4

Trees

Contents

4.1 Forests and Trees

Recall that an *acyclic* graph is one that contains no cycles. A connected acyclic graph is called a *tree*. The trees on six vertices are shown in Figure 4.1. According to these definitions, each component of an acyclic graph is a tree. For this reason, acyclic graphs are usually called *forests*.

In order for a graph to be connected, there must be at least one path between any two of its vertices. The following proposition, an immediate consequence of Exercise 2.2.12, says that trees are the connected graphs which just meet this requirement.

Proposition 4.1 *In a tree, any two vertices are connected by exactly one path.* □

Following Diestel (2005), we denote the unique path connecting vertices x and y in a tree T by xTy.

By Theorem 2.1, any graph in which all degrees are at least two contains a cycle. Thus, every tree contains a vertex of degree at most one; moreover, if the tree is nontrivial, it must contain a vertex of degree exactly one. Such a vertex is called a *leaf* of the tree. In fact, the following stronger assertion is true (Exercise 2.1.2).

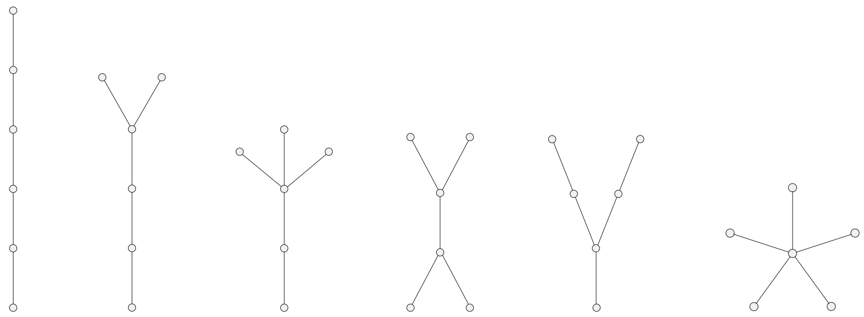

Fig. 4.1. The trees on six vertices

Proposition 4.2 *Every nontrivial tree has at least two leaves.* □

If x is a leaf of a tree T, the subgraph $T-x$ is a tree with $v(T-x) = v(T)-1$ and $e(T-x) = e(T) - 1$. Because the trivial tree has no edges, we have, by induction on the number of vertices, the following relationship between the numbers of edges and vertices of a tree.

Theorem 4.3 *If T is a tree, then $e(T) = v(T) - 1$.* □

Rooted Trees and Branchings

A *rooted tree* $T(x)$ is a tree T with a specified vertex x, called the *root* of T. An orientation of a rooted tree in which every vertex but the root has indegree one is called a *branching*. We refer to a rooted tree or branching with root x as an *x-tree* or *x-branching*, respectively.

There is an evident bijection between x-trees and x-branchings. An x-path thus give rise to a simple example of a branching, a directed x-path. Another example of a branching is shown in Figure 4.2.

Observe that the root of this branching is a source. This is always so, because the sum of the indegrees of a digraph is equal to its number of arcs (Exercise 1.5.2)

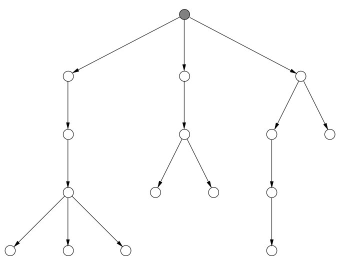

Fig. 4.2. A branching

which, in the case of a branching B, is $v(B) - 1$ by Theorem 4.3. Observe, also, that every vertex of a branching is reachable from its root by a unique directed path. Conversely, in any digraph, reachability from a vertex may be expressed in terms of its branchings. We leave the proof of this fact as an exercise (4.1.6).

Theorem 4.4 *Let x be a vertex of a digraph D, and let X be the set of vertices of D which are reachable from x. Then there is an x-branching in D with vertex set X.* □

Proof Technique: Ordering Vertices

Among the $n!$ linear orderings of the n vertices of a graph, certain ones are especially interesting because they encode particular structural properties. An elementary example is an ordering of the vertices according to their degrees, in decreasing order. More interesting orderings can be obtained by considering the global structure of the graph, rather than just its local structure, as in Exercise 2.2.18. We describe a second example here. Others will be encountered in Chapters 6, 14, and 19, as well as in a number of exercises.

In general, graphs contain copies of many different trees. Indeed, every simple graph with minimum degree k contains a copy of each rooted tree on $k + 1$ vertices, rooted at any given vertex of the graph (Exercise 4.1.9). The analogous question for digraphs (with rooted trees replaced by branchings) is much more difficult. However, in the case of tournaments it can be answered by considering a rather natural ordering of the vertices of the tournament.

A *median order* of a digraph $D = (V, A)$ is a linear order $v_1, v_2, \ldots, v_n$ of its vertex set V such that $|\{(v_i, v_j) : i < j\}|$ (the number of arcs directed from left to right) is as large as possible. In the case of a tournament, such an order can be viewed as a ranking of the players which minimizes the number of upsets (matches won by the lower-ranked player). As we shall see, median orders of tournaments reveal a number of interesting structural properties.

