2 Matching Covering and Packing

Suppose we are given a graph and are asked to find in it as many independent edges as possible. How should we go about this? Will we be able to pair up all its vertices in this way? If not, how can we be sure that this is indeed impossible? Somewhat surprisingly, this basic problem does not only lie at the heart of numerous applications, it also gives rise to some rather interesting graph theory.

A set M of independent edges in a graph $G = (V, E)$ is called a *matching*. M is a matching *of* $U \subseteq V$ if every vertex in U is incident with an edge in M. The vertices in U are then called *matched* (by M); vertices not incident with any edge of M are *unmatched*. *matching* *matched*

A k-regular spanning subgraph is called a *k-factor*. Thus, a subgraph $H \subseteq G$ is a 1-factor of G if and only if $E(H)$ is a matching of V. The problem of how to characterize the graphs that have a 1-factor, i.e. a matching of their entire vertex set, will be our main theme in the first two sections of this chapter. *factor*

A generalization of the matching problem is to find in a given graph G as many disjoint subgraphs as possible that are each isomorphic to an element of a given class $\mathcal{H}$ of graphs. This is known as the *packing* problem. It is related to the *covering* problem, which asks how few vertices of G suffice to meet all its subgraphs isomorphic to a graph in $\mathcal{H}$. Clearly, we need at least as many vertices for such a cover as the maximum number k of graphs from $\mathcal{H}$ that we can pack disjointly into G. If there is no cover by just k vertices, perhaps there is always a cover by at most $f(k)$ vertices, where $f(k)$ may depend on $\mathcal{H}$ but not *packing* *covering* on G? In Section 2.3 we shall prove that when $\mathcal{H}$ is the class of cycles, then there is such a function f.

In Section 2.4 we consider packing and covering in terms of edges: we ask how many edge-disjoint spanning trees we can find in a given graph, and how few trees in it will cover all its edges. In Section 2.5 we prove a path cover theorem for directed graphs, which implies the well-known duality theorem of Dilworth for partial orders.

2.1 Matching in bipartite graphs

$G = (V, E)$
A, B
a, b *etc.*

For this whole section, we let $G = (V, E)$ be a fixed bipartite graph with bipartition $\{A, B\}$. Vertices denoted as a, a' etc. will be assumed to lie in A, vertices denoted as b etc. will lie in B.

How can we find a matching in G with as many edges as possible? Let us start by considering an arbitrary matching M in G. A path in G which starts in A at an unmatched vertex and then contains, alternately, edges from $E \setminus M$ and from M, is an *alternating path* with respect to M. Note that the path is allowed to be trivial, i.e. to consist of its starting vertex only. An alternating path P that ends in an unmatched vertex of B is called an *augmenting path* (Fig. 2.1.1), because we can use it to turn M into a larger matching: the symmetric difference of M with $E(P)$ is again a matching (consider the edges at a given vertex), and the set of matched vertices is increased by two, the ends of P.

alternating path

augmenting path

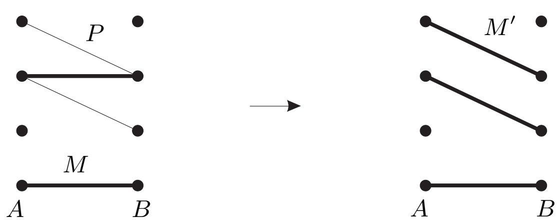

Fig. 2.1.1. Augmenting the matching M by the alternating path P

Alternating paths play an important role in the practical search for large matchings. In fact, if we start with any matching and keep applying augmenting paths until no further such improvement is possible, the matching obtained will always be an optimal one, a matching with the largest possible number of edges (Exercise 1). The algorithmic problem of finding such matchings thus reduces to that of finding augmenting paths – which is an interesting and accessible algorithmic problem.

Our first theorem characterizes the maximal cardinality of a matching in G by a kind of duality condition. Let us call a set $U \subseteq V$ a *(vertex) cover* of E if every edge of G is incident with a vertex in U.

cover

Theorem 2.1.1. (König 1931)
The maximum cardinality of a matching in G is equal to the minimum cardinality of a vertex cover of its edges.

