S of size $r(G)-1$. Clearly $S \cap T$ covers all lines spanned by T and hence $S \cap T \supset Z_i$ for some i. But then $\nu(G'-e-U_i) = r(G'-e-U_i) \leq r(G'-e-Z_i) \leq |S-Z_i| = r(G)-1-|Z_i|$, and so $G'-e-U_i$ contains no perfect matching.

Thus we may apply Theorem 4.2.12 to conclude that G' has at most $2^{3\delta(\delta+1)}\delta^3(2\delta+1)^3$ points with degree ≥ 3. But then even if we reinsert the lines spanned by T, we find that G has at most $2^{3\delta(\delta+1)}\delta^3(2\delta+1)^3 + \delta(\delta+1)$ points of degree ≥ 3. But every point of G has degree ≥ 3 by hypothesis. This proves the theorem. ∎

REMARK. It seems that the number of such basic graphs grows very fast as δ increases. From the proof below we obtain a bound of $O(c^{\delta^2})$ for the maximum size of such a graph, but this is probably far from best possible. The largest known connected r-critical graph G with all degrees at least 3 and $\delta(G) = \delta$ has $O(\delta^2)$ points (Surányi (1976)).

12.1.25. COROLLARY. *Each connected r-critical graph G which has $\delta(G) = \delta$ has cyclomatic number at most $2^{8\delta^2}$.*

PROOF. We prove this corollary by induction on the number of points of G. If $\delta \leq 1$ then the assertion is trivial, so suppose that $\delta \geq 2$. If all points of G have degree ≥ 3, then $p \leq 2^{3\delta(\delta+1)}\delta^3(2\delta+1)^3 + \delta(\delta+1)$ as in the proof of Theorem 12.1.24, and by Corollary 12.1.14, $q \leq (\delta+1)p$, so the cyclomatic number

$$q - p + 1 \leq \delta p + 1 \leq 2^{8\delta^2}.$$

If G has points of degree 2, then upon contracting the two lines incident with any such point, we obtain another r-critical graph with the same δ and cyclomatic number, and so we are finished by induction. ∎

Another way to state Theorem 12.1.24 is that for a given $\delta \geq 1$, there is a finite number of graphs such that if we repeatedly insert points of degree two as in the second half of Lemma 12.1.4, then we obtain precisely the r-critical graphs G with $\delta(G) = \delta$. As we have remarked, the subdivision of a line by two new points is a special case of this operation. Thus the following statement seems stronger than Theorem 12.1.24, but it is actually another corollary of that theorem.

12.1.26. COROLLARY. *For every $\delta \geq 1$, there exist a finite number of graphs such that all connected τ-critical graphs with $\delta(G) = \delta$, and only these, arise from one of the given graphs by subdividing each line by an even number of points.*

PROOF. Consider the set $\mathcal{T}$ of all connected τ-critical graphs with $\delta(G) = \delta$ which do not contain two adjacent points of degree 2. It is obvious by Lemma 12.1.4 that if we subdivide each line of each graph in $\mathcal{T}$ by an even number of points in all possible ways, we obtain all connected τ-critical graphs with $\delta(G) = \delta$ and only those. What remains to be proved is that $\mathcal{T}$ is finite. But by Corollary 12.1.25, every graph in $\mathcal{T}$ has cyclomatic number at most $2^{8\delta^2}$. Let H be any graph in $\mathcal{T}$ and suppose H has q lines, p' points of degree 2 and p'' points of degree ≥ 3. Then we have

$$2q \geq 2p' + 3p''. \tag{12.1.5}$$

Furthermore, since no two points with degree 2 are adjacent, we also have

$$q \geq 2p'. \tag{12.1.6}$$

Dividing both sides of inequality (12.1.6) by 2 and adding the result to (12.1.5), we obtain

$$\tfrac{5}{2}q \geq 3(p' + p'') = 3|V(H)| = 3p.$$

and hence

$$q - p + 1 \geq \tfrac{6}{5}p - p + 1 = \tfrac{p}{5} + 1.$$

But again by Corollary 12.1.25, we also have $q - p + 1 < 2^{8\delta^2}$. So p is bounded and the finiteness of τ follows. ■

So, in fact, there are two finite basis results for τ-critical graphs with fixed $\delta(G)$, depending upon the operation allowed to generate the entire class. The operation of subdividing a line with two points is the more restricted and hence requires more basis graphs. However, the finiteness of either basis easily implies the finiteness of the other, as the argument in the above corollary shows.

12.2. Vertex Packing Polytopes

We have seen that linear programming provides us with a very useful tool in the study of matchings. Not only does it yield a compact and general way to state minimax results, but it leads to the discovery of unexpected connections between matching and other properties (for example, chromatic index) and motivates matching algorithms as well. In this section we try to extend this approach to independent sets of points. Since the vertex packing problem, being NP-complete, is inherently more difficult than matching, we cannot expect such a satisfactory solution as in the matching case. Nevertheless, this study will lead to interesting connections with other graph-theoretic parameters such as the chromatic number, to the definition and characterization of interesting classes of graphs such as perfect graphs, to algorithmic results, and more.

Let G be a graph and for each $A \subseteq V(G)$, let $\boldsymbol{q}^A$ denote the incidence vector of the set A in $\Re^{V(G)}$. Let us define the **vertex packing polytope** of G as the convex hull of the vectors $\boldsymbol{q}^A$, where A ranges over all independent subsets of $V(G)$. We denote this polytope by $VP(G)$. This polytope will be the main object of study in this section.

We try to follow the same approach as in Chapter 7 by attempting to find a description of $VP(G)$ as the solution set of a system of linear inequalities. No nice description of such a set of linear inequalities is known for a general graph, and, in fact, no such system is ever likely to be found! For if we could find such a system with the property that any given member of the system can be shown to belong to the system in polynomial time (a very modest requirement for a "nice" system), then $\alpha(G) \leq k$ could be proved in polynomial time for any graph for which it holds true. But this would mean that the property $\alpha(G) \leq k$ of the pair (G, k) would belong to NP. However, the property $\alpha(G) > k$ is NP-complete, and hence it would follow that NP = co-NP, contrary to the generally accepted hypothesis.

So, instead of trying to find *all* facets of $VP(G)$ (that is, instead of trying to describe $VP(G)$ as the solution set of a system of linear inequalities), we shall content ourselves with trying to find interesting sub-systems of these valid inequalities, and then try to ascertain for which graphs these are sufficient to describe $VP(G)$. This will lead us to several very interesting classes of graphs indeed.

The trivial set of inequalities

$$x_i \geq 0, \qquad \text{for all } i \in V(G) \tag{12.2.1}$$

will always be assumed. These are called the **non-negativity constraints.**

One way to express independence of a set of points is to assume that at most one of the two endpoints of any line is picked. Thus the following inequalities are valid for every vector $\boldsymbol{q}^A$ (where A is independent), and so for each vector $\boldsymbol{x}$ in $VP(G)$:

$$x_i + x_j \le 1 \qquad \text{for all } ij \in E(G). \tag{12.2.2}$$

We shall call these inequalities the line **constraints**. Every vector in $VP(G)$ clearly satisfies all line constraints, but the converse of this is not true in general. Just consider the case when G is a triangle. Then the vector $(1/2, 1/2, 1/2)$ satisfies all line constraints, but is *not* a convex combination of incidence vectors of independent sets.

On the other hand, we do have the following result.

12.2.1. LEMMA. *The non-negativity constraints* (12.2.1) *and the line constraints* (12.2.2) *suffice to describe the vertex packing polytope of a graph G if and only if G is bipartite.*

PROOF. I. Suppose G is bipartite, and let $\boldsymbol{x}$ be any point of the polyhedron P defined by (12.2.1) and (12.2.2). Let (A, B) be a bipartition of G and set

$$U = \{i \in A \mid 0 < x_i < 1\},$$
$$V = \{i \in B \mid 0 < x_i < 1\}.$$

Consider the vectors

$$\boldsymbol{x}' = \boldsymbol{x} - \epsilon \boldsymbol{q}^U + \epsilon \boldsymbol{q}^V,$$
$$\boldsymbol{x}'' = \boldsymbol{x} + \epsilon \boldsymbol{q}^U - \epsilon \boldsymbol{q}^V,$$

where ϵ is a very small positive number.

Then $\boldsymbol{x}'$ satisfies (12.2.1) since ϵ is small. Moreover, $\boldsymbol{x}'$ also satisfies (12.2.2), for suppose $ij \in E(G)$, $i \in A$ and $j \in B$. If $i \in U$ and $j \in V$, or if $i \notin U$ and $j \notin V$, then

$$x_i' + x_j' = x_i + x_j \le 1.$$

If $i \in U$ and $j \notin V$, then

$$x_i' + x_j' < x_i + x_j \le 1.$$

Finally, if $i \notin U$ and $j \in V$, then we must have $x_i = 0$ and so

$$x_i' + x_j' = x_j + \epsilon \le 1.$$

Thus $\boldsymbol{x}' \in P$ and similarly, $\boldsymbol{x}'' \in P$. But $\boldsymbol{x}$ is on the segment connecting $\boldsymbol{x}'$ to $\boldsymbol{x}''$, so it can be a vertex only if $\boldsymbol{x}' = \boldsymbol{x}'' = \boldsymbol{x}$, that is, if $\boldsymbol{x}$ is a 0–1 vector. But then $\boldsymbol{x}$ is the incidence vector of an independent set of points, and hence a vertex of $VP(G)$. This proves that $P = VP(G)$.

II. Suppose that G is non-bipartite, and let C be any odd cycle in G. Then $\boldsymbol{y} = \boldsymbol{q}^{V(C)}/2 \in P$, but $\boldsymbol{y} \notin VP(G)$, since every vector $\boldsymbol{x}$ in $VP(G)$ satisfies $\boldsymbol{x}^{\mathrm{T}} \cdot \mathbf{1} \le (|V(C)| - 1)/2$, but $\boldsymbol{y}$ does not. ∎

An inclusion-wise maximal complete subgraph of graph G is called a **clique** of G. Obviously, for non-negative vectors $\boldsymbol{x}$, the following set of inequalities implies the set (12.2.2).

$$\boldsymbol{x}(K) = \sum_{i \in V(K)} x_i \leq 1 \qquad \text{for all cliques } K. \tag{12.2.3}$$

These inequalities are called the **clique constraints**. (Of course a similar inequality is valid for every complete subgraph. However, such an inequality for a *non-maximal* one is implied by the inequality for the clique in which it lies.) When do these inequalities suffice to describe the polytope $VP(G)$? To answer this question, let us pause to discuss the class of graphs known as "perfect".

A graph G is said to be **perfect** if every induced subgraph G' of G satisfies

$$\chi(G') = \omega(G')$$

(where $\chi(G')$ denotes the chromatic number of G' and $\omega(G')$ denotes the size of a maximum clique in G'). Note that the inequality $\chi(G) \geq \omega(G)$ obviously holds for every graph G. Some perfect graphs are displayed in Figure 12.2.1.

Several results mentioned previously can be stated in terms of the perfection of certain graphs. Note that bipartite graphs are trivially perfect. König's Theorem 1.4.15 on the chromatic index of bipartite graphs implies that line graphs of bipartite graphs are perfect. König's Minimax Theorem 1.1.1 implies that the complement of the line graph of a bipartite graph is perfect. In fact, a matching in a graph G is just a clique in the complement of the line graph of G and if G is bipartite, then a coloring of the complement of its line graph corresponds to a partition of the lines of G into stars, which clearly corresponds to a point cover of G. Finally, König's Theorem on line covers (Corollary 1.1.6) is equivalent to the perfection of the complement of a bipartite graph.

Also note that Dilworth's Theorem 1.4.5 is equivalent to saying that the **comparability graph** of a poset (that is, the graph defined on a poset by connecting two elements by a line if and only if they are comparable) is perfect. The "dual" to Dilworth's result, (Theorem 1.4.9), is equivalent to the fact that the incomparability graph of a poset (that is the complement of the comparability graph) is perfect.

We shall not go into a detailed study of various classes of perfect. graphs, their properties and interrelationships; this could fill an entire volume. (As indeed it does; see the monograph by Golumbic (1980).

Also see Lovász (1983b).) We shall only discuss perfect graphs because of their vertex packing polytopes and we shall try to get to this aspect as quickly as we can.

We shall need one simple lemma about perfection.

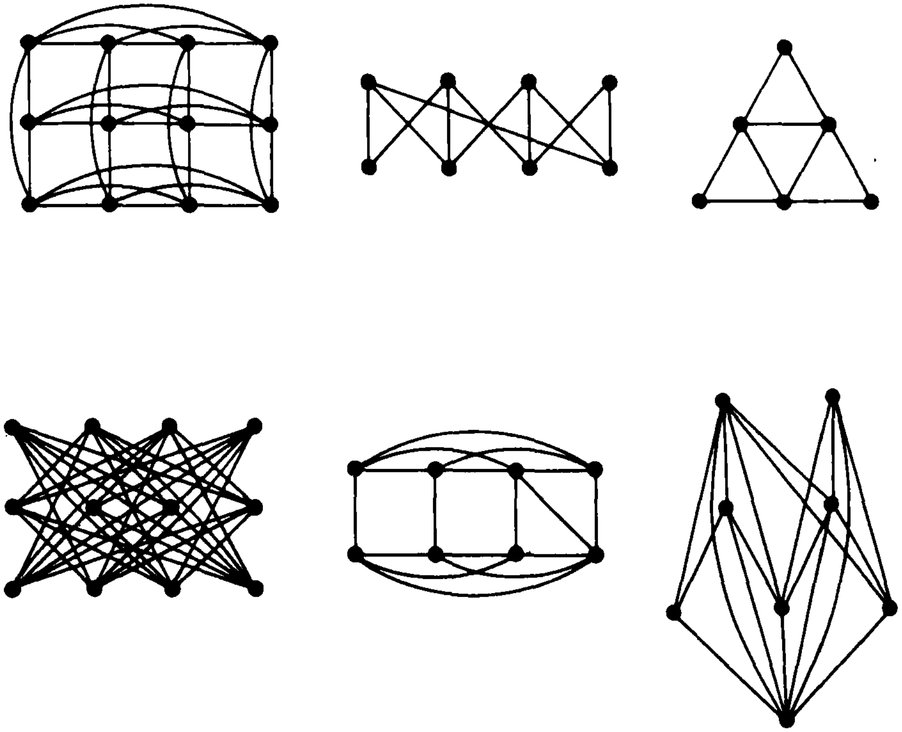


FIGURE 12.2.1. Some perfect graphs

12.2.2. LEMMA. (*The Replication Lemma*). *Let G be a perfect graph and suppose $x \in V(G)$. Create a new point x' and join it to x and to all neighbors of x. Then the resulting graph G' is perfect.*

PROOF. It suffices to show that $\chi(G') = \omega(G')$, since for the induced subgraphs of G' this follows similarly. We distinguish two cases.

Case 1. Suppose x is contained in some maximum clique of G. Then $\omega(G') = \omega(G) + 1$ and hence

$$\chi(G') \leq \chi(G) + 1 = \omega(G) + 1 = \omega(G').$$

This clearly implies that $\chi(G') = \omega(G')$.

Case 2. Now suppose x is not contained in any maximum clique of G. Consider any coloration of G with $\omega(G)$ colors and let A be the color class containing x. Then by the hypothesis of this case, $\omega(G-(A-x)) \leq \omega(G)-1$, since any maximum clique in G must meet A and the point of A it contains cannot be x.

So by the perfection of G, $G-(A-x)$ can be colored with $\omega(G)-1$ colors. Using a further color for the set $A - x \cup x'$ (which is clearly independent), we obtain an $\omega(G)$-coloration of G. ∎

We now come to the main result on vertex packing polytopes of perfect graphs. Stated in the form below, it includes results of Fulkerson (1971, 1972), Lovász (1972e) and Chvátal (1975).

12.2.3. THEOREM. *For any graph G, the following are equivalent:*

(i) G is perfect;

(ii) $\overline{G}$, the complement of G, is perfect;

(iii) $VP(G)$ is given by the non-negativity constraints (12.2.1) and the clique constraints (12.2.3);

(iv) $VP(\overline{G})$ is given by the non-negativity constraints (12.2.1) and the clique constraints (12.2.3) for $\overline{G}$.

PROOF. We will demonstrate the cycle of implications (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (iv) $\Rightarrow$ (i). It suffices to prove the first two of these, since the last two follow by interchanging the roles of G and $\overline{G}$.

I. (i) $\Rightarrow$ (iii). Let $\boldsymbol{x}$ be any vector satisfying (12.2.1) and (12.2.3); we want to prove that $\boldsymbol{x} \in VP(G)$. Trivially, we may assume that $\boldsymbol{x}$ is rational. But then of course there exists a positive integer N such that $\boldsymbol{y} = N\boldsymbol{x}$ is integral. By the non-negativity constraints, the entries y_i $(i \in V(G))$ are non-negative.

Let Y_i $(i \in V(G))$ be $|V(G)|$ disjoint sets, and let $y_i = |Y_i|$. Define a graph G' on $V(G') = \bigcup_{i \in V(G)} Y_i$ by joining every point of Y_i to every point of Y_j whenever $i = j$ or when $ij \in E(G)$. This graph G' arises from G by deleting and doubling points repeatedly, and hence by Lemma 12.2.2, G' is perfect.

Let K' be any clique in G', and let $K = \{i \in V(G) \mid K' \cap Y_i \neq \emptyset\}$. Trivially, K is a clique in G, and

$$|K'| \leq \sum_{i \in K} |Y_i| = \boldsymbol{y}(K) = N\boldsymbol{x}(K) \leq N,$$

by inequalities (12.2.3). Thus $\omega(G') \leq N$ and so G' can be colored with N colors $1, \ldots, N$. Let $A_t = \{i \in V(G) \mid Y_i \text{ has a point with color } t\}$.

Clearly, A_t is an independent set of points in G. Furthermore, each $i \in V(G)$ occurs in exactly $|Y_i| = y_i$ of the sets A_t, since Y_i induces a complete subgraph and hence its elements all have different colors. Thus $\sum_{t=1}^{N} \boldsymbol{q}_{A_t} = \boldsymbol{y}$, or $\frac{1}{N}\sum_{t=1}^{N} \boldsymbol{q}_{A_t} = \boldsymbol{x}$. This shows that $\boldsymbol{x}$ is indeed a convex combination of incidence vectors of independent sets of points in G.

II. (iii) $\Rightarrow$ (ii). It is easy to see that property (iii) is inherited by induced subgraphs, and therefore it suffices to show that $\chi(\overline{G}) = \omega(\overline{G})$. We will proceed by induction and hence assume that every proper induced subgraph of $\overline{G}$ is perfect.

Consider face F of $VP(G)$ defined by the linear equation $\boldsymbol{x}(V(G)) = \alpha(G)$. This face F does not contain $\mathbf{0}$ and hence it is contained in a facet F' of $VP(G)$ not containing $\mathbf{0}$. But by hypothesis (iii), $VP(G)$ is described by the inequalities (12.2.1) and (12.2.3), and so the facet F' is determined by one of these. Since F' does not contain $\mathbf{0}$, it follows that it is determined by one of the inequalities (12.2.3); that is, there is a clique C of G such that every point in F' satisfies $\boldsymbol{x}(C) = 1$. By the choice of F', the incidence vector of every maximum independent set belongs to F', and this implies that every maximum independent set meets C.

From this point on the argument is elementary. Since every maximum independent set meets C, we have $\alpha(G - V(C)) \le \alpha(G) - 1$, or equivalently, $\omega(\overline{G} - V(C)) \le \omega(\overline{G}) - 1$. By the induction hypothesis, $\overline{G} - V(C)$ can be colored with $\omega(\overline{G}) - 1$ colors. Using a new color for the points in C, we obtain an $\omega(G)$-coloration of $\overline{G}$. ■

Note the statement that (i) and (ii) are equivalent does not have anything to do with polyhedra! Restated, it says *the complement of a perfect graph is perfect.* This result is sometimes called the **Perfect Graph Theorem**. It was conjectured by Berge (1961) and proved by Lovász (1972e). On the other hand, all known proofs of this result do use, implicitly or explicitly, some ideas from polyhedral theory.

Theorem 12.2.3 does not give a full description of perfect graphs. What is the complexity of the property of perfection? It follows from the results of Lovász (1972f) that perfection is a property in co-NP; that is, non-perfection of a graph can be exhibited in polynomial time. On the other hand, it is not known whether perfection belongs to NP or maybe even to P!

We might hope that Theorem 12.2.3 could be used to design an algorithm for finding a maximum independent set of points in perfect graphs (just as Edmonds' description of the matching polytope can be used to design a matching algorithm). But when we try to use the Ellipsoid

Method to optimize over $VP(G)$ (cf. Section 9.3), it does not work! It only reduces the problem of maximizing over $VP(G)$ to the problem of maximizing over $VP(\overline{G})$, which is just as difficult! On the other hand, there *does* exist a polynomial-time algorithm to find a maximum independent set of points in a perfect graph, but it is based on a different kind of application of the Ellipsoid Method (see Grötschel, Lovász and Schrijver (1981, 1982)).

Notice that Theorem 12.2.3 can be used to obtain new proofs of some previous results. As mentioned above, bipartite graphs are trivially perfect and hence by Theorem 12.2.3 so are their complements. This fact, mentioned above, is just a restatement of König's Theorem on line covers in bigraphs (Corollary 1.1.6). This in turn, via the Gallai Identities (Lemmas 1.0.1 and 1.0.2), implies König's Minimax Theorem 1.1.1 which says that the complements of line graphs of bigraphs are perfect. But then again by Theorem 12.2.3, it follows that line graphs of bigraphs are perfect, which is equivalent to König's Line Coloring Theorem 1.4.15. In a similar fashion, we can derive Theorems 7.1.2 and 7.1.3. Moreover, Dilworth's Theorem 1.4.7 follows from its easy "dual", Theorem 1.4.11.