Let us first note two basic properties of median orders of tournaments (Exercise 4.1.10). Let T be a tournament and $v_1, v_2, \ldots, v_n$ a median order of T. Then, for any two indices i, j with $1 \leq i < j \leq n$:

(M1) the interval $v_i, v_{i+1}, \ldots, v_j$ is a median order of the induced subtournament $T[\{v_i, v_{i+1}, \ldots, v_j\}]$,

(M2) vertex v_i dominates at least half of the vertices $v_{i+1}, v_{i+2}, \ldots, v_j$, and vertex v_j is dominated by at least half of the vertices $v_i, v_{i+1}, \ldots, v_{j-1}$.

In particular, each vertex v_i, $1 \leq i < n$, dominates its successor v_{i+1}. The sequence $(v_1, v_2, \ldots, v_n)$ is thus a directed Hamilton path, providing an alternative proof (see Locke (1995)) of Rédei's Theorem (2.3): *every tournament has a directed Hamilton path.*

Ordering Vertices (continued)

By exploiting the properties of median orders, Havet and Thomassé (2000) showed that large tournaments contain all large branchings.

Theorem 4.5 *Any tournament on $2k$ vertices contains a copy of each branching on $k+1$ vertices.*

Proof Let $v_1, v_2, \ldots, v_{2k}$ be a median order of a tournament T on $2k$ vertices, and let B be a branching on $k+1$ vertices. Consider the intervals $v_1, v_2, \ldots, v_i$, $1 \leq i \leq 2k$. We show, by induction on k, that there is a copy of B in T whose vertex set includes at least half the vertices of any such interval.
This is clearly true for $k = 1$. Suppose, then, that $k \geq 2$. Delete a leaf y of B to obtain a branching B' on k vertices, and set $T' := T - \{v_{2k-1}, v_{2k}\}$. By (M1), $v_1, v_2, \ldots, v_{2k-2}$ is a median order of the tournament T', so there is a copy of B' in T' whose vertex set includes at least half the vertices of any interval $v_1, v_2, \ldots, v_i$, $1 \leq i \leq 2k-2$. Let x be the predecessor of y in B. Suppose that x is located at vertex v_i of T'. In T, by (M2), v_i dominates at least half of the vertices $v_{i+1}, v_{i+2}, \ldots, v_{2k}$, thus at least $k - i/2$ of these vertices. On the other hand, B' includes at least $(i-1)/2$ of the vertices $v_1, v_2, \ldots, v_{i-1}$, thus at most $k - (i+1)/2$ of the vertices $v_{i+1}, v_{i+2}, \ldots, v_{2k}$. It follows that, in T, there is an outneighbour v_j of v_i, where $i + 1 \leq j \leq 2k$, which is not in B'. Locating y at v_j, and adding the vertex y and arc (x, y) to B', we now have a copy of B in T. It is readily checked that this copy of B satisfies the required additional property. $\square$

Three further applications of median orders are described in Exercises 4.1.16, 4.1.17, and 4.1.18.

Rooted trees and branchings turn out to be basic tools in the design of efficient algorithms for solving a variety of problems involving reachability, as we shall show in Chapter 6.

Exercises

4.1.1

a) Show that every tree with maximum degree k has at least k leaves.
b) Which such trees have exactly k leaves?

4.1.2 Show that the following three statements are equivalent.

a) G is connected and has $n-1$ edges.
b) G is a forest and has $n-1$ edges.
c) G is a tree.

4.1.3 A *saturated hydrocarbon* is a molecule C_mH_n in which every carbon atom (C) has four bonds, every hydrogen atom (H) has one bond, and no sequence of bonds forms a cycle. Show that, for any positive integer m, the molecule C_mH_n can exist only if $n = 2m + 2$.

4.1.4 Let G be a graph and F a maximal forest of G. Show that $e(F) = v(G) - c(G)$.

4.1.5 Prove Theorem 4.3 by induction on the number of edges of G.

$\star$**4.1.6** Prove Theorem 4.4.

4.1.7 Show that a sequence $(d_1, d_2, \ldots, d_n)$ of positive integers is a degree sequence of a tree if and only if $\sum_{i=1}^{n} d_i = 2(n-1)$.

4.1.8 Centre of a Graph

A *centre* of a graph G is a vertex u such that $\max\{d(u,v) : v \in V\}$ is as small as possible.

a) Let T be a tree on at least three vertices, and let T' be the tree obtained from T by deleting all its leaves. Show that T and T' have the same centres.
b) Deduce that every tree has either exactly one centre or two, adjacent, centres.

4.1.9

a) Show that any simple graph with minimum degree k contains a copy of each rooted tree on $k+1$ vertices, rooted at any given vertex of the graph.
b) Deduce that any simple graph with average degree at least $2(k-1)$, where $k-1$ is a positive integer, contains a copy of each tree on $k+1$ vertices.

(P. Erdős and V.T. Sós (see Erdős (1964)) have conjectured that any simple graph with average degree greater than $k-1$ contains a copy of each tree on $k+1$ vertices; see Appendix A.)

4.1.10 Verify the properties (M1) and (M2) of median orders of tournaments.

4.1.11 Let G be a simple graph with vertex set $V := \{1, 2, \ldots, n\}$.

a) Show that the set of transpositions $S := \{(i,j) : ij \in E\}$ generates all permutations of V if and only if G is connected.
b) Deduce that S is a minimal set of transpositions that generates all permutations of V if and only if G is a tree.