Proof. Let M be a matching in G of maximum cardinality. From every edge in M let us choose one of its ends: its end in B if some alternating path ends in that vertex, and its end in A otherwise (Fig. 2.1.2). We shall prove that the set U of these $|M|$ vertices covers E; since any vertex cover of E must cover M, there can be none with fewer than $|M|$ vertices, and so the theorem will follow. M U

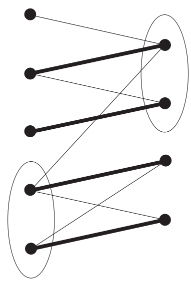

Fig. 2.1.2. The vertex cover U

Note that if an alternating path P ends in a vertex $b \in B$, then $b \in U$: as M is a largest matching, P is not an augmenting path, so b is matched to some $a \in A$ and was put in U when we considered the edge $ab \in M$ while constructing U.

To show that U covers E, let an edge $ab \in E$ be given. If $a \in U$ we are done, so assume that $a \notin U$. To prove $b \in U$, it suffices to show that some alternating path ends in b. If a is unmatched, then ab is such a path. If not, we have $ab' \in M$ for some $b' \in B$. Since $a \notin U$, there exists an alternating path P ending in b'. Depending on whether or not $b \in P$, either Pb or $Pb'ab$ is an alternating path ending in b. □

Let us return to our main problem, the search for some necessary and sufficient conditions for the existence of a 1-factor. In our present case of a bipartite graph, we may as well ask more generally when G contains a matching of A; this will define a 1-factor of G if $|A| = |B|$, a condition that has to hold anyhow if G is to have a 1-factor.

A condition clearly necessary for the existence of a matching of A is that every subset of A has enough neighbours in B, i.e. that

$$|N(S)| \geqslant |S| \qquad \text{for all } S \subseteq A.$$

marriage condition

The following *marriage theorem* says that this obvious necessary condition is in fact sufficient:

[2.2.3] **Theorem 2.1.2.** (Hall 1935)
G contains a matching of A if and only if $|N(S)| \geqslant |S|$ for all $S \subseteq A$.

We give three proofs, of rather different character.[1] In each proof we assume that G satisfies the marriage condition and find a matching of A.

M *First proof.* We show that for every matching M of G that leaves a
a vertex $a \in A$ unmatched there is an augmenting path with respect to M.

Let A' and B' be the sets of vertices in A and B that can be reached by an alternating path from a. Any such path ending at an unmatched $b' \in B'$ is augmenting, so we may assume that all $b' \in B'$ are matched. Their M-neighbours clearly lie in A', but $a \in A'$ is not among these. Hence by the marriage condition, A' also sends an edge $a'b$ to a vertex $b \notin B'$. Appending this edge to an alternating a–a' path yields an alternating a–b path. This places b in B', contradicting its choice. □

Second proof. We apply induction on $|A|$. For $|A| = 1$ the assertion is true. Now let $|A| \geqslant 2$, and assume that the marriage condition is sufficient for the existence of a matching of A when $|A|$ is smaller.

If $|N(S)| \geqslant |S|+1$ for every non-empty set $S \subsetneq A$, we pick an edge $ab \in G$ and consider the graph $G' := G - \{a, b\}$ obtained by deleting its ends. Then every non-empty set $S \subseteq A \setminus \{a\}$ satisfies

$$|N_{G'}(S)| \geqslant |N_G(S)| - 1 \geqslant |S|\,,$$

so by the induction hypothesis G' contains a matching of $A \setminus \{a\}$. Together with the edge ab, this yields a matching of A in G.

Suppose now that A has a non-empty proper subset A' with $|B'| = |A'|$ for $B' := N(A')$. By the induction hypothesis, $G' := G[A' \cup B']$ contains a matching of A'. But $G - G'$ satisfies the marriage condition too: for any set $S \subseteq A \setminus A'$ with $|N_{G-G'}(S)| < |S|$ we would have $|N_G(S \cup A')| < |S \cup A'|$, contrary to our assumption. Again by induction, $G - G'$ contains a matching of $A \setminus A'$. Putting the two matchings together, we obtain a matching of A in G. □

H For our last proof, let H be an edge-minimal subgraph of G that satisfies the marriage condition and contains A. Note that $d_H(a) \geqslant 1$ for every $a \in A$, by the marriage condition with $S = \{a\}$.

Third proof. We show that $d_H(a) = 1$ for every $a \in A$. The edges of H then form a matching of A, since by the marriage condition no two such edges can share a vertex in B.

Suppose a has distinct neighbours b_1, b_2 in H. By definition of H, the graphs $H - ab_1$ and $H - ab_2$ violate the marriage condition. So for

[1] The theorem can also be derived easily from König's theorem; see Exercise 5.