We now formulate a famous conjecture of Berge. Let us call a chordless cycle having length greater than 3 a **hole** and its complement an **antihole**.

12.2.4. CONJECTURE. (*The Strong Perfect Graph Conjecture*). *A graph is perfect if and only if it does not contain, as an induced subgraph, an odd hole or an odd antihole.*

While we cannot contribute anything to this conjecture, it motivates the next class of valid inequalities we consider. The following two classes of inequalities are trivially valid for all vectors in $VP(G)$:

$$x(H) \le (|H|-1)/2 \qquad (H \subseteq V(G) \text{ induces an odd hole}) \tag{12.2.4}$$

$$x(A) \le 2 \qquad (A \subseteq V(G) \text{ induces an odd antihole}). \tag{12.2.5}$$

We call these inequalities the **odd hole constraints and odd antihole constraints**, respectively. In view of the Strong Perfect Graph Conjecture, we might expect that the non-negativity, clique, odd hole, and odd antihole constraints might suffice to describe $VP(G)$. This is, however, not the case, and later on we shall see that $VP(G)$ has facets (that is, essential valid inequalities which *must* occur in any description of the polytope by linear inequalities) having a much more complicated structure.

So let us retreat a bit, regroup and study those graphs whose vertex packing polytopes can be described by some subset of the inequalities (12.2.1) – (12.2.5) other than the clique constraints (12.2.3). While this program is far from complete, some interesting results have already been obtained in this direction. For example, Boulala and Uhry (1979) have proved the following.

12.2.5. THEOREM. *Let G be a series-parallel graph. Then the polytope $VP(G)$ is described by the non-negativity, line, and odd hole constraints* (12.2.1), (12.2.2) *and* (12.2.4), *respectively.* ■

The reader may have noticed that all valid inequalities found so far are special cases of the following:

$$\boldsymbol{x}(U) \leq \alpha(G[U]), \qquad \text{for all } U \subseteq V(G). \tag{12.2.6}$$

These inequalities are called the **rank constraints**. Of course, our previously determined inequalities had the property that $\alpha(G[U])$ could easily be found. On the other hand, a general rank constraint may be useless because we cannot calculate the right hand side in polynomial time! But (the authors asked hopefully!) could it be that only "nice" rank constraints are essential? In fact, many rank constraints follow from others as in the case of a perfect graph G, where all rank constraints follow from the clique constraints. However, no complete description of the essential rank constraints is known, and, sadly, the following result of Chvátal (1975) shows that some essential rank constraints may be very ugly indeed!

12.2.6. THEOREM. *Let G be a connected τ-critical graph. Then the constraint*

$$\boldsymbol{x}(V(G)) \leq \alpha(G) \tag{12.2.7}$$

is essential for $VP(G)$ (that is, it determines a facet of $VP(G)$).

PROOF. Suppose that this constraint is not essential. Then the face F defined by it is not a facet and hence there is a facet F' containing F. Let

$$\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{x} \leq b \tag{12.2.8}$$

be an inequality defining such a facet F (where $\boldsymbol{a} = (a_i) \in \Re^{V(G)}$). Then every vector $\boldsymbol{x} \in VP(G)$ which gives equality in constraint (12.2.7) also gives equality in (12.2.8). In particular, if A is any maximum independent set in G, then the vector $\boldsymbol{q}_A$ gives equality in (12.2.8).

Since (12.2.7) and (12.2.8) define different faces, their left hand sides are not multiples of each other and hence not all a_i's are equal. Then since G is connected, we can find a line $ij \in E(G)$ such that $a_i \neq a_j$. Moreover, since G is τ-critical, graph $G-ij$ contains a set B of $\alpha(G)+1$

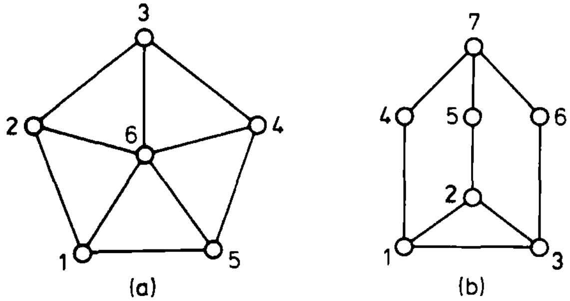


FIGURE 12.2.2.

independent points. But then $B-i$ and $B-j$ are maximum independent sets of points in G, and hence both $\boldsymbol{q}_{B-i}$ and $\boldsymbol{q}_{B-j}$ must satisfy (12.2.8) with equality. However,

$$\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{q}_{B-i} - \boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{q}_{B-j} = a_j - a_i \neq 0,$$

and we have a contradiction. ∎

But not even all the rank constraints are enough to describe $VP(G)$! Consider the 5-wheel, with its points labelled as in Figure 12.2.2(a). Then the inequality

$$x_1 + x_2 + x_3 + x_4 + x_5 + 2x_6 \leq 2 \qquad (12.2.9)$$

is a facet of the vertex packing polytope. (The reader should check this!)

Another example is shown in Figure 12.2.2(b). Here the inequality

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + 2x_7 \leq 3 \qquad (12.2.10)$$

is an essential inequality for the vertex packing polytope containing a coefficient 2.

There is an essential difference between these two examples. In the first, one feels that the odd hole constraint $x_1 + \cdots + x_5 \leq 2$ is somehow involved. Actually, this odd hole constraint determines a facet of the vertex packing polytope of the "rim". This situation can be described in general as follows. Let G be a graph and $i \in V(G)$. Then $VP(G-i)$ is clearly a facet of $VP(G)$ (determined by the non-negativity constraint $x_i \geq 0$). If F is any facet of $VP(G-i)$, then F is contained in *exactly two* facets of $VP(G)$. One of these is $VP(G-i)$; the other is called the **lifting** of F to $VP(G)$. If F is described by the inequality $\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{x} \leq b$ ($\boldsymbol{a} \in \Re^{V(G)-i}$), then its lifting to $VP(G)$ is described by $\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{x} + a_i x_i \leq b$, where

$$a_i = b - \max\{\boldsymbol{a}^{\mathrm{T}} \cdot \boldsymbol{q}_A \mid A \text{ is an independent subset of } V(G) - i - \Gamma(i)\}.$$

Clearly, inequality (12.2.9) arises from lifting the facet determined by the odd hole constraint for the rim of the wheel. However, inequality (12.2.10) is "genuine"; it does not arise by lifting.

The benefit derived from studying facets is not only that it leads to interesting classes of graphs (for example, perfect graphs), but that it can also be used to design linear programming algorithms in a manner similar to that discussed in Section 9.4 for matching. Of course there may well be other nice classes of facets as yet undiscovered. For any such class one may also ask if a given vector satisfies the constraints corresponding to facets in this class. This theory is much less developed then the theory of the matching polytope.

12.3. Hypergraph Matching

Another natural extension of the matching problem arises by considering matchings in hypergraphs. Recall that a **hypergraph** H is a finite set $V(H)$ (called the point set) together with a collection $E(H)$ of subsets (called lines) of $V(H)$. A hypergraph is called **r-uniform** if every line contains precisely r points. (Thus, 2-uniform hypergraphs are precisely the graphs without loops.) Many concepts and problems from graph theory (chromatic number, chromatic index, point cover, etc.) can be extended to hypergraphs in a natural way. We cannot go into much detail here; the interested reader is referred to the monograph of Berge (1973). (See also Lovász (1979c, Chapter 13.))

A subset of lines of a hypergraph is called a **matching** if the lines are disjoint. Extending the notation used for graphs, we denote the maximum number of lines in a matching by $\nu(H)$. A subset of $V(H)$ is called a **point cover** if it meets every line. The minimum size of a point cover is denoted by $\tau(H)$. Trivially,

$$\nu(H) \leq \tau(H). \tag{12.3.1}$$

No complete solution of the hypergraph matching problem (that is, the problem of determining $\nu(H)$) is known. In fact, this problem is NP-complete even for 3-uniform hypergraphs (see Karp (1972) or Garey and Johnson (1979))! In this section, we shall restrict ourselves to the problem of extending König's Theorem to hypergraphs, that is, to finding reasonable general conditions under which equality holds in (12.3.1).

Let us remark that the hypergraph matching problem is equivalent to the vertex packing problem for graphs. Since both are NP-complete, perhaps this is no surprise. But the constructions used in reducing one to the other are really very simple. First, let H be a hypergraph and define its **intersection graph** $L(H)$ as the simple graph whose points are the lines of H, two being adjacent if and only if they intersect. It is clear that

$$\nu(H) = \alpha(L(H)). \tag{12.3.2}$$

Second, let G be any graph. Define the **star hypergraph** $St(G)$ to be the hypergraph whose points are the lines of G, and whose lines are the stars of the points of G. It is then clear that

$$\alpha(G) = \nu(St(G)).$$

There is yet another way to associate with every graph a hypergraph such that the vertex packing problem for the graph is equivalent to the matching problem for the hypergraph. This hypergraph is called the **clique hypergraph** of G and will be denoted by $Cl(G)$. It is defined as follows. The points of $Cl(G)$ are the cliques in G, and for every point of G, the cliques of G containing this point form an line of $Cl(G)$. Then we have

$$\alpha(G) = \nu(Cl(G)).$$

The clique hypergraph cannot in general be constructed in polynomial time because a graph may have exponentially many cliques. Hence it is useless from the point of view of computational complexity. However, it is better suited for certain theoretical investigations than is the star hypergraph. One reason for this is the following nice property of clique hypergraphs. A hypergraph H is said to have the **Helly Property** if the following holds: whenever $E_1, \ldots, E_r \in E(H)$ satisfy $E_i \cap E_j \neq \emptyset$ for every $i, j \in \{1, \ldots, r\}$, then $E_1 \cap \cdots \cap E_r \neq \emptyset$. The following lemma is very easily verified.

12.3.1. LEMMA. *The clique hypergraph of any graph has the Helly Property.* ■

Our first move toward finding sufficient conditions for the König Property to hold is to extend the notion of 2-colorability of a graph. The first definition we introduce does not quite do the job, but it will motivate a subsequent more successful approach.

A hypergraph H is called **2-colorable**, if its points can be 2-colored so that every line contains points of both colors. It is immediately obvious that if a hypergraph has an line which is a singleton then it is not 2-colorable. Usually we do not want to consider this trivial obstacle to 2-colorability, and therefore we shall say that a hypergraph is **essentially 2-colorable** if upon deleting its 1-element lines, the remaining hypergraph is 2-colorable.

An essentially 2-colorable hypergraph, however, need not have the König Property; just consider the hypergraph on 4 points whose lines are all triples. A notion sharper than essential 2-colorability was introduced by Berge (1969, 1970). His idea was to extend the requirement of 2-colorability to certain "sub"-hypergraphs. There are various ways to define "sub"-hypergraphs of a hypergraph and we shall use two of these. Let H be any hypergraph. If H' is obtained from H be deleting lines, then H' is called a **partial hypergraph** of H. If $U \subseteq V(H)$, and H'' is the hypergraph with $V(H'') = U$ and whose lines are the non-empty sets of the form $E \cap U$ $(E \in E(H))$, then H'' is called the **restriction** of H to U.

Following Berge (1969, 1970), we call a hypergraph **balanced** if every restriction of it is essentially 2-colorable. It follows immediately that among all graphs precisely the bipartite ones are balanced. Just as bipartite graphs can be characterized by the absence of odd cycles, we can characterize balanced hypergraphs as follows. A **circuit** in a hypergraph is an alternating sequence $(x_1, E_1, x_2, E_2, \ldots, x_k, E_k)$ where $x_1, \ldots, x_k$ are distinct points, $E_1, \ldots, E_k$ are distinct lines, x_i, $x_{i+1} \in E_i$ $(i = 1, \ldots, k)$ and $x_{k+1} = x_1$. The number k is the **length** of the circuit. The circuit is called **unbalanced** if $E_i \cap \{x_1, \ldots, x_k\} = \{x_i, x_{i+1}\}$ for all $1 \le i \le k$, that is, if there is no incidence between its points and lines other than the trivial ones. The following characterization of balanced hypergraphs is due to Berge (1970).

12.3.2. THEOREM. *A hypergraph is balanced if and only if it does not contain an unbalanced odd circuit.*

PROOF. The "only if" part is trivial: if $(x_1, E_1, \ldots, x_k, E_k)$ is an unbalanced odd circuit, then the restriction of H to the set $\{x_1, \ldots, x_k\}$ cannot be essentially 2-colorable.

On the other hand, suppose H does not contain any unbalanced odd circuit. It suffices to show that H itself is essentially 2-colorable.

The proof is by induction on $|V(H)|$. We may assume then that every restriction of H to some set $U \subset V(H)$ is essentially 2-colorable. Let G be the graph on $V(H)$ formed by those lines of H which have exactly two elements. Since H does not contain any unbalanced odd circuits, G must be bipartite. Let G_0 be a connected component of G. (It may happen that $G_0 = G$ or that G_0 consists of a single isolated point.) Set $U = V(G) - V(G_0)$. We know by hypothesis that the restriction of H to U is essentially 2-colorable; let us 2-color it using red and blue. Also, G_0 is bipartite so let us 2-color G_0 using red and blue as well.

Thus we have colored all points of H. Let us verify that this 2-coloration is essentially good; that is, that every line E of H with more than one point meets both colors. If $|E \cap U| \geq 2$ then this is clear. So suppose that $|E \cap U| \leq 1$, and thus $|E \cap V(G_0)| \geq 1$. But it cannot happen that E contains one point from each of U and $V(G_0)$, since then E would be a line of G connecting G_0 to a point outside G_0, which is impossible since G_0 is a component of G. Thus $|E \cap V(G_0)| \geq 2$.

Suppose now that E is monochromatic, say red. Thus all points of E in $V(G_0)$ are red. Since G_0 is connected, it contains a path P connecting two points in E and such a path must have odd length. We may also assume that no interior point of P belongs to E, for otherwise we could consider an appropriate proper subpath of P instead of P. Then P together with E forms an unbalanced odd circuit. This contradiction proves that no line with more than one point is monochromatic; that is, H is essentially 2-colorable. ■

The following important result on balanced hypergraphs is due to Berge and Las Vergnas (1970):

12.3.3. THEOREM. *Every balanced hypergraph satisfies the König Property.*

PROOF. We imitate the proof of König's Theorem for bigraphs given in Section 1.1. Let H be a balanced hypergraph and let us remove lines from H as long as $\tau(H)$ remains unchanged. (Note that when lines are removed from H, the property of being balanced is preserved.) Let H' denote the hypergraph obtained in this way. Then, $\tau(H') = \tau(H)$, but $\tau(H' - E) < \tau(H)$ for every line E of H.

We claim that H' consists of disjoint lines. This will suffice, since then, trivially, the number of lines in H' is $\tau(H') = \tau(H)$, and hence we will have found $\tau(H)$ disjoint lines in H. This shows that H has the König Property.

Suppose, to the contrary, that H' has two lines E_1 and E_2 which have a point v in common. By the definition of H', the lines of the hypergraph $H' - E_i$ can be covered by fewer than $\tau(H)$ points. For $i = 1, 2$, let T_i be a set of $\tau(H) - 1$ points covering all lines of $H' - E_i$. Obviously, $v \notin T_1 \cup T_2$.

Consider the set $U = (T_1 \oplus T_2) \cup \{v\}$. Since E is balanced, the restriction H'' of H to U is essentially 2-colorable. Let (A, B) be an essential 2-coloration of H''. Without loss of generality, assume that $|A| \leq |B|$. Then $|A| \leq |U|/2$ and since U is odd, $|A| \leq (|U| - 1)/2$.

Next, consider the set $A \cup (T_1 \cap T_2)$. We claim that this set covers every line of H'. Let $E \in E(H')$. If $|E \cap U| \geq 2$, then $A \cap E \neq \varnothing$, since (A, B) is an essential 2-coloration of H''. So suppose that $|E \cap U| \leq 1$. Clearly, $E \neq E_1, E_2$. Then either $T_1 - T_2$ or $T_2 - T_1$ misses E and since both T_1 and T_2 must meet E, it follows that $T_1 \cap T_2$ must meet E.

But then we obtain

$$\tau(H') \leq |A \cup (T_1 \cap T_2)| \leq (|U| - 1)/2 + |T_1 \cap T_2| = \tau(H) - 1,$$

which contradicts the definition of H'. ∎

Define the **chromatic index** $\chi_e(H)$ of a hypergraph H to be the least number of colors sufficient to color the lines of H so that no two lines with the same color have a point in common. Just as in the case of graphs, the maximum degree $\Delta(H)$ (that is, the maximum number of lines containing some one point) is a lower bound for the chromatic index. Then König's Theorem on the chromatic index of bipartite graphs may be extended to balanced hypergraphs.

12.3.4. THEOREM. *If H is a balanced hypergraph, $\chi_e(H) = \Delta(H)$.* ∎

This result can be obtained as a consequence of the following theorem (Lovász (1972e)) which is really just a restatement of the Perfect Graph Theorem.

12.3.5. THEOREM. *For every hypergraph H, the following are equivalent:*

(i) every partial hypergraph H' of H has $\nu(H') = \tau(H')$;
(ii) every partial hypergraph H' of H has $\chi_e(H') = \Delta(H')$;
(iii) $L(H)$ is perfect and H has the Helly Property.

PROOF. I. (i) $\Rightarrow$ (iii). First we show that H has the Helly Property. In fact, let H' be any collection of mutually intersecting lines of H. Then $\nu(H') = 1$ and hence by (i), $\tau(H') = 1$. But this means precisely that all the lines in H' have one point in common.

We claim now that $\overline{L(H)}$ is perfect. This will imply by Theorem 12.2.3 that the graph $L(H)$ is also perfect. To this end, observe that the induced subgraphs of $L(H)$ are just the intersection graphs of partial hypergraphs of H. Hence it suffices to show that $\chi(\overline{L(H)}) = \omega(\overline{L(H)})$, since the relation $\chi(G') = \omega(G')$ follows for every induced subgraph G' of $\overline{L(H)}$ similarly.

By equation (12.3.2), $\omega(\overline{L(H)}) = \alpha(L(H)) = \nu(H)$. Let U be a point cover of H with $\tau(H)$ points. For each $u \in U$, the set of lines of H containing u is independent in the graph $\overline{L(H)}$, and so we have $|U|$ independent sets covering all points of $\overline{L(H)}$. Hence $\chi(\overline{L(H)}) \le |U| = \tau(H) = \nu(H) = \omega(\overline{L(H)})$, which proves that $\overline{L(H)}$ is perfect. Hence the graph $L(H)$ is also perfect.

II. (iii) $\Rightarrow$ (ii). It suffices to show that $\chi_e(H) = \Delta(H)$. Any coloration of $L(H)$ corresponds trivially to a coloration of the lines of H and hence $\chi_e(H) = \chi(L(H))$. Consider any maximum clique in $L(H)$. This gives us $\omega(L(H))$ lines in H which mutually intersect. But since H has the Helly Property, it follows that these $\omega(L(H))$ lines have a point in common, and so $\Delta(H) \ge \omega(L(H)) = \chi(L(H)) = \chi_e(H)$. But then $\Delta(H) = \chi_e(H)$ clearly follows.

III. The implications (ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (i) follow by very similar arguments which we omit. ■

12.4. Vertex Packing in Claw-free Graphs

We conclude this book with an algorithm which solves the vertex-packing problem for **claw-free** graphs; i.e., for graphs having no induced subgraph isomorphic to $K_{1,3}$. The first two polynomial-time algorithms to solve this problem were obtained independently by Minty (1980) and by Sbihi (1980). The algorithm given here is different from both of these, however, although it was inspired by Minty's algorithm in the following sense. As we have remarked in the introduction to this chapter, the class of claw-free graphs includes all line graphs, and so the vertex-packing problem for claw-free graphs includes the matching problem. Minty's algorithm turns this relation around and reduces the vertex-packing problem in a claw-free graph to a matching problem by constructing an auxiliary graph, which he calls the "Edmonds graph". Here we develop a more immediate reduction of the vertex-packing problem for claw-free graphs to a matching problem.

The algorithm of Sbihi solves the problem directly without reference to matchings, but it is not too easy to follow and we shall not discuss it

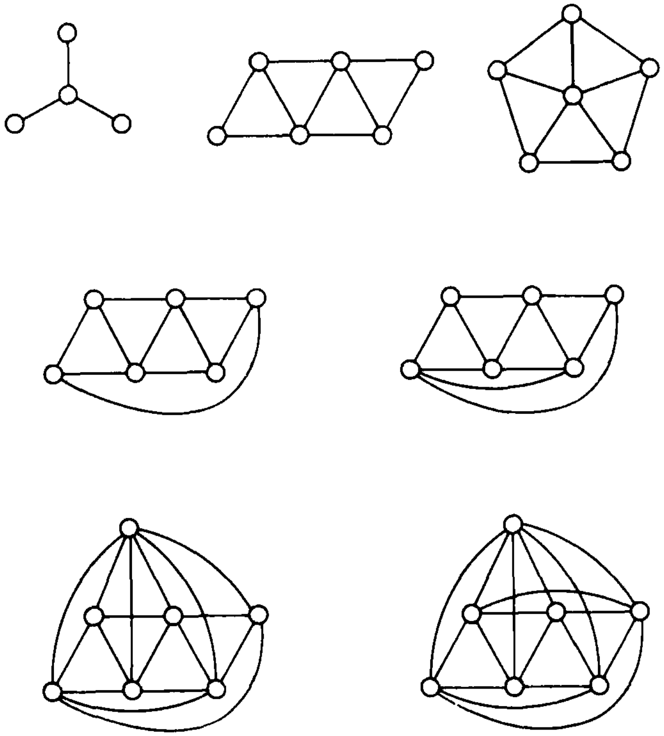


FIGURE 12.4.1. Excluded subgraphs for line graphs

here. Neither do we go into the *weighted* version of the problem, although Minty extends his method to this more general problem as well.