4.1.12 Let $S := \{x_1, x_2, \ldots, x_n\}$ be an n-set, and let $\mathcal{A} := \{A_1, A_2, \ldots, A_n\}$ be a family of n distinct subsets of S. Construct a graph G with vertex set $\mathcal{A}$, two vertices A_i and A_j being joined by an edge if their symmetric difference $A_i \triangle A_j$ is a singleton. Label the edge A_iA_j by this singleton. By studying this labelled graph, prove that there is an element $x_m \in S$ such that the sets $A_1 \cup \{x_m\}, A_2 \cup \{x_m\}, \ldots, A_n \cup \{x_m\}$ are distinct. (J.A. Bondy)

4.1.13 Give an alternative proof of Exercise 4.1.12 by proceeding as follows. Suppose, by way of contradiction, that there is no such element $x_m \in S$, so that, for all $i \in [1, n]$, there exist distinct indices $j(i)$ and $k(i)$ such that $A_{j(i)} \cup \{x_i\} = A_{k(i)}$.

Let $\mathbf{M}$ be the incidence matrix of the hypergraph $(S, \mathcal{A})$ (so that $m_{ij} = 1$ if $x_i \in A_j$ and $m_{ij} = 0$ otherwise), let $\mathbf{c}_i$ denote the column vector with -1 in position $j(i)$, 1 in position $k(i)$, and 0s elsewhere, let $\mathbf{C}$ denote the $n \times n$ matrix whose ith column is $\mathbf{c}_i$, and let $\mathbf{j}$ be the row vector all of whose entries are 1. Show that $\mathbf{M}\,\mathbf{C} = \mathbf{I}$ and $\mathbf{j}\,\mathbf{C} = \mathbf{0}$, and derive a contradiction. (J. Greene)

4.1.14 m identical pizzas are to be shared equally amongst n students.

a) Show how this goal can be achieved by dividing the pizzas into a total of $m + n - d$ pieces, where d is the greatest common divisor of m and n.
b) By considering a suitable bipartite graph, show that no division into a smaller number of pieces will achieve the same objective. (H. Bass)

4.1.15 Rooted trees $T_1(x_1)$ and $T_2(x_2)$ are *isomorphic* if there is an isomorphism from T_1 to T_2 mapping x_1 to x_2. A rooted tree is *uniform* if the degree of a vertex depends only on its distance from the root. Prove that every x-tree on n vertices has exactly n nonisomorphic uniform x-subtrees.

(M.K. Goldberg and I.A. Klipker)

4.1.16 Let $v_1, v_2, \ldots, v_n$ be a median order of an even tournament T. Show that $(v_1, v_2, \ldots, v_n, v_1)$ is a directed Hamilton cycle of T. (S. Thomassé)

4.1.17 A *king* in a tournament is a vertex v from which every vertex is reachable by a directed path of length at most two. Show that every tournament T has a king by proceeding as follows.

Let $v_1, v_2, \ldots, v_n$ be a median order of T.

a) Suppose that v_j dominates v_i, where $i < j$. Show that there is an index k with $i < k < j$ such that v_i dominates v_k and v_k dominates v_j.
b) Deduce that v_1 is a king in T. (F. Havet and S. Thomassé)

4.1.18 A *second outneighbour* of a vertex v in a digraph is a vertex whose distance from v is exactly two. Show that every tournament T has a vertex with at least as many second outneighbours as (first) outneighbours, by proceeding as follows.

Let $v_1, v_2, \ldots, v_n$ be a median order of a tournament T. Colour the outneighbours of v_n red, both v_n and those of its in-neighbours which dominate every red vertex preceding them in the median order black, and the remaining in-neighbours of v_n blue. (Note that every vertex of T is thereby coloured, because T is a tournament.)

a) Show that every blue vertex is a second outneighbour of v_n.
b) Consider the intervals of the median order into which it is subdivided by the black vertices. Using property $(M2)$, show that each such interval includes at least as many blue vertices as red vertices.

c) Deduce that v_n has at least as many second outneighbours as outneighbours.
(F. Havet and S. Thomassé)

(P. D. Seymour has conjectured that every oriented graph has a vertex with at least as many second outneighbours as outneighbours; see Appendix A)

4.1.19

a) Show that the cube of a tree on at least three vertices has a Hamilton cycle.
(M. Sekanina)

b) Find a tree whose square has no Hamilton cycle.

(Fleischner (1974) characterized the graphs whose squares have Hamilton cycles; see also Říha (1991).)

$\star$4.1.20

a) Let T_1 and T_2 be subtrees of a tree T. Show that $T_1 \cap T_2$ and $T_1 \cup T_2$ are subtrees of T if and only if $T_1 \cap T_2 \neq \emptyset$.

b) Let $\mathcal{T}$ be a family of subtrees of a tree T. Deduce, by induction on $|\mathcal{T}|$, that if any two members of $\mathcal{T}$ have a vertex in common, then there is a vertex of T which belongs to all members of $\mathcal{T}$. (In other words, show that the family of subtrees of a tree have the Helly Property (defined in Exercise 1.3.7).)