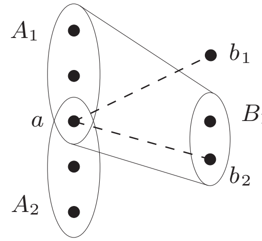

Fig. 2.1.3. B_1 contains b_2 but not b_1

$i = 1, 2$ there is a set $A_i \subseteq A$ containing a such that $|A_i| > |B_i|$ for $B_i := N_{H-ab_i}(A_i)$ (Fig. 2.1.3). Since $b_1 \in B_2$ and $b_2 \in B_1$,

$$\begin{aligned}
|N_H(A_1 \cap A_2 \setminus \{a\})| &\leqslant |B_1 \cap B_2| \\
&= |B_1| + |B_2| - |B_1 \cup B_2| \\
&= |B_1| + |B_2| - \big|N_H(A_1 \cup A_2)\big| \\
&\leqslant |A_1| - 1 + |A_2| - 1 - |A_1 \cup A_2| \\
&= |A_1 \cap A_2| - 2 \\
&= |A_1 \cap A_2 \setminus \{a\}| - 1\,.
\end{aligned}$$

Hence H violates the marriage condition, contrary to assumption. □

This last proof has a pretty 'dual', which begins by showing that $d_H(b) \leqslant 1$ for every $b \in B$. See Exercise 6 and its hint for details.

Corollary 2.1.3. *Every k-regular ($k \geqslant 1$) bipartite graph has a 1-factor.*

Proof. If G is k-regular, then clearly $|A| = |B|$; it thus suffices to show by Theorem 2.1.2 that G contains a matching of A. Now every set $S \subseteq A$ is joined to $N(S)$ by a total of $k\,|S|$ edges, and these are among the $k\,|N(S)|$ edges of G incident with $N(S)$. Therefore $k\,|S| \leqslant k\,|N(S)|$, so G does indeed satisfy the marriage condition. □

In some real-life applications, matchings are not chosen on the basis of global criteria for the entire graph but evolve as the result of independent decisions made locally by the participating vertices. A typical situation is that vertices are not indifferent to which of their incident edges are picked to match them, but prefer some to others. Then if M is a matching and $e = ab$ is an edge not in M such that both a and b prefer e to their current matching edge (if they are matched), then a and b may agree to change M locally by including e and discarding their earlier matching edges. The matching M, although perhaps of maximum size, would thus be *unstable*.

preferences *stable matching*

More formally, call a family $(\leqslant_v)_{v \in V}$ of linear orderings $\leqslant_v$ on $E(v)$ a *set of preferences* for G. Then call a matching M in G *stable* if for every edge $e \in E \setminus M$ there exists an edge $f \in M$ such that e and f have a common vertex v with $e <_v f$. The following result is sometimes called the *stable marriage theorem*; see Exercises 16 and 17 for a discussion of alternative proofs.

[5.4.4] **Theorem 2.1.4.** (Gale & Shapley 1962)
For every set of preferences, G has a stable matching.

Proof. Call a matching M in G *better* than a matching $M' \neq M$ if M makes the vertices in B happier than M' does, that is, if every vertex b in an edge $f' \in M'$ is incident also with some $f \in M$ such that $f' \leqslant_b f$. We shall construct a sequence of better and better matchings. Since these can increase the happiness of a fixed vertex b at most $d(b)$ times, this process will terminate.

Given a matching M, call a vertex $a \in A$ *acceptable to* $b \in B$ if $e = ab \in E \setminus M$ and any edge $f \in M$ at b satisfies $f <_b e$. Call $a \in A$ *happy with* M if a is unmatched or its matching edge $f \in M$ satisfies $f >_a e$ for all edges $e = ab$ such that a is acceptable to b.

Starting with the empty matching, let us construct a sequence of matchings that keep all the vertices in A happy. Given such a matching M, consider a vertex $a \in A$ that is unmatched but acceptable to some $b \in B$. (If no such a exists, terminate the sequence.) Add to M the $\leqslant_a$-maximal edge ab such that a is acceptable to b, and discard from M any other edge at b.

Clearly, each matching in our sequence is better than the previous and keeps the vertices in A happy (which they initially are, when $M = \emptyset$). So the sequence continues until it terminates with a matching M such that no unmatched vertex in A is acceptable to any of its neighbours in B. As every matched vertex in A is happy with M, this matching is stable. $\square$

Despite its seemingly narrow formulation, the marriage theorem counts among the most frequently applied graph theorems, both outside graph theory and within. Often, however, recasting a problem in the setting of bipartite matching requires some clever adaptation. As a simple example, we now use the marriage theorem to derive one of the earliest results of graph theory, a result whose original proof is not all that simple, and certainly not short:

Corollary 2.1.5. (Petersen 1891)
Every regular graph of positive even degree has a 2-factor.