It is quite easy to construct an example which shows that not every claw-free graph is the line graph of some graph. In fact, to characterize line graphs of graphs, we must exclude six further graphs besides $K_{1,3}$. (See Hemminger (1971) and Bermond and Meyer (1973).) A similar characterization of line graphs of *simple* graphs is due to Beineke (1968, 1970) and, independently, to Robertson (unpublished).

12.4.1. THEOREM. *A graph G is the line graph of some graph if and only if it does not contain any of the graphs in Figure* 12.4.1 *as an induced subgraph.* ∎

It follows from the result of van Rooij and Wilf (1965) that it can be decided in polynomial time whether or not a given graph is a line graph. Moreover, given a line graph $L(G)$, it is not difficult to construct G in polynomial time (Roussopolis, (1973)). (See also Lehot (1974).) So for graphs not containing any of the graphs in Figure 12.4.1 as an induced subgraph, the vertex packing problem can be reduced to the matching problem. We shall show that such a reduction is still possible if only the first of these graphs is excluded.

For later reference, let us formulate another (easy) characterization of line graphs due to Krausz (1943).

12.4.2. EXERCISE. A graph G is the line graph of some graph if and only if G can be written as the union of complete subgraphs so that no point of G belongs to more than two of these complete subgraphs.

Now let us sketch the idea of the forthcoming algorithm. First, we shall describe two "local" configurations such that if either one of these occurs, then we can easily reduce the problem to a single smaller graph. Searching for such "local" configurations and carrying out the reductions can be done in polynomial time, and hence these reductions lead to a polynomial-time algorithm. (The reader should note that it is trivial to reduce the problem to determining the independence number of *two* smaller graphs. For example, $\alpha(G) = \max\{\alpha(G-v), \alpha(G-\Gamma(v))\}$. Such a reduction, however, leads to an exponential time algorithm as the reader may easily verify!)

On the other hand, if we cannot carry out either one of these reductions, the graph we are left with will prove to already be a line graph! So in this case we can fall back upon any of the polynomial-time matching algorithms in Chapter 9.

So let us start with a description of the reduction procedures we are going to use. We shall need the following notation. Let G be a graph and suppose $X \subseteq V(G)$. We set $N(X) = \Gamma(X) - X$. So $N(X) = \Gamma(X)$ if and only if X is independent. In particular, $N(v) = \Gamma(v)$ for each point v. Let $N_2(v)$ denote the set of points at distance 2 from point v. So $N_2(v) = N(v \cup N(v))$. We also set $\alpha(X) = \alpha(G[X])$. If four points a, b, c, d induce a claw in G with center a, then we shall denote this claw by $\{a; b, c, d\}$. A point v of a claw-free graph will be called **regular** if $N(v)$ can be partitioned into two complete graphs and **irregular**, otherwise. Our first lemma describes a way to eliminate irregular points. (Note that by Exercise 12.4.2, all points in a line graph are regular, so this procedure will align the structure of our claw-free graph more closely with that of a line graph.)

12.4.3. LEMMA. *Let G be a claw-free graph and a, an irregular point of G. Let Y be the set of those points $y \in N_2(a)$ for which $N(a) - N(y)$ induces a complete graph in G. Let G' be the graph obtained from G by*

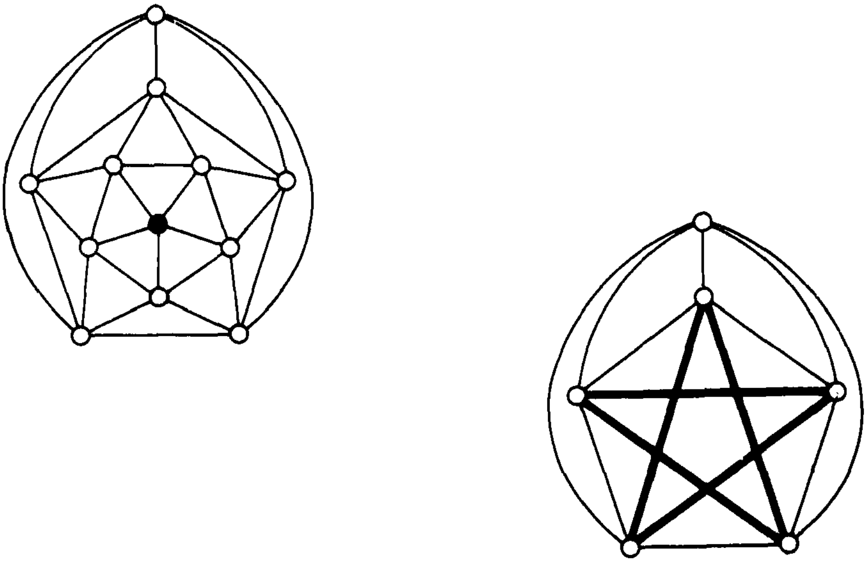


FIGURE 12.4.2. Reduction of an irregular point

deleting $a \cup N(a) \cup Y$ and by joining every two (as yet non-adjacent) points of $N_2(a) - Y$ with a new line. Then G' is claw-free and $\alpha(G') = \alpha(G) - 2$.

This procedure for reducing irregular graphs is illustrated in Figure 12.4.2.

PROOF. The hypothesis that $N(a)$ cannot be covered by two complete subgraphs means that $\overline{G}[N(a)]$ is not bipartite, and so it contains an odd cycle $C = x_0 x_1 \cdots x_{2k}$. Since G is claw-free, we must have $k \geq 2$. If we choose k as small as possible, cycle C will be chordless.

Claim 1. Every point $v \in N(a) - C$ is adjacent to at least $2k-1$ points on C. Moreover, if C contains two points x_i and x_j neither of which is adjacent to v, then $j - i \equiv \pm 2 (\bmod 2k+1)$.

It suffices to verify the second assertion, since it clearly implies the first. But if x_i and x_j are not at distance 2 when traversing cycle C, then we can find an odd cycle in $\overline{G}[N(a)]$ which is shorter than C, contrary to the choice of C.

Claim 2. Every $v \in N_2(a)$ is adjacent to some point of cycle C.

To see this, simply choose any point $u \in N(a) \cap N(v)$, and assume that $u \notin C$. By Claim 1, u must be adjacent to two consecutive points on C, say x_0 and x_1. But then the fact that $\{u; v, x_0, x_1\}$ is not a claw implies that v is adjacent to one of x_0 and x_1.

Claim 3. Every $v \in N_2(a)$ is adjacent to at least k points of C.

We already know that v is adjacent to some point of C, so let, say, x_0 be one of these. Then for any $1 \le t \le k-1$, the quadruple $\{x_0; v, x_{2t}, x_{2t+1}\}$ is not a claw and hence v is adjacent to at least one of x_{2t} and x_{2t+1}. This yields $k-1$ further neighbors of v on C.

Claim 4. $\alpha(N_2(a)) \le 2$.

Suppose, to the contrary, that u, v and w are independent points in $N_2(a)$. Then by Claim 3, each of them is adjacent to at least k points on C. Since $3k > 2k+1$, there is a point $x \in C$ which is adjacent to at least two of u, v and w, say to v and w. But then $\{x_i; a, v, w\}$ is a claw and we have a contradiction.

Claim 5. $\alpha(a \cup N(a) \cup N_2(a)) \le 3$.

To see this, assume that u, v, w and z are four independent points in $a \cup N(a) \cup N_2(a)$. Then, by Claim 4, at most two of u, v, w and z belong to $N_2(a)$. Since, trivially, $\alpha(N(a)) \le 2$, clearly two of them must belong to $N(a)$ and two of them to $N_2(a)$. Let us assume that $u, v \in N(a)$ and $w, z \in N_2(a)$. Let us choose C so that it contains as many of u and v as possible. By the same argument as above, the points w and z cannot have a common neighbor on C. So by Claim 3, every point on C, with at most one exception, is adjacent to one of them. Hence one of u and v, say u, is not on C.

If C has two points not adjacent to u then these are x_{i-1} and x_{i+1} for some i by Claim 1 (where the indices are taken modulo $2k+1$). But then $C' = C - x_i \cup u$ is also a minimum length odd cycle in $\overline{G}[N(a)]$ and since $v \ne x_i$ (because x_i is adjacent to u), it also follows that C' contains more of the two points u and v than C, a contradiction. So we know that u is adjacent to all but at most one point of C.

Now if $v \notin C$, then by the same argument all but at most one point of C are also adjacent to v. So we can find a point on C which is adjacent to u, v and one of w and z. On the other hand, if $v \in C$, then v must be the unique point on C not adjacent to u, and so again we find a point on C adjacent to u, v and one of w and z. In both cases we have a claw, which is a contradiction. This proves Claim 5.

Claim 6. $\alpha(a \cup N(a) \cup Y) \le 2$.

By the definition of Y, this will follow if we show that the set Y induces a complete subgraph of G. Every point in Y is adjacent to at least $k+1$ points of C trivially, and hence any two points of Y have a common neighbor in C. But from this it follows that they are adjacent just as above.

We can now finish the proof of the lemma fairly easily. It is straightforward to verify that G' is claw-free.

Next we show that $\alpha(G') \geq \alpha(G) - 2$. Let A be any maximum independent set in G. Then A contains at most 2 points from $a \cup N(a) \cup Y$ by Claim 6, and hence, $A' = A - a - N(a) - Y$ has cardinality at least $\alpha(G) - 2$. So if A' is independent in G', we are done. If A' is not independent in G', then it must contain at least two points of $N_2(a)$, say u and v. But by Claim 4, it contains only these two points of $N_2(a)$ and by Claim 5, set A then contains at most one point of $a \cup N(a)$, and so A' has cardinality $\alpha(G) - 1$. But then set $A' - u$ is an independent set in G' with cardinality $\alpha(G) - 2$.

Finally, we show that $\alpha(G) \geq \alpha(G') + 2$. To this end, let B be any maximum independent set of points in G'. If $B \cap N_2(a) = \emptyset$, then $B \cup \{x_0, x_1\}$ is an independent set in G of size $\alpha(G') + 2$. If B contains a point v of $N_2(a)$, then clearly this is the only point of $B \cap N_2(a)$ by the definition of G'. Furthermore, since $v \notin Y$, the set $N(a) - N(v)$ contains two non-adjacent points x and y. But then $B \cup \{x, y\}$ is an independent set in G of size $\alpha(G') + 2$. This completes the proof of the lemma. ∎

Our next lemma describes a similar, but simpler, reduction procedure. The proof of this lemma is analogous to that of the last part of the proof of the preceding lemma and is therefore left to the reader.

Let Q be any clique in the graph G. We say that Q is **reducible** if $\alpha(N(Q)) \leq 2$.

12.4.4. LEMMA. *Let Q be any reducible clique in the claw-free graph G. Let G' denote the graph obtained from G by deleting the points of Q and connecting two as yet non-adjacent points u and v of $N(Q)$ by a line if and only if $Q \subseteq N(u) \cup N(v)$. Then G' is claw-free and $\alpha(G') = \alpha(G) - 1$.* ∎

The clique reduction procedure of Lemma 12.4.4 is illustrated in Figure 12.4.3.

It is obvious how to check whether a *given* clique is reducible. It is much more difficult to check if *there exists* any reducible clique! This existence problem will not trouble us however, for in the special case when we have to apply the clique reduction procedure, the reducible clique will arise in a simple manner.

What happens if neither of the reduction principles formulated in the previous lemmas can be applied? The following theorem shows that then, in fact, we have returned to a friendly and familiar case!

12.4.5. THEOREM. *Let G be a graph such that every point of G is contained in two irreducible cliques which cover all neighbors of the point. Then G is a line graph.*

PROOF. The hypothesis implies that G is claw-free and that it does not contain an induced 5-wheel. (In fact, we shall need only these two consequences of the hypothesis.)

Let Q be any clique in G. We say that two neighbors u and v of Q are distant (with respect to Q), if they are non-adjacent and $Q \cap N(u) \cap N(v) = \emptyset$.

Claim 1. Let u and v be two non-adjacent neighbors of a clique Q which are non-distant. Then $Q \subseteq N(u) \cup N(v)$. In fact, if there were a point $z \in Q - N(u) - N(v)$, and also a point $y \in Q \cap N(u) \cap N(v)$, then $\{y; u, v, z\}$ would be a claw.

Claim 2. Let u, v and w be three independent neighbors of a clique Q. If some two of them are distant, then any two of them are distant. For, if u and v are distant, but, say, u and w are not, then $Q \subseteq N(u) \cup N(w)$ by Claim 1, and hence $N(v) \cap Q \subseteq Q - N(u) \subseteq N(w) \cap Q$. So v and w are not distant and hence once again by Claim 1, $Q \subseteq N(v) \cup N(w)$. But since $N(u) \cap N(v) \cap Q = \emptyset$, this implies that $Q \subseteq N(v)$, which is impossible since Q is a clique.

Claim 3. Let u_1, u_2 and u_3 be three independent neighbors of a clique Q, and let v be a neighbor of Q not adjacent to u_1. Then v and u_1 are distant.

To see this, assume that v and u_1 are not distant. Then there is a point $x \in Q \cap N(v) \cap N(u_1)$. Considering the quadruple $\{v; x, u_2, u_3\}$, we see that v cannot be adjacent to both u_2 and u_3. Without loss of generality, let us assume that v is not adjacent to u_2. Then u_1, u_2 and v are three independent neighbors of Q such that v and u_1 are non-distant, but u_1 and u_2 are distant. But this contradicts Claim 2.

We say that the clique Q is **normal** if it has three neighbors which are mutually distant.

Claim 4. Let v_1 and v_2 be two independent neighbors of a normal clique Q. Then v_1 and v_2 are distant.

To prove this claim, consider three mutually distant neighbors u_1, u_2 and u_3 of Q. We may assume that they are distinct from v_1 and v_2, or else we are done by Claim 3. By claw-freeness, each v_i is adjacent to at most two of the u_j's. But we may also assume that each v_i is adjacent to two of the u_j's; for if not, say if v_1 were non-adjacent to u_2 and u_3, then u_2, u_3 and v_1 would be three independent neighbors of Q, of which two would be distant. So by Claim 2, they are mutually distant and hence

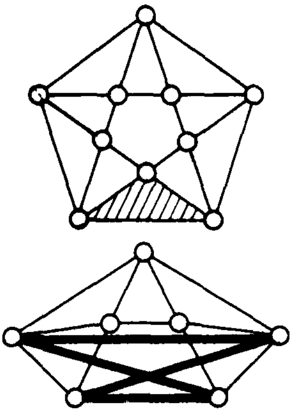


FIGURE 12.4.3. Reduction of a clique

by Claim 3, v_1 and v_2 are also distant. Furthermore, each u_j is adjacent to at least one of the points v_i. For if, say, u_1 were independent from v_1 and v_2, then u_1, v_1 and v_2 would be three independent neighbors of Q, of which u_1 and v_1 are distant by Claim 3. But then v_1 and v_2 would also be distant by Claim 2.

Thus we may assume that v_1 is adjacent to u_1 and u_3, and v_2 is adjacent to u_2 and u_3. It follows by Claim 3 that u_1 and v_2 are distant and so are u_2 and v_1 Suppose now that v_1 and v_2 are not distant. Then there is a point $x \in Q \cap N(v_1) \cap N(v_2)$. If we then consider the quadruple $\{v_1; u_1, u_3, x\}$, we conclude that one of u_1 and u_3 must be adjacent to x. But since u_1 and v_2 are distant, we conclude that u_3 must be adjacent to x.

Next, let y_i be any neighbor of u_i in Q. Considering the quadruple $\{x; v_1, v_2, y_1\}$, we see that y_1 must be adjacent to one of v_1 and v_2. But y_1 cannot be adjacent to v_2, since u_1 and v_2 are distant. So y_1 is adjacent to v_1. Similarly, y_2 is adjacent to v_2. But now the points y_1, v_1, y_2, u_3 and v_2 induce a pentagon which lies entirely in $N(x)$. This is impossible, since by hypothesis, $N(x)$ can be covered by two complete graphs. Thus Claim 4 is proved.

Claim 5. Let a be any point and let Q_1 and Q_2 be irreducible cliques containing a and covering $N(a)$. Then Q_1 and Q_2 are normal. In fact, Q_1 has three independent neighbors u, v and w. At least two of these, say u and v, are not contained in Q_2. But then $a \notin N(u) \cup N(v)$ and so $Q \not\subseteq N(u) \cup N(v)$. Hence u and v are distant by Claim 1 and therefore, u, v and w are mutually distant by Claim 2.

Claim 6. Let Q_1 and Q_2 be normal cliques and suppose $x \in Q_1 \cap Q_2$. Then $N(x) \subseteq Q_1 \cup Q_2$.

To prove this claim, consider any point $y \in N(x)$, and suppose that $y \notin Q_1 \cup Q_2$. Then since Q_2 is normal, y must be adjacent to all points of $Q_1 - Q_2$ by Claim 4 and similarly, it must be adjacent to all points of $Q_2 - Q_1$. We will show that $N(Q_1) \subseteq N(y)$. This will clearly be a contradiction since $N(Q_1)$ contains three independent points.

So suppose $v \in N(Q_1)$. Let u be any neighbor of v in Q_1. We may assume that $u \in Q_1 - Q_2$, since if $u \in Q_1 \cap Q_2$, it then follows that v is adjacent to all points of $Q_1 - Q_2$ by the argument above. So v and y are non-distant neighbors of Q_1 and hence they are adjacent by Claim 4. Hence $v \in N(y)$ as claimed. This completes the proof of Claim 6.

Claim 7. No point of G is contained in more than two normal cliques. For suppose that some point x is contained in three normal cliques Q_1, Q_2 and Q_3. Then by Claim 6, $Q_1 \subseteq Q_2 \cup Q_3$ and similarly, $Q_2 \subseteq Q_1 \cup Q_3$ and $Q_3 \subseteq Q_1 \cup Q_2$. Hence $Q_1 \cup Q_2 \cup Q_3$ is a clique which is clearly a contradiction.

It is now easy to complete the proof of the theorem. Construct the graph H the points of which are the normal cliques of G by joining two normal cliques Q_1 and Q_2 by $|Q_1 \cap Q_2|$ lines. Then $L(H)$ is isomorphic to G. ■

Now we are prepared to prove the main theorem of this section.

12.4.6. THEOREM. *The independence number of a claw-free graph can be computed in polynomial time.*

PROOF. Let G be a claw-free graph. Select any point a of G and check to see if it is regular. This simply means to check whether $N(a)$ induces a bipartite subgraph in the complement of G, and this can be done trivially in polynomial time. If we find that a is irregular, then we reduce the problem to a smaller instance by Lemma 12.4.3. If we find that a is a regular point (i.e., that $N(a)$ can be partitioned into two complete subgraphs T_1 and T_2 of G), then we can extend each $T_i \cup a$ to a clique Q_i and check whether or not these two cliques are reducible. If either one of them is reducible, then again we reduce the problem to a smaller instance, this time using Lemma 12.4.4. If both cliques are irreducible, then we move on to another point. If we have inspected all points without being able to reduce our problem to a smaller graph, then by Theorem 12.4.5 we have a line graph of a graph H. (In fact, the proof of that theorem shows how to construct H.) So we can then apply any of the polynomial-time matching algorithms of Chapter 9. ■

REMARK 1. With some care, the algorithm above can be implemented in $O(p^4)$ time.

REMARK 2. The algorithm described in this section proceeds by reducing the size of the graph directly until a line graph is obtained. If, in addition to claws, we also exclude induced 3-paths, then a similar reduction can be defined which reduces the size of the graph all the way down to zero! (See Hammer, Mahadev and de Werra (1983).)

BOX 12A. Bounds on the Independence Number, or: Can Anything be Done with NP-complete Problems?

Having learned that the computation of $\alpha(G)$ is NP-hard, some might conclude that this problem is mathematically "intractable" and is deserving of no more study; no polynomial-time algorithm or good characterization can be expected. Others, who deny the relevance of polynomiality of algorithms, may say that all we should be concerned with in any case is the improvement of mathematically more or less trivial algorithms by employing various heuristics, programming and data handling tricks, etc. Common to these two extreme views is that they both deny the necessity of further attempts to build a deeper mathematical theory of the independence number of general graphs.

Let us now query a third person, a more traditional mathematician, perhaps. He might well answer that if we cannot determine a function exactly, we should seek good upper and lower bounds. Looking at this suggestion from the point of view of complexity theory, we are led to some profitable approaches to independence number and to other NP-hard problems in general.

It is essential to remark at once that, in the case of independence number, there is a substantial difference between having *upper* and *lower* bounds. A lower bound on the independence number means the existence of an independent set of points of a sufficiently large size. One way to find such a lower bound is to give an algorithm which produces a "large" (although not necessarily a *largest*) independent set in every graph. Such an algorithm is often called a *heuristic* for the vertex packing problem.

Of course, the crucial task, then, is to try to measure how far we are from the optimum. In spite of the obvious practical significance of this problem, we know little about the possibilities of heuristics for the general vertex packing problem. In particular, we do not know if there is any polynomial time heuristic for the vertex packing problem which gives us an independent set whose size tends to infinity as $\alpha(G)$ tends to infinity. To wax even more optimistic, we might seek a heuristic which produces an independent set of size, say, $\alpha(G)/2$ in every graph G. Of course, less constructive methods such as random choice, existence proved by enumeration, etc., may also lead to lower bounds on $\alpha(G)$. Although no really general approaches have been found so far, a very large number of results in extremal graph theory, Ramsey theory and other fields of graph theory are based implicitly on such heuristics. For a survey of some of these, see Griggs (1983).