4.1.21 König's Lemma
Show that every locally-finite infinite tree contains a one-way infinite path.
(D. König)

4.2 Spanning Trees

A *subtree* of a graph is a subgraph which is a tree. If this tree is a spanning subgraph, it is called a *spanning tree* of the graph. Figure 4.3 shows a decomposition of the wheel W_4 into two spanning trees.

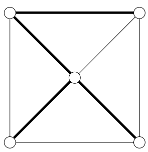

Fig. 4.3. Two spanning trees of the wheel W_4

If a graph G has a spanning tree T, then G is connected because any two vertices of G are connected by a path in T, and hence in G. On the other hand, if G is a connected graph which is not a tree, and e is an edge of a cycle of G, then $G \setminus e$ is a spanning subgraph of G which is also connected because, by Proposition

3.2, e is not a cut edge of G. By repeating this process of deleting edges in cycles until every edge which remains is a cut edge, we obtain a spanning tree of G. Thus we have the following theorem, which provides yet another characterization of connected graphs.

Theorem 4.6 *A graph is connected if and only if it has a spanning tree.* □

It is easy to see that every tree is bipartite. We now use Theorem 4.6 to derive a characterization of bipartite graphs.

Theorem 4.7 *A graph is bipartite if and only if it contains no odd cycle.*

Proof Clearly, a graph is bipartite if and only if each of its components is bipartite, and contains an odd cycle if and only if one of its components contains an odd cycle. Thus, it suffices to prove the theorem for connected graphs.

Let $G[X,Y]$ be a connected bipartite graph. Then the vertices of any path in G belong alternately to X and to Y. Thus, all paths connecting vertices in different parts are of odd length and all paths connecting vertices in the same part are of even length. Because, by definition, each edge of G has one end in X and one end in Y, it follows that every cycle of G is of even length.

Conversely, suppose that G is a connected graph without odd cycles. By Theorem 4.6, G has a spanning tree T. Let x be a vertex of T. By Proposition 4.1, any vertex v of T is connected to x by a unique path in T. Let X denote the set of vertices v for which this path is of even length, and set $Y := V \setminus X$. Then (X,Y) is a bipartition of T. We claim that (X,Y) is also a bipartition of G.

To see this, consider an edge $e = uv$ of $E(G) \setminus E(T)$, and let $P := uTv$ be the unique uv-path in T. By hypothesis, the cycle $P+e$ is even, so P is odd. Therefore the ends of P, and hence the ends of e, belong to distinct parts. It follows that (X,Y) is indeed a bipartition of G. □

According to Theorem 4.7, either a graph is bipartite, or it contains an odd cycle, but not both. An efficient algorithm which finds, in a given graph, either a bipartition or an odd cycle is presented in Chapter 6.

Cayley's Formula

There is a remarkably simple formula for the number of labelled trees on n vertices (or, equivalently, for the number of spanning trees in the complete graph K_n). This formula was discovered by Cayley (1889), who was interested in representing certain hydrocarbons by graphs and, in particular, by trees (see Exercise 4.1.3). A wide variety of proofs have since been found for Cayley's Formula (see Moon (1967)). We present here a particularly elegant one, due to Pitman (1999). It makes use of the concept of a *branching forest*, that is, a digraph each of whose components is a branching.

Theorem 4.8 Cayley's Formula
The number of labelled trees on n vertices is n^{n-2}.

Proof We show, by counting in two ways, that the number of labelled branchings on n vertices is n^{n-1}. Cayley's Formula then follows directly, because each labelled tree on n vertices gives rise to n labelled branchings, one for each choice of the root vertex.

Consider, first, the number of ways in which a labelled branching on n vertices can be built up, one edge at a time, starting with the empty graph on n vertices. In order to end up with a branching, the subgraph constructed at each stage must be a branching forest. Initially, this branching forest has n components, each consisting of an isolated vertex. At each stage, the number of components decreases by one. If there are k components, the number of choices for the new edge (u, v) is $n(k-1)$: any one of the n vertices may play the role of u, whereas v must be the root of one of the $k-1$ components which do not contain u. The total number of ways of constructing a branching on n vertices in this way is thus

$$\prod_{i=1}^{n-1} n(n-i) = n^{n-1}(n-1)!$$

On the other hand, any individual branching on n vertices is constructed exactly $(n-1)!$ times by this procedure, once for each of the orders in which its $n-1$ edges are selected. It follows that the number of labelled branchings on n vertices is n^{n-1}. $\square$

Another proof of Cayley's Formula is outlined in Exercise 4.2.11.

We denote the number of spanning trees in an arbitrary graph G by $t(G)$. Cayley's Formula says that $t(K_n) = n^{n-2}$. There is a simple recursive formula relating the number of spanning trees of a graph G to the numbers of spanning trees in the two graphs $G \setminus e$ and G / e obtained from G by deleting and contracting a link e (Exercise 4.2.1).

Proposition 4.9 *Let G be a graph and e a link of G. Then*

$$t(G) = t(G \setminus e) + t(G / e) \qquad \square$$

Exercises

$\star$**4.2.1** Let G be a connected graph and e a link of G.

a) Describe a one-to-one correspondence between the set of spanning trees of G that contain e and the set of spanning trees of G / e.
b) Deduce Proposition 4.9.