Proof. Let G be any $2k$-regular graph ($k \geqslant 1$), without loss of generality connected. By Theorem 1.8.1, G contains an Euler tour $v_0 e_0 \dots e_{\ell-1} v_\ell$, with $v_\ell = v_0$. We replace every vertex v by a pair (v^-, v^+), and every edge $e_i = v_i v_{i+1}$ by the edge $v_i^+ v_{i+1}^-$ (Fig. 2.1.4). The resulting bipartite graph G' is k-regular, so by Corollary 2.1.3 it has a 1-factor. Collapsing every vertex pair (v^-, v^+) back into a single vertex v, we turn this 1-factor of G' into a 2-factor of G. □ (1.8.1)

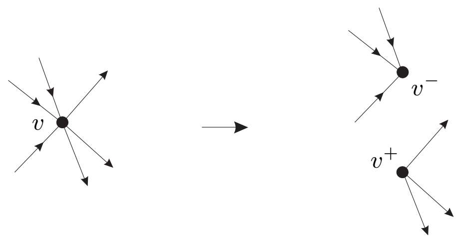

Fig. 2.1.4. Splitting vertices in the proof of Corollary 2.1.5

2.2 Matching in general graphs

Given a graph G, let us denote by $\mathcal{C}_G$ the set of its components, and by $q(G)$ the number of its *odd components*, those of odd order. If G has a 1-factor, then clearly $\mathcal{C}_G$ $q(G)$

Tutte's condition

$$q(G - S) \leqslant |S| \qquad \text{for all } S \subseteq V(G),$$

since every odd component of $G - S$ will send a factor edge to S.

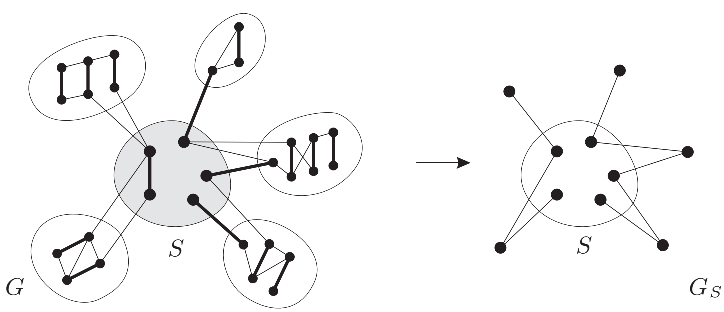

Fig. 2.2.1. Tutte's condition $q(G - S) \leqslant |S|$ for $q = 3$, and the contracted graph G_S from Theorem 2.2.3.

Again, this obvious necessary condition for the existence of a 1-factor is also sufficient:

Theorem 2.2.1. (Tutte 1947)
A graph G has a 1-factor if and only if $q(G-S) \leqslant |S|$ for all $S \subseteq V(G)$.

V, E bad set

Proof. Let $G = (V, E)$ be a graph without a 1-factor. Our task is to find a *bad set* $S \subseteq V$, one that violates Tutte's condition.

We may assume that G is edge-maximal without a 1-factor. Indeed, if G' is obtained from G by adding edges and $S \subseteq V$ is bad for G', then S is also bad for G: any odd component of $G' - S$ is the union of components of $G - S$, and one of these must again be odd.

What does G look like? Clearly, *if* G contains a bad set S then, by its edge-maximality and the trivial forward implication of the theorem,

$$\text{all the components of } G-S \text{ are complete and every vertex } s \in S \text{ is adjacent to all the vertices of } G-s. \qquad (*)$$

But also conversely, if a set $S \subseteq V$ satisfies $(*)$ then either S or the empty set must be bad: if S is not bad we can join the odd components of $G - S$ disjointly to S and pair up all the remaining vertices – unless $|G|$ is odd, in which case $\emptyset$ is bad.

S a, b, c d M_1, M_2

So it suffices to prove that G has a set S of vertices satisfying $(*)$. Let S be the set of vertices that are adjacent to every other vertex. If this set S does not satisfy $(*)$, then some component of $G - S$ has non-adjacent vertices a, a'. Let a, b, c be the first three vertices on a shortest a–a' path in this component; then $ab, bc \in E$ but $ac \notin E$. Since $b \notin S$, there is a vertex $d \in V$ such that $bd \notin E$. By the maximality of G, there is a matching M_1 of V in $G + ac$, and a matching M_2 of V in $G + bd$.