Upper bounds on the independence number, on the other hand, mean the *non*-existence of independent subsets of a certain size. What can be the practical significance of such a negative result? In order to discuss some examples, let us point out that there are two kinds of upper bounds of interest. Sometimes, we may find an upper bound $f(G)$ which is computable in polynomial time. On the other hand, it is also of interest to find upper bounds $f(G)$ which are not necessarily polynomially computable, but for which the property $f(G) \leq k$ is in NP. We call such a bound an **NP-bound.**

If we have a polynomially computable upper bound for $\alpha(G)$, it can be used to design a branch-and-bound algorithm to find $\alpha(G)$; the sharper the upper bound, the better "pruned" is the search tree. (For an introduction to branch-and-bound algorithms for graphs, see Gondran and Minoux (1979, 1984).)

If we want to analyze the performance of a heuristic, that is, if we want to compare its result with the actual value of $\alpha(G)$, we need good upper bounds on $\alpha(G)$.

Given any NP-bound on $\alpha(G)$, we can consider the class of graphs for which this upper bound is attained. For these graphs, the value of $\alpha(G)$ is well-characterized, and hence there is hope that for such graphs it is easier to calculate $\alpha(G)$. Unfortunately, this does not always work! The minimum number of complete graphs covering graph G — in other words, $\chi(\overline{G})$ — is an NP-bound for $\alpha(G)$, but determination of the independence number is already NP-complete for those graphs with $\alpha(G) = \chi(\overline{G})$. It is also NP-complete to decide whether or not $\alpha(G) = \chi(\overline{G})$ holds for a given graph. And yet, a modification of this property, namely requiring it to hold for all induced subgraphs as well, leads to the important notion of *perfect* graph.

Let us mention some further upper bounds on $\alpha(G)$. We can improve the trivial bound $\chi(\overline{G})$ mentioned above by using the idea of "fractional covering" by cliques. Here we suppose that non-negative weights are assigned to the cliques of graph G in such a way that the sum of weights of cliques containing a given point is at least 1, and then try to minimize the total sum of weights. Let us denote this minimum by $\alpha^*(G)$. It is then easy to see that the value of $\alpha^*(G)$ is always an NP-bound for $\alpha(G)$. Unfortunately, however, to compute the number, $\alpha^*(G)$, is NP-hard!

Another way to look at this number is as follows. Let $a_i^{\mathrm{T}} \cdot x \leq b_i$, $i = 1, \ldots, m$, be any set of valid constraints for the vertex packing polytope $VP(G)$. Then an upper bound for $\alpha(G)$ can be obtained by solving the following linear program:

$$\begin{aligned} &\text{maximize} && \mathbf{1}^{\mathrm{T}} \cdot x \\ &\text{subject to} && x \geq \mathbf{0} \\ & && a_i^{\mathrm{T}} \cdot x \leq b_i \qquad (i = 1, \ldots, m) \end{aligned}$$

or equivalently by solving its dual:

$$\begin{aligned} &\text{minimize} && \sum b_i y_i \\ &\text{subject to} && y_i \geq 0 \qquad (i = 1, \ldots, m) \\ & && \sum y_i a_i \geq \mathbf{1}. \end{aligned}$$

Thus from the line constraints we obtain as an NP-bound on $\alpha(G)$ the value $p - \nu_2(G)/2$, which is even polynomially computable, but rather crude. From the clique constraints we obtain the number $\alpha^*(G)$, which is, unfortunately, NP-hard to compute. If we take the odd hole constraints, as discussed Section 12.2, then we obtain a better-behaved upper bound, that is, one which can be computed in polynomial time (Grötschel, Lovász and Schrijver (1984a)). Unfortunately, this upper bound also tends to be rather rough for many graphs, even for most perfect graphs.

Let us conclude with the mention of an upper bound on $\alpha(G)$ which is better than $\alpha^*(G)$ and, in addition, is also computable in polynomial time. This bound is based on quite different ideas, however, and we shall only sketch them here. Let G be a graph and consider all symmetric matrices whose rows and columns are indexed by the points of G. Fix the value of the entry a_{ij} at 1 if $i = j$ and also if i and j are non-adjacent, but let the value of the other entries be arbitrary. Choose the other entries so that the largest eigenvalue of A is as small as possible, and denote this minimum value of the largest eigenvalue by $\vartheta(G)$. Then this number, $\vartheta(G)$, is an upper bound on $\alpha(G)$, and it turns out that this bound is always sharper than $\alpha^*(G)$. (See Lovász (1979d).) Its main advantage is, however, that in spite of its very complicated definition, it can be computed in polynomial time (Grötschel, Lovász and Schrijver (1981)). Since, in particular, for any perfect graph, we have $\alpha(G) = \vartheta(G) = \alpha^*(G) = \chi(\overline{G})$, this result implies that the independence number of any perfect graph can be determined in polynomial time.

For further applications of $\vartheta(G)$ and more generally for more on upper bounds on the independence number of a graph, see Hansen (1979, 1980), Hammer, Hansen and Simeone (1980), Li and Li (1981), Lovász (1982) and Grötschel, Lovász and Schrijver (1984a).

References

The reader should note that, where possible, we have appended a review listing at the end of each reference. If available, we give the *Mathematical Reviews* (MR) code. If that is not available, we provide the code from *Zentralblatt für Mathematik und ihre Grenzgebiete* (Zbl.) and for older entries, we give the code from *Jahrbuch über die Fortschritte der Mathematik* (Jbuch.).

In the case of Russian words, we have adopted the transliteration used by *Mathematical Reviews* when available.

At the end of each entry, the page numbers of the present book upon which the reference is cited are placed in square brackets ([ ]).

H. L. ABBOTT

1966. Some remarks on a combinatorial theorem of Erdős and Rado, *Canad. Math. Bull.* **9**, 1966, 155–160. MR33#5497 [114].

H. L. ABBOTT, D. HANSON AND N. SAUER

1972. Intersection theorems for systems of sets, *J. Combin. Theory Ser. A* **12**, 1972, 381–389. MR45#6633 [114].

I. ADLER, N. MEGIDDO AND M. J. TODD

1984. New results on the average behavior of simplex algorithms, *Bull. Amer. Math. Soc.* **11**, 1984, 378–382 [262].

R. AHARONI

1984. König's duality theorem for infinite bipartite graphs, *J. London Math. Soc.*, 29, 1984, (to appear) [5].

A. V. AHO, J. E. HOPCROFT AND J. D. ULLMAN

1974. *The Design and Analysis of Computer Algorithms*, Addison-Wesley, Reading, Mass., 1974. MR54#1706 [10].

M. AIGNER

1975. *Kombinatorik I*, Springer-Verlag, Berlin, 1975. MR52#13396 [28].

1976. *Kombinatorik II*, Springer-Verlag, Berlin, 1976. MR57#123 [28].

1979. *Combinatorial Theory*, Grundlehren Math. Wiss. **234**, Springer-Verlag, Berlin, 1979. MR80h:05002 [28].

M. AIGNER AND T. A. DOWLING

1971. Matching theory for combinatorial geometries, *Trans. Amer. Math. Soc.* **158**, 1971, 231–245. MR44#3898 [27].

J. AIHARA

1976. A new definition of Dewar-type resonance energies, *J. Amer. Chem. Soc.* **98**, 1976, 2750–2758 [352].

I. ANDERSON

1971. Perfect matchings of a graph, *J. Combin. Theory Ser. B* **10**, 1971, 183–186. MR43#1853 [85].

1973. Sufficient conditions for matching, *Proc. Edinburgh Math. Soc.* **18**, 1973, 129–136. MR52#5479 [115].

B. ANDRÁSFAI

1967. On critical graphs, *Theory of Graphs (International Symposium, Rome, 1966)*, Ed.: P. Rosenstiehl, Gordon and Breach, New York, 1967, 9–19. MR36#5016 [448, 453].

R. P. ANSTEE

1983. A polynomial algorithm for *b*-matchings: an alternative approach, Univ. Waterloo Research Report CORR 83-22, 1983 [387].

K. I. APPEL AND W. HAKEN

1977. Every planar map is four colorable. I. Discharging, *Illinois J. Math.* **21**, 1977, 429–490. MR5827598a [xiii].

K. I. APPEL, W. HAKEN AND J. KOCH

1977. Every planar map is four colorable. II. Reducibility, *Illinois J. Math.* **21**, 1977, 491–567. MR58#27598b [xiii].

D. AVIS

1983. A survey of heuristics for the weighted matching problem, *Networks* **13**, 1983, 475–493 [358].

L. BABAI

1979. Monte-Carlo algorithms in graph isomorphism testing, Univ. Montreal, Dept. Math. and Stat. Report D. M. S. No. 79–10, 1979 [333].

F. BÄBLER

1938. Über die Zerlegung regulärer Streckenkomplexe ungerader Ordnung, *Comment. Math. Helv.* **10**, 1938, 275–287. Jbuch. 64.596 [xii].

1952. Bemerkungen zu einer Arbeit von Herrn R. Catoni, *Comment. Math. Helv.* **26**, 1952, 117–118. MR14, 68b168 [xii].

1954. Über den Zerlegungssatz von Petersen, *Comment. Math. Helv.* **28**, 1954, 155–161. MR16, 57d [xii].

E. BALAS

1981. Integer and fractional matchings, *Studies on Graphs and Discrete Programming*, Ed.: P. Hansen, Ann. Discrete Math., 11, North-Holland, Amsterdam, 1981, 1–13. MR84h:90084 [220].

M. L. BALINSKI

1965. Integer programming: methods, uses and computation, *Management Sci.* **12**, 1965, 253–313. MR33#1149 [291].

1969. Labelling to obtain a maximum matching (with discussion), Chapt. 33 in: *Combinatorial Mathematics and its Applications*, Eds.: R. C. Bose and T. A. Dowling, The University of North Carolina Monograph Series in Probability and Statistics, No. 4, University of North Carolina Press, Chapel Hill, 1969, 585–602. MR41#8270 [369].

1970. On perfect matchings, *SIAM Rev.* **12**, 1970, 570–572. MR44#2637 [85].

M. L. BALINSKI AND K. SPIELBERG

1969. Methods for integer programming: algebraic, combinatorial and enumerative, in: *Progress in Operations Research* **III**, Ed.: J. Aronofsky, Wiley, New York, 1969, 195–292. MR57#18812 [292].

M. O. BALL AND U. DERIGS

1983. An analysis of alternate strategies for implementing matching algorithms, *Networks* **13**, 1983, 517–549 [376].

T. BANG

1979. On matrix-functions giving a good approximation to the van der Waerden permanent conjecture, Københavns Univ.Math. Inst. preprint no. 30, 1979 [315].

F. BARAHONA

1982. On the computational complexity of Ising spin glass models, *J. Phys. A: Math. Gen.* **15**, 1982, 3241–3253. MR84c:82022 [353].

F. BARAHONA, R. MAYNARD, R. RAMMAL AND J. P. UHRY

1982. Morphology of ground states of a 2-dimensional frustration model, *J. Physique A.: Math. Gen.* **15**, 1982, 673–700. MR83c:82045 [354].

G. W. BARTNIK

1978. Algorithmes de couplages dans les graphes, Thèse Doctorat 3ᵉ cycle, Université Paris VI, 7 févier 1978 [369].

L. W. BEINEKE

1968. On derived graphs and digraphs, *Beiträge zur Graphentheorie*, Eds.: H. Sachs, H. J. Voss and H. Walther, Teubner, Leipzig, 1968, 17–23 [472].

1970. Characterizations of derived graphs, *J. Combin. Theory* **9**, 1970, 129–135. MR41#6707 [472].

L. W. BEINEKE, F. HARARY AND M. D. PLUMMER

1967. On the critical lines of a graph, *Pacific J. Math.* **22**, 1967, 205–212. MR35#2772 [448].

L. W. BEINEKE AND M. D. PLUMMER

1967. On the 1-factors of a non-separable graph, *J. Combin. Theory* **2**, 1967, 285–289. MR35#1499 [345].

H.-B. BELCK

1950. Reguläre Factoren von Graphen, *J. Reine Angew. Math.* **188**, 1950, 228–252. MR12, 730d [xii].

C. BERGE

1957. Two theorems in graph theory, *Proc. Nat. Acad. Sci. U. S. A.* **43**, 1957, 842–844. MR20#1323 [13].

1958a. Sur le couplage maximum d'un graphe, *C. R. Acad. Sci. Paris Sér. I Math.* **247**, 1958, 258–259. MR20#7278 [xviii, 90].

1958b. *Théorie des Graphes et ses Applications*, Dunod, Paris, 1958. MR21 #1608 [xvii].

1961. Färbung von Graphen, deren sämtliche bzw. deren ungerade Kreise starr sind, *Wiss. Zeitung, Martin Luther Univ. Halle-Wittenberg*, 1961, 114 [462].

1962. *The Theory of Graphs and its Applications*, John Wiley, New York, 1962. MR24#A2381 [xviii].

1969. The rank of a family of sets and some applications to graph theory, *Recent Progress in Combinatorics*, Ed.: W. T. Tutte, Academic Press, New York, 1969, 49–57. MR41#5231 [468].

1970. Sur certain hypergraphes generalisant les graphes bipartites, in: *Combinatorial Theory and its Applications* I, Eds.: P. Erdős, A. Rényi and V. Sós, Colloq. Math. Soc. János Bolyai, 4, North-Holland, Amsterdam, 1970, 119–133. MR45#6653 [468].

1972. Alternating chain methods: a survey, *Graph Theory and Computing*, Ed.: R. Read, Academic Press, New York, 1972, 1–13. MR50#9685 [448].

1973. *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973. MR50 #9640 [119, 404, 466].

1978a. Regularizable graphs, *Advances in Graph Theory*, Ed.: B. Bollobás, Ann. Discrete Math., 3, North-Holland, Amsterdam, 1978, 11–19. MR80a:05155 [218].

1978b. Regularizable graphs I, *Discrete Math.* **23**, 1978, 85–89. MR80h:05041 [218].

1978c. Regularizable graphs II, *Discrete Math.* **23**, 1978, 91–95. MR80h:05041 [218].

1978d. *Fractional Graph Theory*, Indian Statistical Institute Lecture Notes, No. 1, MacMillan Co. of India, New Delhi, 1978. MR81g:05085 [218].

1979. Regularizable graphs, *Proc. of the Symposium on Graph Theory (Calcutta, 1976)*, Indian Statistical Institute Lecture Notes No. 4, MacMillan of India, New Delhi, 1979. MR80m:05088 [218].

1981. Some common properties for regularizable graphs, edge-critical graphs and B-graphs, *Graph Theory and Algorithms (Proc. Sympos. Res. Inst. Electr. Comm., Tohoku Univ., Sendai, 1980)*, Eds.: N. Saito and T. Nishzeki, Lecture Notes in Comput. Sci. **108**, Springer-Verlag, Berlin, 1981, 108–123. MR83k:05060 [218].

C. BERGE AND M. LAS VERGNAS

1970. Sur un théorème du type König pour hypergraphes, *International Conference on Combinatorial Mathematics*, Eds.: A. Gewirtz and L. Quintas, Ann. New York Acad. Sci. **175**, 1970, 32–40. MR42#1690 [469].

J.-C. BERMOND AND J. C. MEYER

1973. Graphe représentatif des arêtes d'un multigraphe, *J. Math. Pures Appl.* **52**, 1973, 299–308. MR50#9695 [472].

I. BIECHE, R. MAYNARD, R. RAMMAL AND J. P. UHRY

1980. On the ground states of the frustration model of a spin glass by a matching method of graph theory, *J. Phys. A: Math. Gen.* **13**, 1980, 2553–2576. MR81g:82037 [353].

N. BIGGS

1974. *Algebraic Graph Theory*, Cambridge Univ. Press, Cambridge, 1974. MR50#151 [207].

N. L. BIGGS, E. K. LLOYD AND R. J. WILSON

1976. *Graph Theory 1736–1936*, Clarendon Press, Oxford, 1976. MR56#2771 [xii, xvi].

P. BILLINGSLEY

1968. *Convergence of Probability Measures*, John Wiley &Sons, New York, 1968. MR38#1718 [77].

G. BIRKHOFF

1946. Tres observaciones sobre el algebra lineal, *Rev. Univ. Nac. Tucumán, Series A.* **5**, 1946, 147–151. MR8, 561 [xv, 36].

1967. *Lattice Theory*, Amer. Math. Soc. Colloq. Publ. **25** (third ed.), Providence, 1967. MR37#2638 [28].

E. BISHOP

1967. *Foundations of Constructive Analysis*, McGraw-Hill, New York, 1967. MR36#4930 [61].

R. G. BLAND

1977. A combinatorial abstraction of linear programming, *J. Combin. Theory Ser. B* **23**, 1977, 33–57. MR57#5793 [262].

R. G. BLAND, D. GOLDFARB AND M. J. TODD

1981. The ellipsoid method: A survey, *Oper. Res.* **29**, 1981, 1039–1081. MR83e:90081 [265].

R. G. BLAND AND M. LAS VERGNAS

1978. Orientability of matroids, *J. Combin. Theory Ser. B* **24**, 1978, 94–123. MR58#5294 [262].

B. BOLLOBÁS

1977. Extremal problems in graph theory, *J. Graph Theory* **1**, 1977, 117–123. MR56#8426 [114].

1978a. The number of 1-factors in $2k$-connected graphs, *J. Combin. Theory Ser. B* **25**, 1978, 363–366. MR80m:05060 [346].

1978b. *Extremal Graph Theory*, Academic Press, London, 1978. MR80a:05120 [xxxiii, 346].

1979. *Graph Theory. An Introductory Course*, Graduate Texts in Mathematics, 63, Springer-Verlag, Berlin, 1979. Zbl. 411.05032 [422].

B. BOLLOBÁS AND S. E. ELDRIDGE

1976. Maximal matchings in graphs with given minimal and maximal degrees, *Math. Proc. Cambridge Philos. Soc.* **79**, 1976, 221–234. MR52 #10491 [115].

B. BOLLOBÁS AND B. D. MCKAY

1982. The number of matchings in random regular graphs and bipartite graphs, preprint, 1982 [313].

J. A. BONDY AND U. S. R. MURTY

1976. *Graph Theory with Applications*, American Elsevier, New York, 1976 [xxxiii].

K. S. BOOTH AND G. S. LUEKER

1976. Testing for the consecutive ones property, interval graphs and graph planarity testing using PQ-tree algorithms, *J. Comput. System Sci.* **13**, 1976, 335–379. MR55#6932 [55].

K-H. BORGWARDT

1982a. Some distribution-independent results about the asymptotic order of the average number of pivot steps of the simplex method, *Math. Oper. Res.* **7**, 1982, 441–462. MR84i:90079 [262].

1982b. The average number of pivot steps required by the simplex-method is polynomial, *Z. Oper. Res.* **26**, 1982, 157–177. MR84d:90064 [262].

O. BORŮVKA

1926a. O jistém problému minimálním, *Práce Mor. Přírodověd. Spol. v Brně (Acta Societ. Scient. Natur. Moravicae)* **3**, 1926, 37–58. (Czech). Jbuch. 57.1543 [**x**, 248].

1926b. Příspěvek k řešeníotázky ekonomické stavby elektrovodních sítí, *Elektrotechnický obzor* **15**, 1926, 153–154. (Czech) [**x**, 248].

M. BOULALA AND J. P. UHRY

1979. Polytope des independants d'un graphe series-parallele, *Discrete Math.* **27**, 1979, 225–243. MR81f:05134 [464].

H. R. BRAHANA

1917–1918. A proof of Petersen's theorem, *Ann. of Math.* **19**, 1917–1918, 59–63. Jbuch. 46.835 [xii, 110].

L. M. BRÈGMAN

1973. Certain properties of nonnegative matrices and their permanents, *Dokl. Akad. Nauk SSSR* **211**, 1973, 27–30. (Russian). (English translation: *Soviet Math. Dokl.* **14**, 1973, 945–949.) MR48#6130 [311].

R. A. BRUALDI

1975. Transversal theory and graphs, *Studies in Graph Theory, Part I* **11**, MAA Studies in Mathematics, Mathematical Association of America, Washington, D. C., 1975, 23–88. MR53#10589 [30].

1982. Notes on the Birkhoff algorithm for doubly stochastic matrices, *Canad. Math. Bull.* **25**, 1982, 191–199. MR83k:15019 [37].

R. A. BRUALDI AND P. M. GIBSON

1977. Convex polyhedra of doubly stochastic matrices I. Applications of the permanent function, *J. Combin. Theory Ser. A* **22**, 1977, 194–230. MR55#10486 [127].

R. A. BRUALDI, F. HARARY AND Z. MILLER

1980. Bigraphs versus digraphs via matrices, *J. Graph Theory* **4**, 1980, 51–73. MR81b:05077 [127].

S. G. BRUSH

1964. History of the Lenz-Ising model, Lawrence Radiation Laboratory, Livermore, Calif., Report UCRL-7940, 1964 [355].

T. BRYLAWSKI

1982. The Tutte polynomial, *Matroid Theory and its Applications*, Centro Internazionale Matematico Estivo, Liguori editore, Naples, 1982, 125–276 [341].

R. BURKARD AND U. DERIGS

1980. *Assignment and Matching Problems: Solution Methods with FORTRAN Programs*, (With the assistance of T. Bönniger and G. Katzakidis), Lecture Notes in Econom. and Math. Systems **184**, Springer-Verlag, Berlin, 1980. MR82i:90073 [369].

C. CARATHÉODORY

1907. Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen, *Math. Ann.* **64**, 1907, 95–115. Jbuch. 38.448 [257].

P. G. CARTER

1949. An empirical equation for the resonance energy of polycyclic aromatic hydrocarbons, *Trans. Faraday Soc.* **45**, 1949, 597–602 [350].

B. V. ČERKASSKIĬ

1977. Algorithm of construction of maximal flow in networks with complexity $O(n^2p^{1/2})$ operations, *Akad. Nauk SSSR, CEMI, Mathematical Methods for the Solution of Economical Problems* **7**, 1977, 117–126. (Russian). MR58#20342 [57].