4.2.2

a) Let G be a graph with no loops or cut edges. Show that $t(G) \geq e(G)$.

b) For which such graphs does equality hold?

4.2.3 Let G be a connected graph and let x be a specified vertex of G. A spanning x-tree T of G is called a *distance tree* of G with root x if $d_T(x,v) = d_G(x,v)$ for all $v \in V$.

a) Show that G has a distance tree with root x.
b) Deduce that a connected graph of diameter d has a spanning tree of diameter at most $2d$.

$\star$**4.2.4** Show that the incidence matrix of a graph is totally unimodular (defined in Exercise 1.5.7) if and only if the graph is bipartite.

4.2.5 A *fan* is the join $P \vee K_1$ of a path P and a single vertex. Determine the numbers of spanning trees in:

a) the fan F_n on n vertices, $n \geq 2$,
b) the wheel W_n with n spokes, $n \geq 3$.

4.2.6 Let G be an edge-transitive graph.

a) Show that each edge of G lies in exactly $(n-1)t(G)/m$ spanning trees of G.
b) Deduce that $t(G \setminus e) = (m-n+1)t(G)/m$ and $t(G\,/\,e) = (n-1)t(G)/m$.
c) Deduce that $t(K_n)$ is divisible by n, if $n \geq 3$, and that $t(K_{n,n})$ is divisible by n^2.
d) Without appealing to Cayley's Formula (Theorem 4.8), determine $t(K_4)$, $t(K_5)$, and $t(K_{3,3})$.

4.2.7

a) Let G be a simple graph on n vertices, and let H be the graph obtained from G by replacing each edge of G by k multiple edges. Show that $t(H) = k^{n-1}t(G)$.
b) Let G be a graph on n vertices and m edges, and let H be the graph obtained from G by subdividing each edge of G $k-1$ times. Show that $t(H) = k^{m-n+1}t(G)$.

$\star$**4.2.8** Using Theorem 4.7 and Exercise 3.4.11b, show that a digraph contains a directed odd cycle if and only if some strong component is not bipartite.

$\star$**4.2.9** A branching in a digraph is a *spanning branching* if it includes all vertices of the digraph.

a) Show that a digraph D has a spanning x-branching if and only if $\partial^+(X) \neq \emptyset$ for every proper subset X of V that includes x.
b) Deduce that a digraph is strongly connected if and only if it has a spanning v-branching for every vertex v.

4.2.10 Nonreconstructible Infinite Graphs
Let $T := T_\infty$ denote the infinite tree in which each vertex is of countably infinite degree, and let $F := 2T_\infty$ denote the forest consisting of two disjoint copies of T_∞. Show that (T,F) is a nonreconstructible pair.

4.2.11 Prüfer Code
Let K_n be the labelled complete graph with vertex set $\{1, 2, \ldots, n\}$, where $n \geq 3$. With each spanning tree T of K_n one can associate a unique sequence $(t_1, t_2, \ldots, t_{n-2})$, known as the *Prüfer code* of T, as follows. Let s_1 denote the first vertex (in the the ordered set $(1, 2, \ldots, n)$) which is a leaf of T, and let t_1 be the neighbour of s_1 in T. Now let s_2 denote the first vertex which is a leaf of $T - s_1$, and t_2 the neighbour of s_2 in $T - s_1$. Repeat this operation until t_{n-2} is defined and a tree with just two vertices remains. (If $n \leq 2$, the Prüfer code of T is taken to be the empty sequence.)

a) List all the spanning trees of K_4 and their Prüfer codes.
b) Show that every sequence $(t_1, t_2, \ldots, t_{n-2})$ of integers from the set $\{1, 2, \ldots, n\}$ is the Prüfer code of a unique spanning tree of K_n.
c) Deduce Cayley's Formula (see Theorem 4.8). (H. Prüfer)

4.2.12

a) For a sequence $d_1, d_2, \ldots, d_n$ of n positive integers whose sum is equal to $2n-2$, let $t(n; d_1, d_2, \ldots, d_n)$ denote the number of trees on n vertices $v_1, v_2, \ldots, v_n$ such that $d(v_i) = d_i$, $1 \leq i \leq n$. Show that

$$t(n; d_1, d_2, \ldots, d_n) = \binom{n-2}{d_1 - 1, d_2 - 1, \ldots, d_n - 1}$$

b) Apply the Multinomial Theorem to deduce Cayley's Formula.

4.2.13 By counting the number of branchings whose root lies in the m-set of $K_{m,n}$, show that $t(K_{m,n}) = m^{n-1}n^{m-1}$.