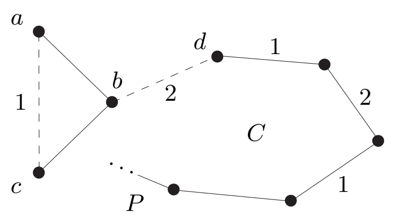

Fig. 2.2.2. Deriving a contradiction if S does not satisfy $(*)$

v

Let $P = d \dots v$ be a maximal path in G starting at d with an edge from M_1 and containing alternately edges from M_1 and M_2 (Fig. 2.2.2). If the last edge of P lies in M_1, then $v = b$, since otherwise we could continue P. Let us then set $C := P + bd$. If the last edge of P lies in M_2, then by the maximality of P the M_1-edge at v must be ac, so $v \in \{a, c\}$; then let $C := dPvbd$. This is again a cycle, since b cannot be an inner vertex of P: those are incident with M_2-edges from G, which b is not.

In every case C is an even cycle with every other edge in M_2, and whose only edge not in E is bd. Replacing in M_2 its edges on C with the edges of $C - M_2$, we obtain a matching of V in E, a contradiction. $\square$

Corollary 2.2.2. (Petersen 1891)
Every bridgeless cubic graph has a 1-factor.

Proof. We show that any bridgeless cubic graph G satisfies Tutte's condition. Let $S \subseteq V(G)$ be given, and consider an odd component C of $G - S$. Since G is cubic, the degrees (in G) of the vertices in C sum to an odd number, but only an even part of this sum arises from edges of C. So G has an odd number of S–C edges, and therefore has at least 3 such edges (since G has no bridge). The total number of edges between S and $G - S$ thus is at least $3q(G - S)$. But it is also at most $3|S|$, because G is cubic. Hence $q(G - S) \leqslant |S|$, as required. □

In order to shed a little more light on the techniques used in matching theory, we now give a second proof of Tutte's theorem. In fact, we shall prove a slightly stronger result, a result that places a structure interesting from the matching point of view on an *arbitrary* graph. If the graph happens to satisfy the condition of Tutte's theorem, this structure will at once yield a 1-factor.

factor-critical

A non-empty graph $G = (V, E)$ is called *factor-critical* if G has no 1-factor but for every vertex $v \in G$ the graph $G - v$ has a 1-factor. We call a vertex set $S \subseteq V$ *matchable to* $\mathcal{C}_{G-S}$ if the (bipartite[2]) graph G_S, which arises from G by contracting the components $C \in \mathcal{C}_{G-S}$ to single vertices and deleting all the edges inside S, contains a matching of S. (Formally, G_S is the graph with vertex set $S \cup \mathcal{C}_{G-S}$ and edge set $\{\, sC \mid \exists c \in C : sc \in E \,\}$; see Fig. 2.2.1.)

matchable

G_S

Theorem 2.2.3. (Gallai 1964; Edmonds 1965)
Every graph G has a set S of vertices with the following two properties:

(i) *S is matchable to $\mathcal{C}_{G-S}$;*

(ii) *Every component of $G - S$ is factor-critical.*

Given any such set S, the graph G contains a 1-factor if and only if $|S| = |\mathcal{C}_{G-S}|$.

For any given G, the assertion of Tutte's theorem follows easily from this result. Indeed, by (i) and (ii) we have $|S| \leqslant |\mathcal{C}_{G-S}| = q(G - S)$ (since factor-critical graphs have odd order); thus Tutte's condition of $q(G - S) \leqslant |S|$ implies $|S| = |\mathcal{C}_{G-S}|$, and the existence of a 1-factor follows from the last statement of Theorem 2.2.3.

Proof of Theorem 2.2.3. Note first that the last assertion of the theorem follows at once from the assertions (i) and (ii): if G has a 1-factor, we have $q(G - S) \leqslant |S|$ and hence $|S| = |\mathcal{C}_{G-S}|$ as above; conversely if (2.1.2)

[2] except for the – permitted – case that S or $\mathcal{C}_{G-S}$ is empty

$|S| = |\mathcal{C}_{G-S}|$, then the existence of a 1-factor follows straight from (i) and (ii).