G. CHARTRAND, D. L. GOLDSMITH AND S. SCHUSTER

1979. A sufficient condition for graphs with 1-factors, *Colloq. Math.* **41**, 1979, 339–344. MR82f:05062 [113].

G. CHARTRAND AND L. NEBESKÝ

1979. A note on 1-factors in graphs, *Period. Math. Hungar.* **10**, 1979, 41–46. MR58#21834 [113].

N. CHRISTOFIDES

1975. *Graph Theory: An Algorithmic Approach*, Academic Press, London, 1975. MR55#2623 [243].

V. CHVÁTAL

1970. On finite delta-systems of Erdős and Rado, *Acta Math. Acad. Sci. Hungar.* **21**, 1970, 341–355. Zbl. 207.25 [114].

1973a. Edmonds polytopes and a hierarchy of combinatorial problems, *Discrete Math.* **4**, 1973, 305–337. MR47#1635 [283].

1973b. New directions in Hamiltonian graph theory, *New Directions in the Theory of Graphs*, Ed.: F. Harary, Academic Press, New York, 1973, 65–95. MR50#9689 [117].

1973c. Tough graphs and Hamiltonian circuits, *Discrete Math.* **5**, 1973, 215–228. MR47#4849 [117].

1975. On certain polytopes associated with graphs, *J. Combin. Theory Ser. B* **18**, 1975, 138–154. MR51#7949 [283, 461, 464].

1983. *Linear Programming*, W. H. Freeman, New York, 1983 [257, 261].

V. CHVÁTAL AND D. HANSON

1976. Degrees and matchings, *J. Combin. Theory Ser. B* **20**, 1976, 128–138. MR55#169 [114].

E. G. COFFMAN, JR. AND R. L. GRAHAM

1972. Optimal scheduling for two processor systems, *Acta Inform.* **1**, 1972, 200–213. MR48#13231 [viii].

S A. COOK

1971. The complexity of theorem proving procedures, *Proceedings of the Third Annual ACM Symposium on Theory of Computing (Shaker Heights, 1971)*, ACM, New York, 1971, 151–158. Zbl. 253.68020 [8, 226].

G. CORNUÉJOLS, D. HARTVIGSEN AND W. R. PULLEYBLANK

1982. Packing subgraphs in a graph, *O. R. Letters* **1**, 1982, 139–143. MR84b: 68039 [230, 231, 390].

G. CORNUÉJOLS AND W. R. PULLEYBLANK

1980a. Perfect triangle-free 2-matchings, *Combinatorial Optimization II (Proc. Conf. Univ. East Anglia, Norwich, 1979)*, Math. Programming Stud. No. 13, North-Holland, Amsterdam, 1980, 1–7. MR81m:05110 [230].

1980b. A matching problem with side conditions, *Discrete Math.* **29**, 1980, 135–159. MR83b:05100 [231].

1982. The travelling salesman polytope and $\{0, 2\}$-matchings, *Bonn Workshop on Combinatorial Optimization*, Eds.: A. Bachem, M. Grötschel and B. Korte, Ann. Discrete Math., 16, North-Holland, Amsterdam, 1982, 27–55. MR84c:90037 [230].

1983. Critical graphs, matchings and tours or a hierarchy of relaxations for the travelling salesman problem, *Combinatorica* **3**, 1983, 35–52. Zbl. 478.05074 [228, 229].

H. CRAPO

1979. Structural rigidity, *Structural Topology* **1**, 1979, 26–45. MR82g:51028 [420].

H. CRAPO AND G.-C. ROTA

1968. *Combinatorial Geometries*, M. I. T. Press, Cambridge, Mass., 1968. MR45#74 [28].

A. B. CRUSE

1977. A note on 1-factors of certain regular multigraphs, *Discrete Math.* **18**, 1977, 213–216. MR57#3004 [113].

W. H. CUNNINGHAM AND A. B. MARSH III

1978. A primal algorithm for optimum matching, *Polyhedral Combinatorics (dedicated to the memory of D. R. Fulkerson)*, Eds.: M. L. Balinski and A. J. Hoffman, Math. Programming Stud. No. 8, North-Holland, Amsterdam, 1978, 50–72. MR81a:68067 [280, 376].

D. M. CVETKOVIĆ, M. DOOB AND H. SACHS

1979. *Spectra of Graphs, Theory and Application*, VEB Deutscher Verlag der Wissenschaften, Berlin (D. D. R.), 1979, 245–251. MR81i:05054 [335].

G. B. DANTZIG

1951. Maximization of a linear function of variables subject to linear inequalities, Chapter 21 in: *Activity Analysis of Production and Allocation,* (Cowles Commission Monograph No. 13), Ed.: T. C. Koopmans, John Wiley and Sons, New York, 1951, 339–347. MR15-47 [xix, 261].

1963. *Linear Programming and Extensions*, Princeton University Press, Princeton, N. J., 1963. MR34#1073 [xix, 261].

1983. Reminiscences about the origins of linear programming, in: *Mathematical Programming, The State of the Art: Bonn, 1982,* Eds.: A. Bachem, M. Grötschel and B. Korte, Springer-Verlag, Berlin, 1983, 78–86. MR84h:01063 [xix].

R. W. DEMING

1979. Independence numbers of graphs — an extension of the König-Egerváry theorem, *Discrete Math.* **27**, 1979, 23–33. MR80j:05074 [223].

J. DÉNES AND A. KEEDWELL

1974. *Latin Squares and their Applications,* English Universities Press, London, 1974. MR50#4338 [348].

M. D. DEVINE

1973. A model for minimizing the cost of drilling dual completion oil wells, *Management Sci.* **20**, 1973, 532–535 [viii].

E. W. DIJKSTRA

1959. A note on two problems in connexion with graphs, *Numer. Math.* **1**, 1959, 269–271. MR21#6334 [243].

R. P. DILWORTH

1950. A decomposition theorem for partially ordered sets, *Ann. of Math.* **51**, 1950, 161–166. MR11, 309f [xvi, 32, 33].

E. A. DINIC

1970. Algorithm for solution of a problem of maximum flow in a network with power estimation, *Dokl. Akad. Nauk SSSR* **194**, 1970, 754–757. (Russian). MR44#5178. (English translation: *Soviet Math. Dokl.* **11**, 1970, 1277–1280.) [48, 56, 57].

R. M. DUDLEY

1968. Distances of probability measures and random variables, *Ann. Math. Statist.* **39**, 1968, 1563–1572. MR37#5900 [76].

I. S. DUFF

1977. A survey of sparse matrix research, *Proc. IEEE* **65**, 1977, 500–535 [141].

A. L. DULMAGE AND N. S. MENDELSOHN

1958. Coverings of bipartite graphs, *Canad. J. Math.* **10**, 1958, 517–534. MR20#3549 [99, 137].

1959. A structure theory of bipartite graphs of finite exterior dimension, *Trans. Roy. Soc. Canada Ser. III* **53**, 1959, 1–13. Zbl. 91.375 [99, 137].

1967. Graphs and matrices, *Graph Theory and Theoretical Physics*, Ed.: F. Harary, Academic Press, New York, 1967, 167–277. MR40#5468 [99].

J. EDMONDS

1965a. Paths, trees, and flowers, *Canad. J. Math.* **17**, 1965, 449–467. MR31 #2165 [xviii, xix, 13, 92, 93, 357, 369].

1965b. Maximum matching and a polyhedron with (0, 1) vertices, *J. Res. Nat. Bur. Standards Sect. B* **69B**, 1965, 125–130. MR32#1012 [xx, 255, 274, 357, 370, 376].

1965c. Minimum partition of a matroid into independent subsets, *J. Res. Nat. Bur. Standards Sect. B* **69B**, 1965, 67–72. MR32#7441 [8].

1967a. Systems of distinct representatives and linear algebra, *J. Res. Nat. Bur. Standards Sect. B* **71B**, 1967, 241–247. MR37#5114 [316].

1967b. Optimum branchings, *J. Res. Nat. Bur. Standards Sect. B* **71B**, 1967, 233–240. MR37#2632 [250].

1970. Submodular functions, matroids, and certain polyhedra, *Combinatorial Structures and their Applications*, Eds.: R. Guy, H. Hanani, N. Sauer and J. Schönheim, Gordon and Breach, New York, 1970, 69–87. MR42#5828 [24, 26, 253, 418].

1971. Matroids and the greedy algorithm, *Math. Programming* **1**, 1971, 127–136. MR45#6414 [28].

1973. Edge-disjoint branchings, *Combinatorial Algorithms*, Ed. R. Rustin, Academic Press, New York, 1973, 91–96. MR50#4377 [253, 254].

J. EDMONDS AND D. R. FULKERSON

1965. Transversals and matroid partitions, *J. Res. Nat. Bur. Standards Sect. B* **69B**, 1965, 147–153. MR32#5531 [88, 93].

J. EDMONDS AND R. GILES

1977. A min-max relation for submodular functions on graphs, *Studies in Integer Programming*, Eds.: P. L. Hammer, E. L. Johnson and B. H. Korte, Ann. Discrete Math., 1, North-Holland, Amsterdam, 1977, 185–204. MR57#165 [24, 254, 272].

J. EDMONDS AND E. L. JOHNSON

1973. Matching, Euler tours and the Chinese postman, *Math. Programming* **5**, 1973, 88–124. MR48#168 [233, 236].

J. EDMONDS AND R. M. KARP

1970. Theoretical improvements in algorithmic efficiency for network flow problems, *Combinatorial Structures and their Applications*, Eds.: R. Guy, H. Hanani, N. Sauer and J. Schönheim, Gordon and Breach, New York, 1970, 93–96. MR42#1583 [48, 53, 57].

1972. Theoretical improvements in algorithmic efficiency for network flow problems, *J. Assoc. Comput. Mach.* **19**, 1972, 248–264. Zbl. 318.90024 [48, 53, 57].

J. EDMONDS, L. LOVÁSZ AND W. R. PULLEYBLANK

1982. Brick decompositions and the matching rank of graphs, *Combinatorica* **2**, 1982, 247–274. Zbl. 521.05035 [296, 303, 304].

E. EGERVÁRY

1931. On combinatorial properties of matrices, *Mat. Lapok*, 38, 1931. 16–28. (Hungarian with German summary). Jbuch. 57.1340 [xvi, 15, 271].

G. P. EGORYČEV

1980. Solution of the van der Waerden problem for permanents, IFSO-13M, *Akad. Nauk SSSR Sibirsk. Otdel. Inst. Fiz.*, Krasnoyarsk, preprint, 1980. (Russian). MR82e:15006. (English translation: *Soviet Math. Dokl.*, American Mathematical Society, Providence **23**, 1982, 619–622. MR82e:15006, MR83b:15002a, MR83b:15002b and MR83b:15003.) [310].

1981. The solution of van der Waerden's problem for permanents, *Adv. in Math.* **42**, 1981, 299–305. MR83b:15002b [310].

A. EHRENFEUCHT, V. FABER AND H. A. KIERSTEAD

1982. A new method of proving theorems on chromatic index, Los Alamos National Laboratory preprint LA-UR-82-661, 1982 [286].

P. ELIAS, A. FEINSTEIN AND C. E. SHANNON

1956. Note on maximum flow through a network, *IRE Trans. Inform. Theory,* IT-**2**, 1956, 117–119 [xvi, 45].

P. ERDŐS

1967. On even subgraphs of graphs, *Mat. Lapok* **18**, 1967, 283–288. (Hungarian). MR39#95 [240].

P. ERDŐS AND T. GALLAI

1960. Graphs with prescribed degrees, *Mat. Lapok* **11**, 1960, 264–274. (Hungarian) [384, 405].

1961. On the minimal number of vertices representing the edges of a graph, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **6**, 1961, 181–203. MR26#1878 [101, 450].

P. ERDŐS, A. HAJNAL AND J. MOON

1964. A problem in graph theory, *Amer. Math. Monthly* **71**, 1964, 1107–1110. MR30#577 [452].

P. ERDŐS AND R. RADO

1960. Intersection theorems for systems of sets, *J. London Math. Soc.* **35**, 1960, 85–90. MR22#2554 [114].

P. ERDŐS AND J. SPENCER

1974. *Probabilistic Methods in Combinatorics*, Akadémiai Kiadó, Budapest, 1974. MR52#2895 [332].

A. ERRERA

1922. Une demonstration du théorème de Petersen, *Mathesis* **36**, 1922, 56–61. Jbuch. 48.664 [xii].

L. EULER

1736. Solutio problematis ad geometriam situs pertinentis, *Commentarii Academiae Scientarum Imperialis Petropolitanae* **8**, 1736, 128–140 [xi].

S. EVEN

1976. The max-flow algorithm of Dinic and Karzanov: an exposition, M. I. T., LCS. TM-80, preprint, Dec., 1976 [56, 57].

1979. *Graph Algorithms*, Computer Science Press, Potomac, Maryland, 1979. MR82e:68066 [43, 55, 56, 57].

S. EVEN AND O. KARIV

1975. An $O(n^{5/2})$ algorithm for maximum matching in general graphs, *16th Annual Symposium on Foundations of Computer Science (Berkeley, 1975)*, IEEE Computer Society Press, New York, 1975, 100–112. MR55#1800 [369].

S. EVEN AND R. E. TARJAN

1975. Network flow and testing graph connectivity, *SIAM J. Comput.* **4**, 1975, 507–518. MR55#9898 [55].

1976. Computing an *st* numbering, *Theoret. Comput. Sci.* **2**, 1976, 339–344. MR54#2508 [55].

V. FABER, A. EHRENFEUCHT AND H. A. KIERSTEAD

1981. A new method of proving theorems on chromatic index, Los Alamos National Labs preprint LA-UR-82-661, 1981 [286].

R. FAGIN

1974. Generalized first order spectra and polynomial-time recognizable sets, *Complexity of Computation*, Ed. R. Karp, SIAM-AMS Proc. **7**, 1974, 43–73. MR51#7840 [10].

D. I. FALIKMAN

1981. A proof of the van der Waerden conjecture on the permanent of a doubly stochastic matrix, *Mat. Zametki* **29**, 1981, 931–938. (Russian). (English translation: *Mathematical Notes of the Academy of Science of the USSR*, Consultants Bureau, New York **29**, 1981, 475–479. MR82k: 15007.) [310].

W. FELLER

1950. *An Introduction to Probability Theory and its Applications*, Wiley, New York, 1950. MR12-424 [344].

S. FIORINI AND R. WILSON

1977. *Edge Colourings of Graphs*, Res. Notes in Math. **16**, Pitman, London, 1977. MR58#27599 [287].

J. FOLKMAN AND D. R. FULKERSON

1969. Edge colorings in bipartite graphs, *Combinatorial Mathematics and its Applications*, Eds.: R. C. Bose and T. A. Dowling, The University of North Carolina Monograph Series in Probability and Statistics, No. 4, University of North Carolina Press, Chapel Hill, 1969, 561–577. MR41#6722 [74].

J. FOLKMAN AND J. LAWRENCE

1978. Oriented matroids, *J. Combin. Theory Ser. B* **25**, 1978, 199–236. MR81g:05045 [262].

J. FONLUPT AND A. ZEMIRLINE

1983. On the number of common bases of two matroids, *Discrete Math.* **45**, 1983, 217–228 [305].

L. R. FORD, JR. AND D. R. FULKERSON

1956. Maximal flow through a network, *Canad. J. Math.* **8**, 1956, 399–404. MR18-56 [xvi, xix, 45, 57].

1957. A simple algorithm for finding maximal network flows and an application to the Hitchcock problem, *Canad. J. Math.* **9**, 1957, 210–218. MR19-1244 [xix].

1958. Network flow and systems of representatives, *Canad. J. Math.* **10**, 1958, 78–84. MR20#4502 [31, 255].

1962. *Flows in Networks*, Princeton University Press, Princeton, N. J., 1962. MR28#2917 [xix, 3, 31, 33, 42, 46, 62].

A. FRANK

1980. On the orientation of graphs, *J. Combin. Theory Ser. B* **28**, 1980, 251–261. MR81i:05075 [74].

1982. An algorithm for submodular functions on graphs, *Bonn Workshop on Combinatorial Optimization*, Eds.: A. Bachem, M. Grötschel and B. Korte, Ann. Discrete Math., 16, North-Holland, Amsterdam, 1982, 97–120 [24].

A. FRANK AND A. GYÁRFÁS

1978. How to orient the edges of a graph?, *Combinatorics*, I, Eds.: A. Hajnal and V. T. Sós, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam, 1978, 353–364. MR80c:05078 [74].

A. FRANK, A. SEBŐ AND É. TARDOS

1983. Covering directed odd cuts, preprint, May, 1983 [245].

A. FRANK AND É. TARDOS

1984. Generalized polymatroids and submodular flows, (in preparation) [305].

S. FRIEDLAND

1979. A lower bound for the permanent of a doubly stochastic matrix, *Ann. of Math.* **110**, 1979, 167–176. MR81i:15006 [315].

O. FRINK JR.

1925–1926. A proof of Petersen's theorem, *Ann. of Math.* **27**, 1925–1926, 491–493. Jbuch. 52.577 [xii, 110].

1926. A proof of Petersen's theorem, *Bull. Amer. Math. Soc.* **32**, 1926, 201. Jbuch. 52.605 [110].

G. FROBENIUS

1912. Über Matrizen aus nicht negativen Elementen, *Sitzungsber. König. Preuss. Akad. Wiss.* **26**, 1912, 456–477. Jbuch. 43.204 [xiii].

1917. Über zerlegbare Determinanten, *Sitzungsber. König. Preuss. Akad. Wiss.* **XVIII**, 1917, 274–277. Jbuch. 46.144 [xv, 6].

M. FUJII, T. KASAMI AND N. NINOMIYA

1969. Optimal sequencing of two equivalent processors, *SIAM J. Appl. Math.* **17**, 1969, 784–789. MR40#6945. (Erratum: *SIAM J. Appl. Math.* **20**, 1971, 141.) [viii].

K. FUKUDA

1982. Oriented matroid programming, Ph. D. thesis, Dept. of Combinatorics and Optimization, Univ. Waterloo, 1982 [262].

D. R. FULKERSON

1956. Note on Dilworth's decomposition theorem for partially ordered sets, *Proc. Amer. Math. Soc.* **7**, 1956, 701–702. MR17, 1176g [33].

1971. Blocking and anti-blocking pairs of polyhedra, *Math. Programming* **1**, 1971, 168–194. MR45#3222 [461].

1972. Anti-blocking polyhedra, *J. Combin. Theory Ser. B* **12**, 1972, 50–71. MR44#2629 [461].

1974. Packing rooted directed cuts in a weighted directed graph, *Math. Programming* **6**, 1974, 1–13. MR52#7953 [250].

D. R. FULKERSON, A. J. HOFFMAN AND M. H. MCANDREW

1965. Some properties of graphs with multiple edges, *Canad. J. Math.* **17**, 1965, 166–177. MR31#2166 [119].

H. N. GABOW

1973. Implementation of algorithms for maximum matching on non-bipartite graphs, Ph.D. Thesis, Stanford University Dept. Comput. Sci., 1973 [369, 375].

1976. An efficient implementation of Edmonds' algorithm for maximum matching on graphs, *J. Assoc. Comput. Mach.* **23**, 1976, 221–234. MR53#9715 [369].

1979. Algorithmic proofs of two relations between connectivity and the 1-factors of a graph, *Discrete Math.* **26**, 1979, 33–40. MR80i:05066 [347].

P. GÁCS AND L. LOVÁSZ

1981. Khachian's algorithm for linear programming, *Mathematical Programming at Oberwolfach*, Eds.: H. König, B. Korte and K. Ritter, Math. Programming Stud. No. 14, North-Holland, Amsterdam, 1981, 61–68. MR83e:90088a [265].

D. GALE

1968. Optimal assignments in an ordered set: An application of matroid theory, *J. Combin. Theory* **4**, 1968, 176–180. MR37#2624 [28].

D. GALE, H. W. KUHN AND A. W. TUCKER

1950. On symmetric games, in: *Contributions to the Theory of Games*, Eds.: H. W. Kuhn and A. W. Tucker, Ann. Math. Studies **24**, Princeton Univ. Press, Princeton, N. J., 1950, 81–87. MR12-513 [260].

Z. GALIL

1978. A new algorithm for the maximal flow problem, *19th Annual Symposium on Foundations of Computer Science (Ann Arbor, 1978)*, IEEE Computer Society Press, New York, 1978, 231–245. MR80e:68103 [57].

1980. An $O(V^{5/3}E^{2/3})$ algorithm for the maximal flow problem, *Acta Inform.* **14**, 1980, 221–242. MR81i:90073 [57].

1983. Efficient algorithms for finding maximal matching in graphs, preprint, 1983 [15, 16, 369].

Z. GALIL, C. M. HOFFMAN, E. M. LUKS, C. P. SCHNORR AND A. WEBER

1982. An $O(n^3 \log n)$ deterministic and $O(n^3)$ probabilistic isomorphism test for trivalent graphs, *23rd Annual Symposium on Foundations of Computer*

Science (Chicago, 1982), IEEE Computer Society Press, New York, 1982, 118–125 [333].

Z. GALIL, S. MICALI AND H. GABOW

1982. Priority queues with variable priority and an $O(EV \log V)$ algorithm for finding a maximal weighted matching in general graphs, *23rd Annual Symposium on Foundations of Computer Science (Chicago, 1982)*, IEEE Computer Society Press, New York, 1982, 255–261 [376].

Z. GALIL AND A. NAAMAD

1979. Network flow and generalized path compression, *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing (Atlanta, 1979)*, ACM, New York, 1979, 13–26. MR81b:90068 [57].

1980. An $O(EV \log^2 V)$ algorithm for the maximal flow problem, *J. Comput. System Sci.* **21**, 1980, 203–217. MR81m:90046 [57].

T. GALLAI

1950. On factorization of graphs, *Acta Math. Acad. Sci. Hungar.* **1**, 1950, 134–153. MR12, 626b [xii].

1957/58. Maximum-minimum theorems for networks I, II, *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.*, 7, 1957, 305–338; **8**, 1958, 1–40. MR23#A1552a [254].