4.2.14 Show that the Petersen graph has 2000 spanning trees.

4.2.15 Let T be a tree with vertex set V, and let $f : V \to V$ be a mapping with no fixed point. For $v \in V$, denote by v^+ the successor of v on the path $vTf(v)$, and by D_f the digraph with vertex set V and arc set $\{(v, v^+) : v \in V\}$.

a) Show that each component of D_f contains a unique directed 2-cycle.
b) The *centroid* of T is the set of all vertices v for which the largest component of $T - v$ has as few vertices as possible. For $v \in V$, let $f(v)$ be a vertex of a largest component of $T - v$, and let (x, y, x) be a directed 2-cycle of D_f. Show that the centroid of T is contained in the set $\{x, y\}$, and hence consists either of one vertex or of two adjacent vertices. (C. Jordan)
c) An *endomorphism* of a simple graph G is a mapping $f : V \to V$ such that, for every $xy \in E$, either $f(x) = f(y)$ or $f(x)f(y) \in E$. Let f be an endomorphism of T, and let (x, y, x) be a directed 2-cycle of D_f.
   i) Show that $f(x) = y$ and $f(y) = x$.
   ii) Deduce that every endomorphism of a tree T fixes either a vertex or an edge of T. (L. Lovász)

d) Let T be a spanning tree of the n-cube Q_n, let $f(v)$ be the antipodal vertex of vertex v in Q_n (that is, the unique vertex whose distance from v is n), and let (x, y, x) be a directed 2-cycle of D_f.
   i) Show that $d_T(f(x), f(y)) \geq 2n - 1$.
   ii) Deduce that every spanning tree of Q_n has a fundamental cycle of length at least $2n$. (R.L. Graham)

4.2.16 Let G be a connected simple graph and T a spanning tree of G. Consider the mapping $\phi : \binom{V}{2} \setminus T \to \binom{T}{2}$ (where T is regarded as a subset of E) defined by $\phi(xy) := \{e, f\}$, where e and f are the first and last edges of the path xTy.

a) Show that the mapping ϕ is a bijection.
b) Deduce that $\binom{n}{2} - |T| = \binom{|T|}{2}$.
c) Deduce Theorem 4.3. (N. Graham, R.C. Entringer, and L. Székely)

4.3 Fundamental Cycles and Bonds

The spanning trees of a connected graph, its even subgraphs, and its edge cuts are intimately related. We describe these relationships here. Recall that, in the context of even subgraphs, when we speak of a *cycle* we typically mean its edge set. Likewise, by a *spanning tree*, we understand in this context the edge set of the tree. Throughout this section, G denotes a connected graph and T a spanning tree of G.

Cotrees

The complement $E \setminus T$ of a spanning tree T is called a *cotree*, and is denoted $\overline{T}$. Consider, for example, the wheel W_4 shown in Figure 4.4a, and the spanning tree $T := \{1, 2, 4, 5\}$ indicated by solid lines. The cotree $\overline{T}$ is simply the set of light edges, namely $\{3, 6, 7, 8\}$.

By Proposition 4.1, for every edge $e := xy$ of a cotree $\overline{T}$ of a graph G, there is a unique xy-path in T connecting its ends, namely $P := xTy$. Thus $T + e$ contains a unique cycle. This cycle is called the *fundamental cycle* of G with respect to T and e. For brevity, we denote it by C_e, the role of the tree T being implicit. Figure 4.4b shows the fundamental cycles of W_4 with respect to the spanning tree $\{1, 2, 4, 5\}$, namely $C_3 = \{1, 2, 3, 4\}$, $C_6 = \{1, 5, 6\}$, $C_7 = \{1, 2, 5, 7\}$, and $C_8 = \{4, 5, 8\}$.

One can draw interesting conclusions about the structure of a graph from the properties of its fundamental cycles with respect to a spanning tree. For example, if all the fundamental cycles are even, then every cycle of the graph is even and hence, by Theorem 4.7, the graph is bipartite. (This is the idea behind the proof of Theorem 4.7.) The following theorem and its corollaries show why fundamental cycles are important.

Theorem 4.10 *Let T be a spanning tree of a connected graph G, and let S be a subset of its cotree $\overline{T}$. Then $C := \triangle\{C_e : e \in S\}$ is an even subgraph of G. Moreover, $C \cap \overline{T} = S$, and C is the only even subgraph of G with this property.*

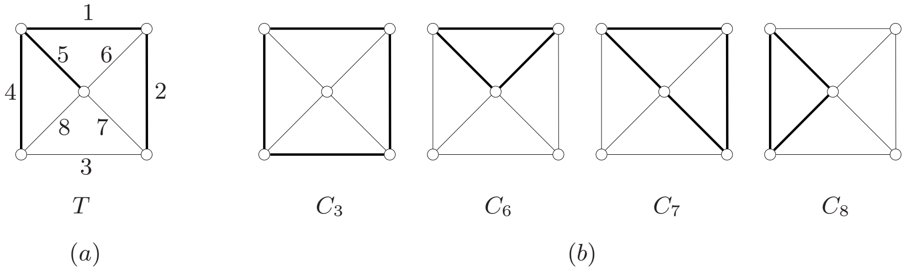

Fig. 4.4. (a) A spanning tree T of the wheel W_4, and (b) the fundamental cycles with respect to T

Proof As each fundamental cycle C_e is an even subgraph, it follows from Corollary 2.16 that C is an even subgraph, too. Furthermore, $C \cap \overline{T} = S$, because each edge of S appears in exactly one member of the family $\{C_e : e \in S\}$.