1958. Maximum-minimum Sätze über Graphen, *Acta Math. Acad. Sci. Hungar.* **9**, 1958, 395–434. MR23#A1552b [xx, 6].

1959. Über extreme Punkt- und Kantenmengen, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **2**, 1959, 133–138. MR24#A1222 [1, 2].

1961. Maximum-minimum Sätze und verallgemeinerte Faktoren von Graphen, *Acta Math. Acad. Sci. Hungar.* **12**, 1961, 131–173. MR24#A55 [254, 425].

1963a. Neuer Beweis eines Tutte'schen Satzes, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8**, 1963, 135–139. MR29#4050 [85, 89].

1963b. Kritische Graphen II, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **8**, 1963, 373–395. MR32#5541 [93].

1964a. Maximale Systeme unabhängiger Kanten, *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **9**, 1964, 401–413. MR32#7445 [xviii, 93].

1964b. König Dénes (1884–1944), *Mat. Lapok* **15**, 1964, 277–293. (Hungarian) [xvii].

1978. The life and scientific work of Dénes König (1884–1944), *Linear Algebra Appl.* **21**, 1978, 189–205. MR80b:01021 [xvii].

M. R. GAREY AND D. S. JOHNSON

1979. *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, San Francisco, 1979. MR80g:68056 [10, 226, 443, 467].

M. R. GAREY, D. S. JOHNSON AND L. STOCKMEYER

1976. Some simplified NP-complete graph problems, *Theoret. Comput. Sci.* **1**, 1976, 237–267. MR53#14978 [443].

R. GILES

1978. Facets and other faces of branching polyhedra, *Combinatorics*, I, Eds.: A. Hajnal and V. T. Sós, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam, 1978, 401–418. MR80i:05043 [305].

C. D. GODSIL

1981a. Matchings and walks in graphs, *J. Graph Theory* **5**, 1981, 285–297. MR83b:05102 [336].

1981b. Matching behaviour is asymptotically normal, *Combinatorica* **1**, 1981, 369–376. MR83c:05099 [xx, 341, 342, 343, 344].

C. D. GODSIL AND I. GUTMAN

1981. On the matching polynomial of a graph, *Algebraic Methods in Graph Theory*, I, Eds.: L. Lovász and V. T. Sós, Colloq. Math. Soc. János Bolyai, 25, North-Holland, Amsterdam, 1981, 241–249. MR83b:05101 [330].

M. GOLUMBIC

1980. *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980. MR81e:68081 [459].

R. E. GOMORY

1958. Outline of an algorithm for integer solutions to linear programs, *Bull. Amer. Math. Soc.* **64**, 1958, 275–278. MR21#1230 [283].

1963. An algorithm for integer solutions to linear programs, *Recent Advances in Mathematical Programming*, Eds.: R. L. Graves and P. Wolfe, McGraw-Hill, New York, 1963, 269–302. MR30#4594 [283].

R. GOMORY AND T. C. HU

1961. Multi-terminal network flows, *J. SIAM* **9**, 1961, 551–570. MR24 #B1671 [61].

M. GONDRAN AND M. MINOUX

1979. *Graphes et Algorithmes*, Editions Eyrolles, Paris, 1979. MR82g:68059 [369, 382, 481].

1984. *Graphs and Algorithms*, Wiley-Interscience, Chichester, 1984 [369, 382, 481].

R. L. GRAHAM AND P. HELL

1982. On the history of the minimum spanning tree problem, Simon Fraser Univ. Dept. of Comput. Sci. Tech. Report TR 82-5, 1982 [248].

R. L. GRAHAM, B. L. ROTHSCHILD AND J. SPENCER

1980. *Ramsey Theory,* John Wiley, New York, 1980. MR82b:05001 [332].

J. E. GRAVER AND W. B. JURKAT

1980. f-factors and related decompositions of graphs, *J. Combin. Theory Ser. B* **28**, 1980, 66–84. MR81c:05074 [384].

J. R. GRIGGS

1983. Bounds on the independence number in terms of the degrees, *J. Combin. Theory Ser. B* **34**, 1983, 22–39. MR84g:05060 [481].

M. GRÖTSCHEL

1982. Approaches to hard combinatorial optimization problems, Part IV, Chapt. 2 in: *Modern Applied Mathematics — Optimization and Operations Research,* Ed.: B. Korte, North-Holland, Amsterdam, 1982, 437–515. MR84h:90085 [382].

M. GRÖTSCHEL, L. LOVÁSZ AND A. SCHRIJVER

1981. The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica* **1**, 1981, 169–197. MR84a:90044 [xx, 265, 266, 380, 417, 463, 482].

1982. Polynomial algorithms for perfect graphs, *Bonn Workshop on Combinatorial Optimization,* Eds.: A. Bachem, M. Grötschel and B. Korte, Ann. Discrete Math., 16, North-Holland, Amsterdam, 1982, 439–500 [463].

1984a. Relaxations of vertex packing, Univ. Augsburg Preprint No. 35, 1984 [482].

1984b. Corrigendum to our paper: "The ellipsoid method and its consequences in combinatorial optimization", *Combinatorica,* to appear [418].

1984c. Geometric methods in combinatorial optimization, *Progress in Combinatorial Optimization,* Ed.: W. R. Pulleyblank, Academic Press, Orlando, Florida, 1984, 167–183 [266].

B. GRÜNBAUM

1970a. Problem 2. 1-factors. *Combinatorial Structures and their Applications,* Eds.: R. Guy, H. Hanani, N. Sauer and J. Schönheim, Gordon and Breach, New York, 1970, 492 [407].

1970b. Polytopes, graphs and complexes, *Bull. Amer. Math. Soc.* **76**, 1970, 1131–1201. MR42#959 [119].

M. GUAN

1962. Graphic programming using odd and even points, *Chinese Math.* **1**, 1962, 273–277. MR28#5828 [ix, 231, 247].

R. P. GUPTA

1967. A decomposition theorem for bipartite graphs (results), *Theory of Graphs (International Symposium, Rome, 1966)*, Ed.: P. Rosenstiehl, Gordon and Breach, New York, 1967, 135–136 [39].

1974. On decompositions of a multi-graph into spanning subgraphs, *Bull. Amer. Math. Soc.* **80**, 1974, 500–502. MR49#149 [287].

I. GUTMAN

1977. The acyclic polynomial of a graph, *Publ. Inst. Math. (Beograd) (N. S.)* **22**, 1977, 63–69. MR58#27615 [352].

1982. Topological properties of benzenoid molecules, *Bull. de la Societe Chemique Beograd* **47**, 1982, 453–471 [349, 352].

I. GUTMAN, M. MILUN AND N. TRINAJSTIĆ

1975. Topological definition of delocalisation energy, *Match* **1**, 1975, 171–175 [352].

I. GUTMAN AND B. MOHAR

1981. More difficulties with topological resonance energy, *Chem. Phys. Letters* **77**, 1981, 567–570 [352].

L. G. HAČIJAN

1979. A polynomial algorithm in linear programming, *Akad. Nauk SSSR. Dokl.* **244**, 1979, 1093–1096. (Russian). (English translation: *Soviet Math. Dokl.*, American Mathematical Society, Providence **20**, 1979, 191–194. MR80g:90071.) [xx, 262].

F. O. HADLOCK

1975. Finding a maximum cut of a planar graph in polynomial time, *SIAM J. Comput.* **4**, 1975, 221–225. MR53#196. (Erratum: MR55#13865) [640].

M. HAIMOVICH

1983. The simplex algorithm is very good! — on the expected number of pivot steps and related properties of random linear programs, preprint, 1983 [263].

A. HAJNAL

1965. A theorem on k-saturated graphs, *Canad. J. Math.* **17**, 1965, 720–724. MR31#3354 [449, 450, 451, 453].

S. L. HAKIMI

1962. On realizability of a set of integers as degrees of the vertices of a linear graph, I, *J. SIAM* **10**, 1962, 496–506. MR26#5558 [406].

1965. On the degrees of the vertices of a directed graph, *J. Franklin Inst.* **279**, 1965, 290–308. MR31#4736 [74].

M. HALL JR.

1945. An existence theorem for Latin squares, *Bull. Amer. Math. Soc.* **51**, 1945, 387–388. MR7, 106 [xvii].

1948. Distinct representatives of subsets, *Bull. Amer. Math. Soc.* **54**, 1948, 922–926. MR10, 238g [347].

1956. An algorithm for distinct representatives, *Amer. Math. Monthly*, 1956, 716–717. MR18, 867 [xix].

P. HALL

1935. On representatives of subsets, *J. London Math. Soc.* **10**, 1935, 26–30. Zbl. 10.345 [xvi, 5, 29].

P. R. HALMOS

1950. *Measure Theory*, Van Nostrand, Princeton, N. J., 1950. MR11, 504 [74].

1958. *Finite Dimensional Vector Spaces*, Van Nostrand, New York, 1958. MR19-725 [328].

P. R. HALMOS AND H. E. VAUGHN

1950. The marriage problem, *Amer. J. Math.* **72**, 1950, 214–215. MR11, 423 [105].

J. H. HALTON

1966. A combinatorial proof of a theorem of Tutte, *Math. Proc. Cambridge Philos. Soc.* **62**, 1966, 683–684. MR34#1225 [85].

P. L. HAMMER, P. HANSEN AND B. SIMEONE

1980. Upper planes of quadratic 0–1 functions and stability in graphs, *Nonlinear Programming 4 (Madison, Wisc. 1980)*, Academic Press, New York, 1980, 394–414. MR83k:90065 [482].

P. L. HAMMER, N. V. R. MAHADEV AND D. DE WERRA

1983. The struction of a graph: application to CN-free graphs, Univ. Waterloo Dept. of Combinatorics and Optimization Research Report CORR 83-21, 1983 [480].

P.HANSEN

1979. Upper bounds for the stability number of a graph, *Rev. Roumaine Math. Pures Appl.* **24**, 1979, 1195–1199. MR80m:05040 [482].

1980. Bornes et algorithmes pour les stables d'un graphe, *Regards sur la Théorie des Graphes*, Eds.: P. Hansen and D. de Werra, Presses Polytechniques Romandes, Lausanne, 1980, 38–53. MR83d:05082 [482].

F. HARARY

1968. The cell growth problem and its attempted solutions, in: *Beiträge zur Graphentheorie*, Eds.: H. Sachs, H-J. Voss and H. Walther, Teubner, Leipzig, 1968, 49–60 [351].

1969. *Graph Theory*, Addison-Wesley, Reading, Mass., 1969. MR41#1566 [117].

D. J. HARTFIEL

1970. A simplified form for nearly reducible and nearly decomposable matrices, *Proc. Amer. Math. Soc.* **24**, 1970, 388–393. MR40#5635. (Erratum: See MR41#1965) [141].

D. HAUSMANN AND B. KORTE

1981. Algorithmic versus axiomatic definitions of matroids, *Mathematical Programming at Oberwolfach*, Eds.: H. König, B. Korte and K. Ritter, Math. Programming Stud. No. 14, North-Holland, Amsterdam, 1981, 98–111. MR82f:05025 [414].

V. HAVEL

1955. Eine Bemerkung über die Existenz der endlichen Graphen, *Časopis Pěst. Mat.* **80**, 1955, 477–480. (Czech). MR19–627 [406].

O. J. HEILMANN AND E. H. LIEB

1970. Monomers and dimers, *Phys. Rev. Lett.* **24**, 1970, 1412–1414 [xx, 339].

1972. Theory of monomer dimer systems, *Comm. Math. Phys.* **25**, 1972, 190–232. MR45#6337 [xx, 339].

T. HELGASON

1974. Aspects of the theory of hypermatroids, *Hypergraph Seminar*, Eds.: C. Berge and D. K. Ray-Chaudhuri, Lecture Notes in Math. **411**, 1974, 191–214. MR51#7909 [410].

P. HELL AND D. G. KIRKPATRICK

1983. Packings by cliques and by finite families of graphs, preprint, 1983 [231].

R. L. HEMMINGER

1971. Characterization of the line graph of a multigraph, *Notices Amer. Math. Soc.* 18, 1971, 934 [472].

W C. HERNDON

1974. Resonance theory and the enumeration of Kekulé structures, *J. Chem. Ed.* **51**, 1974, 10–15 [349].

G. HETYEI

1964. Rectangular configurations which can be covered by 2×1 rectangles, *Pécsi Tan. Főisk. Közl.* 8, 1964, 351–367. (Hungarian) [122, 123, 174, 191].

1972. A new proof of a factorization theorem, *Acta Acad. Paedagog. Civitate Pécs Ser. 6 Math. Phys. Chem. Tech.* **16**, 1972, 3–6. (Hungarian). MR58 #27647 [85].

D. HILBERT

1889. Über die Endlichkeit des Invariantensystems für binäre Grundformen, *Math. Ann.* **33**, 1889, 223–226. Jbuch. 20.110 [xi].

A. J. HOFFMAN

1960. Some recent applications of the theory of linear inequalities to extremal combinatorial analysis, *Proc. Sympos. Appl. Math.* **10**, Eds.: R. Bellman and M. Hall, Jr., American Mathematical Society, Providence, 1960, 113–127. MR22#5578 [xx, 3].

1974. A generalization of max flow-min cut, *Math. Programming* **6**, 1974, 352–359. MR50#15906 [272].

A. J. HOFFMAN AND J. B. KRUSKAL

1956. Integral boundary points of convex polyhedra, *Linear Inequalities and Related Systems*, Eds.: H. W. Kuhn and A. W. Tucker, Annals of Math. Study No. 38, Princeton University Press, Princeton, N. J., 1956, 223–246. MR18, 980b [xx, 255, 272].

A. J. HOFFMAN AND H. W. KUHN

1956a. Systems of distinct representatives and linear programming, *Amer. Math. Monthly* **63**, 1956, 455–460. MR18, 370 [30].

1956b. On systems of distinct representatives, *Linear Inequalities and Related Systems*, Eds.: H. W. Kuhn and A. W. Tucker, Ann. of Math. Studies **38**, Princeton University Press, Princeton, N. J., 1956, 199–206. MR18, 416 [30].

A. J. HOFFMAN AND H. W. WIELANDT

1953. The variation of the spectrum of a normal matrix, *Duke Math. J.* **20**, 1953, 37–39. MR14, 611 [36].

I. HOLYER

1981. The NP-completeness of edge-coloring, *SIAM J. Comput.* **10**, 1981, 718–720. MR83b:68050 [254, 285, 389].

M. P. HOMENKO AND T. M. VIVROT

1971. Structure of prime graphs, *Topological Aspects of Graph Theory, Izdanie Inst. Mat. Akad. Nauk Ukrain. SSR*, Kiev, 1971, 64–82 and 299–300. (Ukrainian). MR 53#178 [86].

1973. Structure of hyperprimitive graphs, *Φ-transformations of Graphs*, Vidannja Inst. Mat. Akad. Nauk Ukrain. RSR, Kiev, 1973, 307–317. (Ukrainian). MR 53#2760 [86].

J. E. HOPCROFT AND R. M. KARP

1971. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs, *Proceedings of the 12th Annual Symposium on Switching and Automata Theory (East Lansing, 1971)*, IEEE Computer Society Press, New York, 1971, 122–125 [15, 141].

1973. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs, *SIAM J. Comput.* **2**, 1973, 225–231. MR49#2468 [15, 141].

J. E. HOPCROFT AND R. E. TARJAN

1973a. Dividing a graph into triconnected components, *SIAM J. Comput.* **2**, 1973, 135–158. MR48#5733 [158].

1973b. Algorithm 447: Efficient algorithms for graph manipulation, *Comm. ACM* **16**, 1973, 372–378 [55].

1974. Efficient planarity testing, *J. Assoc. Comput. Mach.* **21**, 1974, 549–568. MR50#11841 [55].

T. C. HU

1963. Multi-commodity network flows, *Oper. Res.* **11**, 1963, 344–360 [254].

1969. *Integer Programming and Network Flows*, Addison-Wesley, Reading, Mass., 1969. MR41#8025 [42, 62].

E. ISING

1925. Beitrag zur Theorie des Ferromagnetismus, *Z. Physik* **31**, 1925, 253–258 [352].

K. JACOBS

1969. Der Heiratssatz, *Selecta Mathematica I*, Springer-Verlag, Heidelberg, 1969, 103–141. MR44#1 [3, 5].

P. M. JENSEN AND B. KORTE

1982. Complexity of matroid property algorithms, *SIAM J. Comput.* **11**, 1982, 184–190. MR83m:68078 [414].

R. G. JEROSLOW

1973. The simplex algorithm with the pivot rule of maximizing criterion improvement, *Discrete Math.* **4**, 1973, 367–377. MR51#7611 [262].

D. M. JOHNSON, A. L. DULMAGE AND N. S. MENDELSOHN

1960. On an algorithm of G. Birkhoff concerning doubly stochastic matrices, *Canad. Math. Bull.* **3**, 1960, 237–242. MR24#A133 [37].

D. B. JUDIN AND A. S. NEMIROVSKIĬ

1976. Informational complexity and effective methods for convex extremal problems, *Ékonom. i Mat. Metody* **12**, 1976, 357–369. (Russian). MR56 #10965 [xx, 263].

M. JÜNGER, W. R. PULLEYBLANK AND G. REINELT

1983. On partitioning the edges of graphs into connected subgraphs, Univ. of Waterloo, Dept. of Combinatorics and Optimization, Research Report CORR 83-8, March, 1983 [110].

T. KAMAE, U. KRENGEL AND G. L. O'BRIEN

1977. Stochastic inequalities on partially ordered spaces, *Ann. Probab.* **5**, 1977, 899–912. MR58#13308 [78].

T. KAMEDA AND I. MUNRO

1974. A $O(|V||E|)$ algorithm for maximum matching of graphs, *Computing* **12**, 1974, 91–98. MR53#4607 [369].

L. V. KANTOROVICH

1939. Mathematical methods in the organization and planning of production, Publishing House of the Leningrad State Univ., 1939, 68 pg. (Russian). (translated in: *Management Sci.* **6**, 1960, 366–422. MR23 #B2053.) [xix, 261].

1942. On the translocation of masses, *Comp. Rend. Acad. USSR (Doklady)* **37**, 1942, 199–201. (Russian). (translated in: *Management Sci.* **5**, 1958, 1–4. MR20#3035.) [261].

L. V. KANTOROVICH AND M. K. GAVURIN

1949. The application of mathematical methods to problems of freight flow analysis, *Akad. Nauk SSSR*, 1949. (English translation: *Akad. Nauk SSSR*, 1949.) [261].

O. KARIV

1976. An $O(n^{5/2})$ algorithm for finding a maximum matching in a general graph, Weizmann Institute of Science, Ph.D. Thesis, Rehovot, 1976 [369].

N. KARMARKAR

1984. A new polynomial-time algorithm for linear programming, preprint, 1984 (to appear in *Combinatorica* (1985)) [266].

R. M. KARP

1972. Reducibility among combinatorial problems, *Complexity of Computer Computations*, Eds.: R. E. Miller and J. W. Thatcher, Plenum Press, New York, 1972, 85–103. MR51#14644 [226, 249, 252, 413, 443, 467].

1975. On the computational complexity of combinatorial problems, *Networks* **5**, 1975, 45–68. Zbl. 324.05003 [226, 249, 252].

R. M. KARP AND C. H. PAPADIMITRIOU

1980. On linear characterizations of combinatorial optimization problems, *21st Annual Symposium on Foundations of Computer Science (Syracuse,* *1980)*, IEEE Computer Society Press, New York, 1980, 1–9. MR81k: 90064 [266, 380].

1982. On linear characterization of combinatorial optimization problems, *SIAM J. Comput.* **11**, 1982, 620–632. MR83k:90070 [266, 380].

A. V. KARZANOV

1974. The problem of finding the maximal flow in a network by the method of preflows, *Dokl. Akad. Nauk SSSR* **215**, 1974, 49–52. (Russian). MR49 #8619. (English translation by American Mathematical Society: *Soviet Math. Dokl.* **215**, 1974, 434–437.) [56, 57].

P. W. KASTELEYN

1961. The statistics of dimers on a lattice. I. The number of dimer arrangements on a quadratic lattice, *Physica* **27**, 1961, 1209–1225 [xx].

1963. Dimer statistics and phase transitions, *J. Math. Phys.* **4**, 1963, 287–293. MR27#3394 [xx, 321].

1967. Graph theory and crystal physics, *Graph Theory and Theoretical Physics*, Ed.: F. Harary, Academic Press, New York, 1967, 43–110. MR40 #6903 [321, 355].

D. G. KIRKPATRICK AND P. HELL

1978. On the completeness of a generalized matching problem, *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing (San Diego, 1978)*, ACM, New York, 1978, 240–245. MR80d:68053 [231].

1983. On the complexity of general graph factor problems, *SIAM J. Comput.* **12**, 1983, 601–609 [231].

V. KLEE AND G. J. MINTY

1972. How good is the simplex algorithm?, *Inequalities III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)*, Ed.: O. Shisha, Academic Press, New York, 1972, 159–175. MR48#10492 [262].

D. KÖNIG

1915. Line systems and determinants, *Math. Termész. Ért.* **33**, 1915, 221–229. (Hungarian). Jbuch. 45.1240 [xiv].

1916a. Über Graphen und ihre Andwendung auf Determinantentheorie und Mengenlehre, *Math. Ann.* **77**, 1916, 453–465. Jbuch. 46.146–147 [xiv, xv, 15, 35, 37, 137, 285].

1916b. Graphok és alkalmazásuk a determinánsok és a halmazok elméletére, *Math. Termész. Ért.* **34**, 1916, 104–119. Jbuch. 46.1451–1452 [xiv, xv, 15, 35, 37, 137, 285].

1931. Graphs and matrices, *Mat. Fiz. Lapok* **38**, 1931, 116–119. (Hungarian). Zbl. 3.328 [xvi, 15].

1933. Über trennende Knotenpunkte in Graphen (nebst Anwendungen auf Determinanten und Matrizen), *Acta Sci. Math. (Szeged)* **6**, 1933, 155–179. Zbl. 7.329 [xv, xvi].

1936. *Theorie der endlichen und unendlichen Graphen*, Akademischen Verlagsgesellschaft, Leipzig, 1936. Zbl. 13.228. (reprinted: Chelsea, New York, 1950. MR12-195.) [xii, xv, xvi, xvii, 15, 70, 110].

E. KORACH

1982. On dual integrality, min-max equalities and algorithms in combinatorial programming, Univ. of Waterloo, Dept. of Combinatorics and Optimization, Ph.D. Thesis, 1982 [223, 247].

A. KOTZIG

1959a. On the theory of finite graphs with a linear factor I, *Mat.-Fyz. Časopis Slovensk. Akad. Vied* **9**, 1959, 73–91. (Slovak). MR26#6955 [xviii, 143, 159, 345].

1959b. On the theory of finite graphs with a linear factor II, *Mat.-Fyz. Časopis Slovensk. Akad. Vied* **9**, 1959, 136–159. (Slovak). MR26#6955 [xviii, 143, 159, 172].

1960. On the theory of finite graphs with a linear factor III, *Mat.-Fyz. Časopis Slovensk. Akad. Vied* **10**, 1960, 205–215. (Slovak). MR26#6955 [xviii, 143, 159].

J. KRAUSZ

1943. Démonstration nouvelle d'un théorème de Whitney sur les réseaux, *Mat. Fiz. Lapok* **50**, 1943, 75–85. MR8, 284h [473].

J. B. KRUSKAL, JR.

1956. On the shortest spanning subtree of a graph and the traveling salesman problem, *Proc. Amer. Math. Soc.* **7**, 1956, 48–50. MR17, 1231d [x, 248].

H. W. KUHN

1955. The Hungarian method for the assignment problem, *Naval Res. Logist. Quart.* **2**, 1955, 83–97. MR17, 759d [xix, 15].

1956. Variants of the Hungarian method for assignment problems, *Naval Res. Logist. Quart.* **3**, 1956, 253–258. MR19-1024 [xix, 15].

S. KUNDU

1973. The k-factor conjecture is true, *Discrete Math.* **6**, 1973, 367–376. MR49#152 [407, 408].

S. KUNDU AND E. L. LAWLER

1973. A matroid generalization of a theorem of Mendelsohn and Dulmage, *Discrete Math.* **4**, 1973, 159–163. MR47#57 [30].

H. KUNZ

1970. Location of the zeros of the partition function for some classical lattice systems, *Phys. Lett. (A)*, 1970, 311–312 [339].

K. KURATOWSKI

1930. Sur le problème des courbes gauches en topologie, *Fund. Math.* **16**, 1930, 271–283. Jbuch. 56.1141 [323].

M. LAS VERGNAS

1970. Transversales disjointes d'une famille d'ensembles, preprint, 1970 [20].

1975. A note on matchings in graphs, *Colloque sur la Théorie des Graphes (Paris, 1974)*, Cahiers Centre Études Rech. Opér. **17**, 1975, 257–260. MR54#171 [109].

E. L. LAWLER

1976. *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, New York, 1976. MR55#12005 [15, 42, 57, 369, 376].

P. G. H. LEHOT

1974. An optimal algorithm to detect a line graph and output its root graph, *J. Assoc. Comput. Mach.* **21**, 1974, 569–575. MR50#192 [472].

A. LEMPEL, S. EVEN AND I. CEDERBAUM

1967. An algorithm for planarity testing for graphs, *Theory of Graphs (International Symposium, Rome, 1966)*, Ed.: P. Rosenstiehl, Gordon and Breach, New York, 1967, 215–232. MR36#3669 [55].

M. LESK, M. D. PLUMMER AND W. R. PULLEYBLANK

1984. Equi-matchable graphs, *Graph Theory and Combinatorics: Proceedings of the Cambridge Combinatorial Conference in Honour of Paul Erdős*, Ed.: B. Bollobás, Academic Press, London, 1984, 239–254 [102].

L. A. LEVIN

1973. Universal search problems, *Problemy Peredači Informacii* **9**, 1973, 115–116. MR49#4799 [226].

W. W. LI AND S. R. LI

1981. Independence numbers of graphs and generators of ideals, *Combinatorica* **1**, 1981, 55–61. MR82h:05029 [482].

S. LINS

1981. A minimax theorem on circuits in projective graphs, *J. Combin. Theory Ser. B* **30**, 1981, 253–262. MR82j:05074 [253, 254].

C. H. C. LITTLE

1972. The parity of the number of 1-factors of a graph, *Discrete Math.* **2**, 1972, 179–181. MR45#6664 [309].

1973. Kasteleyn's theorem and arbitrary graphs, *Canad. J. Math.* **25**, 1973, 758–764. MR48#5923 [324].

1974a. An extension of Kasteleyn's method of enumerating the 1-factors of planar graphs, *Combinatorial Mathematics, Proc. Second Australian Conference*, Ed.: D. Holton, Lecture Notes in Math. **403**, Springer-Verlag, Berlin, 1974, 63–72. MR52#2950 [323].

1974b. A theorem on connected graphs in which every edge belongs to a 1-factor, *J. Austral. Math. Soc.* **18**, 1974, 450–452. MR52#7971 [132, 177].

C. H. C. LITTLE, D. D. GRANT AND D. A. HOLTON

1975. On defect-d matchings in graphs, *Discrete Math.* **13**, 1975, 41–54. MR54#5043a [113].

1976. Erratum: "On defect-d matchings", *Discrete Math.* **14**, 1976, 203. MR54#5043b [113].

C. H. C. LITTLE AND J. PLA

1972. Sur l'utilisation d'un pfaffien dans l'étude des couplages parfaits d'un graphe, *C. R. Acad. Sci. Paris Sér. I Math.* **274**, 1972, A447. MR44#6532 [324].

C. L. LIU

1968. *Introduction to Combinatorial Mathematics*, McGraw-Hill, New York, 1968. MR38#3154 [43].

L. C. LORENTZEN

1966. Notes on covering of arcs by nodes in an undirected graph, Tech. Rept. ORC 66.16, Univ. California, Berkeley, 1966 [292].

L. LOVÁSZ

1970a. The factorization of graphs, *Combinatorial Structures and their Applications*, Eds.: R. Guy, H. Hanani, N. Sauer and J. Schönheim, Gordon and Breach, New York, 1970, 243–246. MR42#2970 [384].

1970b. Subgraphs with prescribed valencies, *J. Combin. Theory* **8**, 1970, 391–416. MR42#113 [384, 402].

1970c. A generalization of König's theorem, *Acta Math. Acad. Sci. Hungar.* **21**, 1970, 443–446. MR42#5811 [20].

1970d. Generalized factors of graphs, *Combinatorial Theory and its Applications* II, Eds.: P. Erdős, A. Rényi and V. T. Sós, Colloq. Math. Soc. János Bolyai, 4, North-Holland, Amsterdam, 1970, 773–781. MR46#93 [254].

1972a. On the structure of factorizable graphs, *Acta Math. Acad. Sci. Hungar.* **23**, 1972, 179–195. MR47#77 [143, 159, 173].

1972b. On the structure of factorizable graphs, II, *Acta Math. Acad. Sci. Hungar.* **23**, 1972, 465–478. MR47#77 [384].

1972c. The factorization of graphs II, *Acta Math. Acad. Sci. Hungar.* **23**, 1972, 223–246. MR48#3811 [384, 389].

1972d. A note on factor-critical graphs, *Studia Sci. Math. Hungar.* 7, 1972, 279–280. MR49#153 [196].

1972e. Normal hypergraphs and the perfect graph conjecture, *Discrete Math.* **2**, 1972, 253–276. MR46#1624 [461, 462, 470].

1972f. A characterization of perfect graphs, *J. Combin. Theory Ser. B* **13**, 1972, 95–98. MR46#8885 [462].

1973. Antifactors of graphs, *Period. Math. Hungar.* **1973**, 4, 121–123. MR51 #10165 [389].

1974. Valencies of graphs with 1-factors, *Period. Math. Hungar.* **5**, 1974, 149–151. MR50#6927 [407].

1975a. 2-matchings and 2-covers of hypergraphs, *Acta Math. Acad. Sci. Hungar.* **26**, 1975, 433–444. MR54#7320 [234, 240].

1975b. Three short proofs in graph theory, *J. Combin. Theory Ser. B* **19**, 1975, 269–271. MR53#211 [4, 85].

1976a. On two minimax theorems in graph theory, *J. Combin. Theory Ser. B* **21**, 1976, 96–103. MR55#174 [250].

1976b. On some connectivity properties of Eulerian graphs, *Acta Math. Acad. Sci. Hungar.* **28**, 1976, 129–138. MR55#10321 [254].

1977. Flats in matroids and geometric graphs, *Combinatorial Surveys (Proceedings of the Sixth British Combinatorial Conference)*, Ed.: P. Cameron, Academic Press, New York, 1977, 45–86. MR58#310 [410, 411, 450, 452].

1978. Some finite basis theorems in graph theory, *Combinatorics*, II, Eds.: A. Hajnal and V. T. Sós, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam, 1978, 717–729. MR80c:05087 [134, 453].

1979a. Graph theory and integer programming, *Discrete Optimization I*, Eds.: P. L. Hammer, E. L. Johnson and B. Korte, Ann. Discrete Math., 4, North-Holland, Amsterdam, 1979, 141–158. Zbl. 407.05053 [274].

1979b. On determinants, matchings and random algorithms, *Fundamentals of Computation Theory, FCT '79, (Proc. Conf. Algebraic, Arithmetic and Categorical Methods in Computation Theory, Berlin/Wendisch-Rietz 1979)*, Ed.: L. Budach, Math. Research 2, Akademie-Verlag, Berlin, 1979, 565–574. MR81k:05094 [333, 420].

1979c. *Combinatorial Problems and Exercises*, North-Holland, Amsterdam, 1979. MR80m:05001 [110, 220, 328, 340, 466].

1979d. On the Shannon capacity of a graph, *IEEE Trans. Inform. Theory* **25**, 1979, 1–7. MR81g:05095 [482].

1980. Matroid matching and some applications, *J. Combin. Theory Ser. B* **28**, 1980, 208–236. MR82g:05038 [421, 425].

1981. The matroid matching problem, *Algebraic Methods in Graph Theory*, II, Eds.: L. Lovász and V. T. Sós, Colloq. Math. Soc. János Bolyai, 25, North-Holland, Amsterdam, 1981, 495–517. MR84h:05033 [411, 414, 419, 430].

1982. Bounding the independence number of a graph, *Bonn Workshop on Combinatorial Optimization*, Eds.: A. Bachem, M. Grötschel and B. Korte, Ann. Discrete Math., 16, North-Holland, Amsterdam, 1982, 213–223. MR84i:05067 [482].

1983a. Submodular functions and convexity, *Mathematical Programming, the State of the Art: Bonn, 1982*, Eds.: A. Bachem, M. Grötschel and B. Korte, Springer-Verlag, Heidelberg, 1983, 235–257 [24, 411].

1983b. Perfect graphs, in: *Selected Topics in Graph Theory 2*, Eds.: L. W. Beineke and R. J. Wilson, Academic Press, New York, 1983, 55–87 [459].

1983c. Ear-decompositions of matching-covered graphs, *Combinatorica* **2**, 1983, 395–407. Zbl. 516.05047 [186, 225].

1984. Algorithmic aspects of some notions in classical mathematics, Univ. Bonn Research Report No. WP-84329-OR, 1984 [61].

L. LOVÁSZ AND P. MAJOR

1973. A note on a paper of Dudley, *Studia Sci. Math. Hungar.* 8, 1973, 151–152. MR48#7329 [76].

L. LOVÁSZ AND M. D. PLUMMER

1975a. On bicritical graphs, *Infinite and Finite Sets (Colloq. Keszthely, Hungary, 1973)*, II, Eds.: A. Hajnal, R. Rado and V. T. Sós, Colloq. Math. Soc. János Bolyai, 10, North-Holland, Amsterdam, 1975, 1051–1079. MR51#10164 [200].

1975b. On a family of planar bicritical graphs, *Proc. London Math. Soc. Ser. (3)* **30**, 1975, 160–176. MR55#7825 [205].

1977. On minimal elementary bipartite graphs, *J. Combin. Theory Ser. B* **23**, 1977, 127–138. MR58#16122 [122, 127, 128].

L. LOVÁSZ AND Y. YEMINI

1982. On generic rigidity in the plane, *SIAM J. Algebraic Discrete Methods* **3**, 1982, 91–99. MR83b:52007 [420].

E. LUCAS

1882. *Recreations Mathematiques* **1**, Gauthiers-Villars, Paris, 1882. Jbuch. 15.158 [53].

C. L. LUCCHESI AND D. H. YOUNGER

1978. A minimax theorem for directed graphs, *J. London Math. Soc. Ser. (2)* **17**, 1978, 369–374. MR80e:05062 [250].

W. MAAK

1935. Eine neue Definition der fastperiodischen Funktionen, *Abh. Math. Sem. Hamburg* **11**, 1935, 240–244. Zbl. 13.111 [5].

W. MADER

1971. Minimale n-fach kantenzusammenhangende Graphen, *Math. Ann.* **191**, 1971, 21–28. MR45#98 [208].

1973. 1-Faktoren von Graphen, *Math. Ann.* **201**, 1973, 269–282. MR50 #12807 [85, 345].

1976. Über die Anzahl der 1-Faktoren in 2-fach zusammenhängenden Graphen, *Math. Nachr.* **74**, 1976, 217–232. MR55#2665 [346].

1978a. Über die Maximalzahl kantendisjunkter A-Wege, *Arch. Math. (Basel)* **30**, 1978, 325–336. MR80a:05121 [254].

1978b. Über die Maximalzahl kreuzungsfreier H-Wege, *Arch. Math. (Basel)* **31**, 1978, 387–402. MR80f:05058 [254, 425].

1979a. Über ein graphentheoretisches Problem von T. Gallai, *Arch. Math. (Basel)* **33**, 1979, 239–257. MR81f:05109 [254].

1979b. Connectivity and edge-connectivity in finite graphs, *Surveys in Combinatorics (Proceedings of the Seventh British Combinatorial Conference, Cambridge, 1979)*, Ed.: B. Bollobás, London Math. Soc. Lecture Note Ser. No. 38, Cambridge University Press, London, 1979, 66–95. MR81d:05044 [254].

V. M. MALHOTRA, M. P. KUMAR AND S. N. MAHESHWARI

1978. An $O(V^3)$ algorithm for finding maximum flows in networks, *Inform. Process. Lett.* **7**, 1978, 277–278. MR80b:90058 [57].

A. MANDEL

1982. Topology of oriented matroids, Ph.D. Thesis, Dept. of Combinatorics and Optimization, Univ. of Waterloo, 1982 [262].

A. B. MARSH, III

1979. Matching algorithms, Johns Hopkins Univ., Ph.D. Thesis, 1979 [387].

C. MARTEL

1981. Generalized network flows with an application to multiprocessor scheduling, Univ. Calif. Berkeley, Comput. Sci. Dept., preprint, 1981 [24].

F. G. MAUNSELL

1952. A note on Tutte's paper: "The factorization of linear graphs", *J. London Math. Soc.* **27**, 1952, 127–128. MR13, 572e [85].

C. J. H. MCDIARMID

1975. Rado's theorem for polymatroids, *Math. Proc. Cambridge Philos. Soc.* **78**, 1975, 263–281. MR52#153 [410].

B. D. MCKAY

1982. Asymptotics for 0–1 matrices with prescribed line sums, preprint, 1982 [313].

N. S. MENDELSOHN AND A. L. DULMAGE

1958. Some generalizations of the problem of distinct representatives, *Canad. J. Math.* **10**, 1958, 230–241. MR20#1635 [30].

K. MENGER

1927. Zur allgemeinen Kurventheorie, *Fund. Math.* **10**, 1927, 96–115. Jbuch. 53.561 [xvi, 68, 70].

S. MICALI AND V. V. VAZIRANI

1980. An $O(V^{1/2}E)$ algorithm for finding maximum matching in general graphs, *21st Annual Symposium on Foundations of Computer Science (Syracuse, 1980)*, IEEE Computer Society Press, New York, 1980, 17–27. MR81i:68008 [369].

M. M. MILIĆ

1974. General passive networks-solvability, degeneracies and order of complexity, *IEEE Trans. Circuits and Systems*, CAS-**21**, 1974, 177–183. MR56#2685 [424].

H. MINC

1972. Nearly decomposable matrices, *Linear Algebra Appl.* **5**, 1972, 181–187. MR47#1832 [141].

1978. *Permanents*, Encyclopedia of Mathematics and its Applications **6**, Addison-Wesley, Reading, 1978. MR80d:15009 [310].

G. J. MINTY

1980. On maximal independent sets of vertices in claw-free graphs, *J. Combin. Theory Ser. B* **28**, 1980, 284–304. MR81f:68076 [445, 471].

L. MIRSKY

1963. Results and problems in the theory of doubly-stochastic matrices, *Z. Warsch. Verw. Gebiete* **1**, 1963, 319–334. MR27#3007 [37].

1971. *Transversal Theory*, Academic Press, New York, 1971. MR44#87 [5, 30, 40].

E. W. MONTROLL

1964. Lattice Statistics, *Applied Combinatorial Mathematics*, Ed.: E. Beckenbach, Wiley, New York, 1964, 96–143 [326].

T. S. MOTZKIN

1936. Beiträge zur Theorie der linearen Ungleichungen, dissertation, Basel, 1933, Jerusalem, 1936 [259].

A. MOWSHOWITZ

1972. The characteristic polynomial of a graph, *J. Combin. Theory Ser. B* **12**, 1972, 177–193. MR45#5011 [335].

T. MUIR

1882. *A Treatise on the Theory of Determinants*, MacMillan and Co., London, 1882. Jbuch. 14.100 [318].

1906. *The Theory of Determinants*, MacMillan and Co., London, 1906. Jbuch. 37.180 [318].

D. NADDEF

1982. Rank of maximum matchings in a graph, *Math. Programming* **22**, 1982, 52–70. MR83c:05100 [294].

D. NADDEF AND W. R. PULLEYBLANK

1982. Ear decompositions of elementary graphs and GF(2)-rank of perfect matchings, *Bonn Workshop on Combinatorial Optimization*, Eds.: A. Bachem, M. Grötschel and B. Korte, Ann. Discrete Math., 16, North-Holland, Amsterdam, 1982, 241–260. Zbl. 501.05050 [178].

C. ST. J. A. NASH-WILLIAMS

1960. On orientations, connectivity and odd vertex pairings in finite graphs, *Canad. J. Math.* **12**, 1960, 555–567. MR22#9455 [74].

1961. Edge-disjoint spanning trees of finite graphs, *J. London Math. Soc.* **36**, 1961, 445–450. MR24#A3087 [254].

1964. Decomposition of finite graphs into forests, *J. London Math. Soc.* **39**, 1964, 12. MR28#4541 [254].

G. L. NEMHAUSER AND L. E. TROTTER, JR.

1974. Properties of vertex packing and independence system polyhedra, *Math. Programming* **6**, 1974, 48–61. MR52#200 [292].

J. VON NEUMANN

1934. Zum Haarschen Mass in topologischen Gruppen, *Compositio Math.* **1**, 1934, 106–114. Zbl. 8.246 [79].

1947. On a maximization problem (manuscript), Institute for Advanced Study, Princeton Univ., Nov., 1947 [xix, 260].

1953. A certain zero-sum two-person game equivalent to the optimal assignment problem. *Contributions to the Theory of Games* II, Ed.: H. W. Kuhn, Ann. of Math. Studies, 28, Princeton University Press, Princeton, N. J., 1953, 5–12. MR14, 998 [xv, 36].

T. NISHIZEKI

1978. Lower bounds on the cardinality of the maximum matchings of graphs, *Proceedings of the Ninth Southeastern Conference on Combinatorics, Graph Theory and Computing*, Eds.: F. Hoffman, et.al., Congress. Numer. XXI, Utilitas Mathematica, Winnipeg, 1978, 527–547. MR81d:05041 [118].

1979. On the relationship between the genus and the cardinality of the maximum matchings of a graph, *Discrete Math.* **25**, 1979, 149–156. MR80g:05057 [118].

T. NISHIZEKI AND I. BAYBARS

1979. Lower bounds on the cardinality of the maximum matchings of planar graphs, *Discrete Math.* **28**, 1979, 255–267. MR81a:05079 [119].

R. Z. NORMAN AND M. O. RABIN

1959. An algorithm for a minimum cover of a graph, *Proc. Amer. Math. Soc.* **10**, 1959, 315–319. MR21#5583 [2].

H. OKAMURA AND P. D. SEYMOUR

1981. Multi-commodity flows in planar graphs, *J. Combin. Theory Ser. B* **31**, 1981, 75–81. MR82j:90033 [254].

P. E. O'NEIL

1969. Asymptotics and random matrices with row-sum and column sum-restrictions, *Bull. Amer. Math. Soc.* **75**, 1969, 1276–1282. MR41#1770 [313].

O. ORE

1955. Graphs and matching theorems, *Duke Math. J.* **22**, 1955, 625–639. MR17, 394d [xvii, 17].

1957. Graphs and subgraphs, *Trans. Amer. Math. Soc.* **84**, 1957, 109–136. MR18, 751g [384].

1962. *Theory of Graphs*, American Mathematical Society Colloquium Publications **38**, American Mathematical Society, Providence, 1962. MR27 #740 [xviii].