Let C' be any even subgraph of G such that $C' \cap \overline{T} = S$. Then

$$(C \bigtriangleup C') \cap \overline{T} = (C \cap \overline{T}) \bigtriangleup (C' \cap \overline{T}) = S \bigtriangleup S = \emptyset$$

Therefore the even subgraph $C \bigtriangleup C'$ is contained in T. Because the only even subgraph contained in a tree is the empty even subgraph, we deduce that $C' = C$. □

Corollary 4.11 *Let T be a spanning tree of a connected graph G. Every even subgraph of G can be expressed uniquely as a symmetric difference of fundamental cycles with respect to T.*

Proof Let C be an even subgraph of G and let $S := C \cap \overline{T}$. It follows from Theorem 4.10 that $C = \bigtriangleup\{C_e : e \in S\}$ and that this is the only way of expressing C as a symmetric difference of fundamental cycles with respect to T. □

The next corollary, which follows from Theorem 4.10 by taking $S := \overline{T}$, has several interesting applications (see, for example, Exercises 4.3.9 and 4.3.10).

Corollary 4.12 *Every cotree of a connected graph is contained in a unique even subgraph of the graph.* □

We now discuss the relationship between spanning trees and edge cuts. We show that, for each of the above statements concerning even subgraphs, there is an analogous statement concerning edge cuts. As before, let G be a connected graph and let T be a spanning tree of G. Note that, because T is connected and spanning, every nonempty edge cut of G contains at least one edge of T. Thus the only edge cut contained in the cotree $\overline{T}$ is the empty edge cut (just as the only even subgraph contained in T is the empty even subgraph).

In order to be able to state the cut-analogue of Theorem 4.10, we need the notion of a fundamental bond. Let $e := xy$ be an edge of T. Then $T \setminus e$ has exactly two components, one containing x and the other containing y. Let X denote the vertex set of the component containing x. The bond $B_e := \partial(X)$ is contained in $\overline{T} \cup \{e\}$ and includes e. Moreover, it is the only such bond. For, let B be any bond contained in $\overline{T} \cup \{e\}$ and including e. By Corollary 2.12, $B \triangle B_e$ is an edge cut. Moreover, this edge cut is contained in $\overline{T}$. But, as remarked above, the only such edge cut is the empty edge cut. This shows that $B = B_e$. The bond B_e is called the *fundamental bond* of G with respect to T and e. For instance, the fundamental bonds of the wheel W_4 with respect to the spanning tree $\{1, 2, 4, 5\}$ (indicated in Figure 4.5a) are $B_1 = \{1, 3, 6, 7\}$, $B_2 = \{2, 3, 7\}$, $B_4 = \{3, 4, 8\}$, and $B_5 = \{5, 6, 7, 8\}$ (see Figure 4.5b).

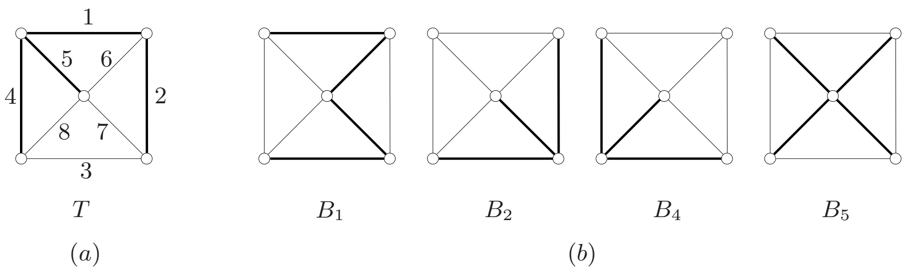

Fig. 4.5. (a) A spanning tree T of the wheel W_4, and (b) the fundamental bonds with respect to T

The proofs of the following theorem and its corollaries are similar to those of Theorem 4.10 and its corollaries, and are left as an exercise (Exercise 4.3.5).

Theorem 4.13 *Let T be a spanning tree of a connected graph G, and let S be a subset of T. Set $B := \triangle\{B_e : e \in S\}$. Then B is an edge cut of G. Moreover $B \cap T = S$, and B is the only edge cut of G with this property.* □

Corollary 4.14 *Let T be a spanning tree of a connected graph G. Every edge cut of G can be expressed uniquely as a symmetric difference of fundamental bonds with respect to T.* □

Corollary 4.15 *Every spanning tree of a connected graph is contained in a unique edge cut of the graph.* □

Corollaries 4.11 and 4.14 imply that the fundamental cycles and fundamental bonds with respect to a spanning tree of a connected graph constitute bases of its cycle and bond spaces, respectively, as defined in Section 2.6 (Exercise 4.3.6). The dimension of the cycle space of a graph is referred to as its *cyclomatic number*.

In this section, we have defined and discussed the properties of fundamental cycles and bonds with respect to spanning trees in connected graphs. All the above theorems are valid for disconnected graphs too, with maximal forests playing the role of spanning trees.

Exercises

4.3.1 Determine the fundamental cycles and fundamental bonds of W_4 with respect to the spanning tree shown in Figure 4.3 (using the edge labelling of Figure 4.4).

4.3.2 TREE EXCHANGE PROPERTY
Let G be a connected graph, let T_1 and T_2 be (the edge sets of) two spanning trees of G, and let $e \in T_1 \setminus T_2$. Show that:

a) there exists $f \in T_2 \setminus T_1$ such that $(T_1 \setminus \{e\}) \cup \{f\}$ is a spanning tree of G,
b) there exists $f \in T_2 \setminus T_1$ such that $(T_2 \setminus \{f\}) \cup \{e\}$ is a spanning tree of G.