G. I. ORLOVA AND JA.G. DORFMAN

1972. Finding the maximum cut in a graph, *Engrg. Cybernetics* **10**, 1972, 502–506. MR48#8301 [252].

M. W. PADBERG AND M. R. RAO

1981. The Russian method for linear programming III: Bounded integer programming, New York Univ. Grad. School of Business Administration preprint, 1981 [266, 380].

1982. Odd minimum cut-sets and b-matchings, *Math. Oper. Res.* **7**, 1982, 67–80. MR84e:05063 [xxiii, 62, 251, 380, 381].

H. PERFECT

1966. Symmetrized form of P. Hall's theorem on distinct representatives, *Quart. J. Math. Oxford Ser. (2)* **17**, 1966, 303–306. MR35#75 [316].

J. PETERSEN

1891. Die Theorie der regulären Graphen, *Acta Math.* **15**, 1891, 193–220. Jbuch. 23.115 [xi, 110, 218, 404].

1898. Sur le théorème de Tait, *L'Intermediaire des Mathematiciens* **5**, 1898, 225–227 [xiii].

J. M. PLA

1965. Sur l'utilisation d'un pfaffien dans l'étude des couplages parfaits d'un graph, *C. R. Acad. Sci. Paris Sér. I Math.* **260**, 1965, 2967–2970. MR30 #4251 [324].

J. PLESNÍK

1972. Connectivity of regular graphs and the existence of 1-factors, *Mat. Časopis Sloven. Akad. Vied* **22**, 1972, 310–318. MR47#6548 [111].

1979. Remark on matchings in regular graphs, *Acta Fac. Rerum Natur. Univ. Comenian. Math.* **34**, 1979, 63–67. MR81g:05089 [113].

M. D. PLUMMER

1980. On n-extendable graphs, *Discrete Math.* **31**, 1980, 201–210. MR81k: 05083 [206].

S. POLJAK

1974. A note on stable sets and coloring of graphs, *Comment. Math. Univ. Carolin.* **15**, 1974, 307–309. MR50#4369 [443].

L. PONTRJAGIN

1939. *Topological Groups*, Princeton Univ. Press, Princeton, N. J., 1939. MR1-44 [79].

W. R. PULLEYBLANK

1973. Faces of Matching Polyhedra, Univ. of Waterloo, Dept. Combinatorics and Optimization, Ph.D. Thesis, 1973 [199, 274, 279, 387].

1979. Minimum node covers and 2-bicritical graphs, *Math. Programming* **17**, 1979, 91–103. MR81e:05088 [217].

1980. The matching rank of Halin graphs, *Methods of O. R.*, 40, Athenäum /Hain/Scriptor/Haustein, Königstein/Ts., 1980, 401–404 [205].

1983. Polyhedral combinatorics, in: *Mathematical Programming, the State of the Art: Bonn, 1982*, Eds.: A. Bachem, M. Grötschel and B. Korte, Springer-Verlag, Berlin, 1983, 312–345 [xx].

W. R. PULLEYBLANK AND J. EDMONDS

1974. Facets of 1-matching polyhedra, *Hypergraph Seminar*, Eds.: C. Berge and D. Ray-Chaudhuri, Lecture Notes in Math. **411**, Springer-Verlag, Berlin, 1974, 214–242. MR52#7960 [274].

M. O. RABIN

1976. Probabilistic algorithms, in: *Algorithms and Complexity, Proc. Sympos. Carnegie-Mellon Univ., Pittsburgh, Pa., 1976)*, Ed.: J. F. Traub, Academic Press, New York, 1976, 21–40. MR57#4603 [333].

1980. Probabilistic algorithm for testing primality, *J. Number Theory* **12**, 1980, 128–138. MR81f:10003 [333].

R. RADO

1942. A theorem on independence relations, *Quart. J. Math. Oxford* **13**, 1942, 83–89. MR4, 269c [xvii, 28].

1957. Note on independence functions, *Proc. London Math. Soc.* 7, 1957, 300–320. MR19, 522b [28].

R. VON RANDOW

1975. *Introduction to the Theory of Matroids*, Lecture Notes in Econom. and Math. Systems **109**, Springer-Verlag, Berlin, 1975. MR52#10460 [23, 28, 93].

A. RECSKI

1980. Sufficient conditions for the unique solvability of linear networks containing memoryless 2-ports, *Internat. J. Circuit Theory Appl.* 8, 1980, 95–103. MR81g:94050 [424].

J. RIORDAN

1958. *An Introduction to Combinatorial Analysis*, John Wiley and Sons, New York, 1958, 164–194. MR20#3077 [334].

J. T. ROBACKER

1955. On network theory, The RAND Corp., Research Memorandum RM-1498, May, 1955 [3].

H. E. ROBBINS

1939. A theorem on graphs with an application to a problem of traffic control, *Amer. Math. Monthly* **46**, 1939, 281–283. Zbl. 21.357 [74].

A. C. M. VAN ROOIJ AND H. S. WILF

1965. The interchange graph of a finite graph, *Acta Math. Acad. Sci. Hungar.* **16**, 1965, 263–269. MR33#3959 [472].

G.-C. ROTA AND L. H. HARPER

1971. Matching theory, an introduction, *Advances in Probability Theory and Related Topics*, Ed.: P. Ney **I**, Marcel Dekker, New York, 1971, 169–215. MR44#89 [78].

B. ROTHSCHILD AND A. WHINSTON

1966a. On two-commodity network flows, *Oper. Res.* **14**, 1966, 377–387 [254].

1966b. Feasibility of two commodity network flows, *Oper. Res.* **14**, 1966, 1121–1129. MR34#8809 [254].

N. D. ROUSSOPOLIS

1973. A $\max\{m, n\}$ algorithm for determining the graph H from its line graph G, *Inform. Process. Lett.* **2**, 1973, 108–112. MR54#12397 [472].

H. SACHS

1984. Perfect matchings in hexgonal systems, *Combinatorica* 4, 1984, 89–99 [352].

N. SBIHI

1980. Algorithme de recherche d'un stable de cardinalité maximum dans un graphe sans étoile, *Discrete Math.* **29**, 1980, 53–76. MR81e:68087 [445, 471].

H. SCHNEIDER

1977. The concepts of irreducibility and full indecomposability of a matrix in the works of Frobenius, König and Markov, *Linear Algebra Appl.* **18**, 1977, 139–162. MR56#5170 [xv].

A. SCHRIJVER

1978. A short proof of Minc's conjecture, *J. Combin. Theory Ser. A* **25**, 1978, 80–83. MR58#10481 [311].

1980a. A counterexample to a conjecture of Edmonds and Giles, *Discrete Math.* **32**, 1980, 213–214. MR81k:05084 [254].

1980b. On cutting planes, *Combinatorics '79*, Eds.: M. Deza and I. Rosenberg, Ann. Discrete Math., 9, North-Holland, Amsterdam, 1980, 291–296. MR82b:90080 [284].

1982. Min-max relations for directed graphs, *Bonn Workshop on Combinatorial Optimization*, Eds.: A. Bachem, M. Grötschel and B. Korte, Ann. Discrete Math. 16, North-Holland, Amsterdam, 1982, 261–280. MR84c:05045 [254].

1983a. Min-max results in combinatorial optimization, *Mathematical Programming, the State of the Art: Bonn, 1982*, Eds.: A. Bachem, M. Grötschel and B. Korte, Springer-Verlag, Berlin, 1983, 439–500 [xvi, 234, 387].

1983b. Bounds on permanents, and the number of 1-factors and 1-factorizations of bipartite graphs, *Surveys in Combinatorics*, Ed.: E. K. Lloyd, London Math. Soc. Lecture Note Ser.: 82, Cambridge Univ. Press, Cambridge, 1983, 107–134 [315].

A. SCHRIJVER AND W. G. VALIANT

1980. On lower bounds for permanents, *Nederl. Akad. Wetensch. Indag. Math.* **42**, 1980, 425–427. MR82a:15004 [312, 315].

J. T. SCHWARTZ

1979. Probabilistic algorithms for verification of polynomial identities, *Symbolic and Algebraic Computation (EUROSAM '79, Internat. Sympos., Marseille, 1979)*, Ed.: E. W. Ng, Lecture Notes in Comput. Sci. **72**, Springer-Verlag, Berlin, 1979, 200–215. MR81g:68060 [332].

1980. Fast probabilistic algorithms for verification of polynomial identities, *J. Assoc. Comput. Mach.* **27**, 1980, 701–717. MR82m:68078 [332].

A. SEBŐ

1984a. On the structure of odd joins, preprint, MO/54, Aug., 1984 [245].

1984b. Finding the t-join-structure of graphs, Eötvös Loránd Univ. preprint, 1984 [245].

S. SESHU AND M. B. REED

1961. *Linear Graphs and Electrical Networks*, Addison-Wesley, Reading, 1961. MR26#4638 [422].

P. D. SEYMOUR

1977. The matroids with the max-flow min-cut property, *J. Combin. Theory Ser. B* **23**, 1977, 189–222. MR57#2960 [240].

1978. A two-commodity cut theorem, *Discrete Math.* **23**, 1978, 177–181. MR80a:90068 [253].

1979a. On multicolourings of cubic graphs, and conjectures of Fulkerson and Tutte, *Proc. London Math. Soc. Ser. (3)* **38**, 1979, 423–460. MR81j:05061 [288, 289].

1979b. Sums of circuits, *Graph Theory and Related Topics*, Eds.: J. A. Bondy and U. S. R. Murty, Academic Press, New York, 1979, 341–355. MR81b:05068 [254].

1979c. A short proof of the two-commodity flow theorem, *J. Combin. Theory Ser. B* **26**, 1979, 370–371. MR80d:90037 [254].

1980. Decomposition of regular matroids, *J. Combin. Theory Ser. B* **28**, 1980, 305–359. MR82j:05046 [29].

1981a. On odd cuts and plane multicommodity flows, *Proc. London Math. Soc. Ser. (3)* **42**, 1981, 178–192. MR82g:05080 [236].

1981b. Matroids and multicommodity flows, *European J. Combin.* **2**, 1981, 257–290. MR82m:05030 [234].

Y. SHILOACH

1978. An $O(nI(\log I)^2)$ maximum flow algorithm, Stanford Univ., Comput. Sci. Dept., Tech. Report STAN-CS-78-802, 1978 [57].

1979. Edge-disjoint branchings in directed multigraphs, *Inform. Process. Lett.* 8, 1979, 24–27. MR81f:05112 [253].

Z. SKUPIEŃ

1973. A short proof of the theorem on the structure of maximal graphs without 1-factors, *Grafy i hipergrafy*, Práce Nauk. Inst. Mat. Fiz. Teoret. Politechn. Wrocław Ser. Stud. Materiały **9**, 1973, 9–14. MR50 #12810 [86].

D. D. K. SLEATOR

1980. An $O(nm \log n)$ algorithm for maximum network flow, Stanford Univ., Comput. Sci. Dept., Tech. Report STAN-CS-80-831, 1980 [57].

D. D. K. SLEATOR AND R. E. TARJAN

1983. A data structure for dynamic trees, *J. Comput. System Sci.* **26**, 1983, 362–391. Zbl. 509.68058 [57].

S. SMALE

1982. On the average speed of the simplex method of linear programming, preprint, 1982 [262].

1983. The problem of the average speed of the simplex method, *Mathematical Programming, the State of the Art: Bonn, 1982*, Eds.: A. Bachem, M. Grötschel and B. Korte, Springer-Verlag, Berlin, 1983, 530–539 [262].

R. SOLOVAY AND V. STRASSEN

1977. A fast Monte-Carlo test for primality, *SIAM J. Comput.* **6**, 1977, 84–85. MR55#2732 [333].

N. Z. ŠOR

1970. Utilization of the operation of space dilatation in the minimization of convex functions, *Kibernetika* **6**, 1970, 6–12. (Russian). (English translation: *Cybernetics* **6**, 1970, 7–15.) MR45#4836 [xx, 263].

1977. Cut-off method with space extension in convex programming problems, *Kibernetika* **13**, 1977, 94–95. (Russian). (English translation: *Cybernetics* **13**, 1977, 94–96.) MR56#14711 [xx, 263].

E. SPERNER

1928. Ein Satz über Untermengen einer Endlichen Menge, *Math. Z.* **27**, 1928, 544–548. Jbuch. 54.90 [7].

S. STAHL

1979. Fractional edge colorings, *Cahiers Centre Études Rech. Opér.* **21**, 1979, 127–131. MR80i:90091 [288].

E. STEINITZ

1922. Polyhedra und Raumenteilungen, *Encyklopädie der Mathematischen Wissenschaften* III, AB **12**, 1922, 1–139 [xiii].

F. STERBOUL

1979. A characterization of the graphs in which the transversal number equals the matching number, *J. Combin. Theory Ser. B* **27**, 1979, 228–229. MR81c:05080 [223].

V. STRASSEN

1965. The existence of probability measures with given marginals, *Ann. Math. Statist.* **36**, 1965, 423–439. MR31#1693 [76].

D. P. SUMNER

1974. On Tutte's factorization theorem, *Graphs and Combinatorics*, Eds.: R. Bari and F. Harary, Lecture Notes in Math. **406**, Springer- Verlag, New York, 1974, 350–355. MR51#287 [109].

1976. 1-factors and anti-factor sets, *J. London Math. Soc. Ser. (2)* **13**, 1976, 351–359. MR53#13047 [109].

1979. Randomly matchable graphs, *J. Graph Theory* **3**, 1979, 183–186. MR80k:05088 [102].

L. SURÁNYI

1975. On line-critical graphs, *Infinite and Finite Sets (Colloq. Keszthely, Hungary, 1973)*, III, Eds.: A. Hajnal, R. Rado and V. T. Sós, Colloq. Math. Soc. János Bolyai, 10, North-Holland, Amsterdam, 1975, 1411–1444. MR53#5381 [450, 451, 453].

1976. A note on a conjecture of Gallai concerning α-critical graphs, *Combinatorics* **III**, Eds.: A. Hajnal and V. T. Sós, Colloq. Math. Soc. János Bolyai, 18, North-Holland, Amsterdam, 1978, 1065–1074. MR80c: 05088 [455].

R. SWINBORNE-SHELDRAKE, W. C. HERNDON AND I. GUTMAN

1975. Kekulé structures and resonance energies of benzenoid hydrocarbons, *Tetrahedron Lett.*, No. **10**, 1975, 755–758 [350].

P. G. TAIT

1878–80a. On the colouring of maps, *Proc. Roy. Soc. Edinburgh* **10**, 1878–80, 501–503. Jbuch. 12.408 [xii].

1878–80b. Remarks on the previous communication, *Proc. Roy. Soc. Edinburgh* **10**, 1878–80, 729. Jbuch. 12.409 [xii, 285].

1880. Note on a theorem in the geometry of position, *Trans. Roy. Soc. Edinburgh* **29**, 1880, 657–660. Jbuch. 12.409 [xii, 285].

R. E. TARJAN

1972. Depth-first search and linear graph algorithms, *SIAM J. Comput.* **1**, 1972, 146–160. MR46#3313 [55].

G. TARRY

1895. Le problème des labyrinthes, *Nouv. Ann. Math.* **14**, 1895, 187–190. Jbuch. 26.257 [53].

N. TRINAJSTIĆ

1983. *Chemical Graph Theory*, CRC Press, Boca Raton, Florida, 1983 [349].

L. E. TROTTER, JR.

1973. Solution characteristics and algorithms for the vertex packing problem, Cornell Univ., Dept. of O. R., Tech. Report #168, Jan., 1973 [292].

A. C. TUCKER

1977. A note on convergence of the Ford-Fulkerson flow algorithm, *Math. Oper. Res.* **2**, 1977, 143–144. MR57#5042 [58].

W. T. TUTTE

1946. On Hamiltonian circuits, *J. London Math. Soc.* **21**, 1946, 98–101. MR8, 397d [xiii].

1947. The factorization of linear graphs, *J. London Math. Soc.* **22**, 1947, 107–111. MR9, 297d [xvii, 83, 84, 317].

1952. The factors of graphs, *Canad. J. Math.* **4**, 1952, 314–328. MR14, 67e [xvii, 85, 384, 402, 403].

1953. The 1-factors of oriented graphs, *Proc. Amer. Math. Soc.* **4**, 1953, 922–931. MR16, 57f [215, 216].

1954. A short proof of the factor theorem for finite graphs, *Canad. J. Math.* **6**, 1954, 347–352. MR16, 57e [384, 385].

1961. On the problem of decomposing a graph into n connected factors, *J. London Math. Soc.* **36**, 1961, 221–230. MR25#3858 [254].

1965. Lectures on matroids, *J. Res. Nat. Bur. Standards Sect. B* **69B**, 1965, 1–47. MR31#4023 [28].

1966. Introduction to the theory of matroids, Rand Report R-448-PR, 1966 [28].

1971. *Introduction to the Theory of Matroids*, Modern Analytic and Computational Methods in Science and Mathematics, No. 37, American Elsevier, New York, 1971. MR43#1865 [28, 93].

L. G. VALIANT

1979a. The complexity of computing the permanent, *Theoret. Comput. Sci.* **8**, 1979, 189–201. MR80f:86054 [xx, 145, 310].

1979b. The complexity of enumeration and reliability problems, *SIAM J. Comput.* **8**, 1979, 410–421. MR80f:68055 [310].

V. V. VAZIRANI

1984. Maximum matchings without blossoms, Ph.D. Thesis, Univ. California at Berkeley, 1984 [369].

V. G. VIZING

1964. On an estimate of the chromatic class of a p-graph, *Diskret. Analiz.* **3**, 1964, 25–30. (Russian). MR31#4740 [285].

1965. The chromatic class of a multigraph, *Kibernetika* **3**, 1965, 29–39. (Russian). MR32#7333 [285].

M. VOORHOEVE

1979. A lower bound for the permanents of certain (0, 1) matrices, *Nederl. Akad. Wetensch. Indag. Math.* **82**, 1979, 83–86. MR80c:15005 [312].

B. L. VAN DER WAERDEN

1926. Aufgabe 45, *Jber. Deutsch. Math.-Verein.* **35**, 1926, 117 [310].

1931. *Moderne Algebra*, J. Springer, Berlin, 1931. Zbl. 16.339 [27].

J. M. WEINSTEIN

1963. On the number of disjoint edges in a graph, *Canad. J. Math.* **15**, 1963, 106–111. MR27#2973 [115].

1974. Large matchings in graphs, *Canad. J. Math.* **26**, 1974, 1498–1508. MR51#5418 [115].

D. J. A. WELSH

1968. Kruskal's theorem for matroids, *Math. Proc. Cambridge Philos. Soc.* **64**, 1968, 3–4. MR37#2623 [28].

1976. *Matroid Theory*, Academic Press, London, 1976. MR55#148 [23, 29, 30, 93].

1983. Randomised algorithms, *Discrete Appl. Math.* **5**, 1983, 133–145. MR84e: 68046 [333].

D. DE WERRA

1971. Equitable colorations of graphs, *R. A. I. R. O.*, 5, Ser. R- **3**, 1971, 3–8. MR58#27610 [38].

1972. Decomposition of bipartite multigraphs into matchings, *Z. Oper. Res. Ser. A-B* **16**, 1972, A85-A90. MR46#5182 [272].

1975. A few remarks on chromatic scheduling, *Combinatorial Programming: Methods and Applications*, Ed.: B. Roy, D. Reidel, Dordrecht, 1975, 337–342. MR53#5357 [38].

W. WESSEL

1970. Kanten-kritische Graphen mit der Zusammenhangzahl 2, *Manuscripta Math.* **2**, 1970, 309–334. MR42#124 [448].

H. WEYL

1935. Elementare Theorie der konvexen Polyheder, *Comment. Math. Helv.* **7**, 1935, 290–306. Zbl. 11.411. (English translation in: *Contributions to the Theory of Games*, Eds.: H. W. Kuhn and A. W. Tucker, Ann. Math. Studies **24**, Princeton Univ. Press, Princeton, N. J., 1950, 3–18. MR12-352.) [256].

A. T. WHITE

1973. *Graphs, Groups and Surfaces*, North-Holland, Amsterdam, 1973. MR49 #4783 [117].

H. WHITNEY

1932. Non-separable and planar graphs, *Trans. Amer. Math. Soc.* **34**, 1932, 339–362. Zbl. 4.131 [124].

1935. On the abstract properties of linear dependence, *Amer. J. Math.* **57**, 1935, 507–533. Zbl. 12.004 [27, 29].

C. WITZGALL AND C. ZAHN, JR.

1965. Modification of Edmonds' maximum matching algorithm, *J. Res. Nat. Bur. Standards Sect. B* **69B**, 1965, 91–98. MR32#5548 [369].

D. R. WOODALL

1973. The binding number of a graph and its Anderson number, *J. Combin. Theory Ser. B* **15**, 1973, 225–255. MR48#5915 [115, 117].

D. H. YOUNGER

1969. Maximum families of disjoint directed cut sets, *Recent Progress in Combinatorics*, Ed. W. T. Tutte, Academic Press, New York, 1969, 329–333. MR41#3324 [250].

N. ZADEH

1973. A bad network problem for the simplex method and other minimum cost flow algorithms, *Math. Programming* **5**, 1973, 255–266. MR49 #8620 [262].

J. ZAKS

1969. On the 1-factors of n-connected graphs, *Combinatorial Structures and their Applications*, Eds.: R. Guy, H. Hanani, N. Sauer and J. Schönheim, Gordon and Breach, New York, 1970, 481–488. MR41 #8290 [345].

1971. On the 1-factors of n-connected graphs, *J. Combin. Theory Ser. B* **11**, 1971, 169–180. MR43#7365 [345, 346].

Index of Terms

Those page numbers printed in **bold face** denote the location of the *definition* of the corresponding concept. Note that in some cases, a definition appears in several different places for the convenience of the reader.

This Page Intentionally Left Blank

Index of Symbols

We list the symbol, the name of the concept symbolized, if any, and then the number of the page upon which the symbol is first defined.