(Each of these two facts is referred to as a *Tree Exchange Property*.)

4.3.3 Let G be a connected graph and let S be a set of edges of G. Show that the following statements are equivalent.

a) S is a spanning tree of G.
b) S contains no cycle of G, and is maximal with respect to this property.
c) S meets every bond of G, and is minimal with respect to this property.

4.3.4 Let G be a connected graph and let S be a set of edges of G. Show that the following statements are equivalent.

a) S is a cotree of G.
b) S contains no bond of G, and is maximal with respect to this property.
c) S meets every cycle of G, and is minimal with respect to this property.

4.3.5

a) Prove Theorem 4.13.
b) Deduce Corollaries 4.14 and 4.15.

4.3.6

a) Let T be a spanning tree of a connected graph G. Show that:
    i) the fundamental cycles of G with respect to T form a basis of its cycle space,
    ii) the fundamental bonds of G with respect to T form a basis of its bond space.
b) Determine the dimensions of these two spaces.

(The cycle and bond spaces were defined in Section 2.6.)

4.3.7 Let G be a connected graph, and let $\mathbf{M}$ be its incidence matrix.

a) Show that the columns of $\mathbf{M}$ corresponding to a subset S of E are linearly independent over $GF(2)$ if and only if $G[S]$ is acyclic.
b) Deduce that there is a one-to-one correspondence between the bases of the column space of $\mathbf{M}$ over $GF(2)$ and the spanning trees of G.

(The above statements are special cases of more general results, to be discussed in Section 20.2.)

4.3.8 Algebraic Duals
An *algebraic dual* of a graph G is a graph H for which there is a bijection $\theta : E(G) \to E(H)$ mapping each cycle of G to a bond of H and each bond of G to a cycle of H.

a) Show that:
   i) the octahedron and the cube are algebraic duals,
   ii) $K_{3,3}$ has no algebraic dual.
b) Let G be a connected graph and H an algebraic dual of G, with bijection θ.
   i) Show that T is a spanning tree of G if and only if $\theta(T)$ is a cotree of H.
   ii) Deduce that $t(G) = t(H)$.

4.3.9 Show that any graph which contains a Hamilton cycle has a covering by two even subgraphs.

$\star$**4.3.10** Show that any graph which contains two edge-disjoint spanning trees has:

a) an eulerian spanning subgraph,
b) a covering by two even subgraphs.

4.4 Related Reading

Matroids

One of the characteristic properties of spanning trees of a connected graph is the Tree Exchange Property noted in Exercise 4.3.2a. Because the spanning trees of G correspond to bases of the incidence matrix $\mathbf{M}$ of G (Exercise 4.3.7), the Tree Exchange Property may be seen as a special case of the appropriate exchange property of bases of a vector space. Whitney (1935) observed that many essential properties of spanning trees, such as the ones described in Section 4.3, and more generally of bases of a vector space, may be deduced from that exchange property. Motivated by this observation, he introduced the notion of a matroid.

A *matroid* is an ordered pair $(E, \mathcal{B})$, consisting of a finite set E of *elements* and a nonempty family $\mathcal{B}$ of subsets of E, called *bases*, which satisfy the following *Basis Exchange Property*.

$$\textit{If } B_1, B_2 \in \mathcal{B} \textit{ and } e \in B_1 \setminus B_2 \textit{ then there exists } f \in B_2 \setminus B_1 \textit{ such that } (B_1 \setminus \{e\}) \cup \{f\} \in \mathcal{B}$$

Let $\mathbf{M}$ be a matrix over a field $\mathbb{F}$, let E denote the set of columns of $\mathbf{M}$, and let $\mathcal{B}$ be the family of subsets of E which are bases of the column space of $\mathbf{M}$. Then $(E, \mathcal{B})$ is a matroid. Matroids which arise in this manner are called *linear matroids*. Various linear matroids may be associated with graphs, one example being the matroid on the edge set of a connected graph in which the bases are the edge sets of spanning trees. (In the matroidal context, statements concerning connected graphs extend easily to all graphs, the role of spanning trees being played by maximal forests when the graph is not connected.)

Much of matroid-theoretic terminology is suggested by the two examples mentioned above. For instance, subsets of bases are called *independent sets*, and minimal dependent sets are called *circuits*. In the matroid whose bases are the spanning trees of a connected graph G, the independent sets of the matroid are the forests of G and its circuits are the cycles of G. For this reason, this matroid is called the *cycle matroid* of G, denoted $M(G)$.

The *dual* of a matroid $M = (E, \mathcal{B})$ is the matroid $M^* = (E, \mathcal{B}^*)$, where $\mathcal{B}^* := \{E \setminus B : B \in \mathcal{B}\}$. When M is the linear matroid associated with a matrix $\mathbf{M}$, the bases of M^* are those subsets of E which are bases of the orthogonal complement of the column space of $\mathbf{M}$. When M is the cycle matroid of a connected graph G, the bases of M^* are the cotrees of G, and its circuits are the bonds of G. For this reason, the dual of the cycle matroid of a graph G is called the *bond matroid* of G, denoted $M^*(G)$. Many manifestations of this cycle–bond duality crop up throughout the book. The reader is referred to Oxley (1992) for a thorough account of the theory of matroids.