We now prove the existence of a set S satisfying (i) and (ii), by induction on $|G|$. For $|G| = 0$ we may take $S = \emptyset$. Now let G be given with $|G| > 0$, and assume the assertion holds for graphs with fewer vertices.

Consider the sets $T \subseteq V(G)$ for which Tutte's condition fails worst, i.e. for which

d $$d(T) := d_G(T) := q(G-T) - |T|$$

S is maximum, and let S be a largest such set T. Note that $d(S) \geqslant d(\emptyset) \geqslant 0$.

$\mathcal{C}$ We first show that every component $C \in \mathcal{C}_{G-S} =: \mathcal{C}$ is odd. If $|C|$ is even, pick a vertex $c \in C$, and consider $T := S \cup \{c\}$. As $C - c$ has odd order it has at least one odd component, which is also a component of $G - T$. Therefore

$$q(G-T) \geqslant q(G-S) + 1 \quad \text{while} \quad |T| = |S| + 1\,,$$

so $d(T) \geqslant d(S)$ contradicting the choice of S.

Next we prove the assertion (ii), that every $C \in \mathcal{C}$ is factor-critical. Suppose there exist $C \in \mathcal{C}$ and $c \in C$ such that $C' := C - c$ has no 1-factor. By the induction hypothesis (and the fact that, as shown earlier, for fixed G our theorem implies Tutte's theorem) there exists a set $S' \subseteq V(C')$ with

$$q(C' - S') > |S'|\,.$$

Since $|C|$ is odd and hence $|C'|$ is even, the numbers $q(C' - S')$ and $|S'|$ are either both even or both odd, so they cannot differ by exactly 1. We may therefore sharpen the above inequality to

$$q(C' - S') \geqslant |S'| + 2\,,$$

giving $d_{C'}(S') \geqslant 2$. Then for $T := S \cup \{c\} \cup S'$ we have

$$d(T) \geqslant d(S) - 1 - 1 + d_{C'}(S') \geqslant d(S)\,,$$

where the first '-1' comes from the loss of C as an odd component and the second comes from including c in the set T. As before, this contradicts the choice of S.

It remains to show that S is matchable to $\mathcal{C}_{G-S}$. If not, then by the marriage theorem there exists a set $S' \subseteq S$ that sends edges to fewer than $|S'|$ components in $\mathcal{C}$. Since the other components in $\mathcal{C}$ are also components of $G - (S \smallsetminus S')$, the set $T = S \smallsetminus S'$ satisfies $d(T) > d(S)$, contrary to the choice of S. $\square$

Let us consider once more the set S from Theorem 2.2.3, together with any matching M in $G = (V, E)$. As before, we write $\mathcal{C} := \mathcal{C}_{G-S}$. Let us denote by k_S the number of edges in M with at least one end in S, and by $k_\mathcal{C}$ the number of edges in M with both ends in $G - S$. Since each $C \in \mathcal{C}$ is odd, at least one of its vertices is not incident with an edge of the second type. Therefore every matching M satisfies S $V, \mathcal{C}$ $k_S, k_\mathcal{C}$

$$k_S \leqslant |S| \quad \text{and} \quad k_\mathcal{C} \leqslant \tfrac{1}{2}\Big(|V| - |S| - |\mathcal{C}|\Big). \tag{1}$$

Moreover, G contains a matching M_0 with equality in both cases: first choose $|S|$ edges between S and $\bigcup \mathcal{C}$ according to (i), and then use (ii) to find a suitable set of $\frac{1}{2}\big(|C| - 1\big)$ edges in every component $C \in \mathcal{C}$. This matching M_0 thus has exactly M_0

$$|M_0| = |S| + \tfrac{1}{2}\Big(|V| - |S| - |\mathcal{C}|\Big) \tag{2}$$

edges.

Now (1) and (2) together imply that *every* matching M of maximum cardinality satisfies both parts of (1) with equality: by $|M| \geqslant |M_0|$ and (2), M has at least $|S| + \frac{1}{2}\big(|V| - |S| - |\mathcal{C}|\big)$ edges, which implies by (1) that neither of the inequalities in (1) can be strict. But equality in (1), in turn, implies that M has the structure described above: by $k_S = |S|$, every vertex $s \in S$ is the end of an edge $st \in M$ with $t \in G - S$, and by $k_\mathcal{C} = \frac{1}{2}\big(|V| - |S| - |\mathcal{C}|\big)$ exactly $\frac{1}{2}\big(|C| - 1\big)$ edges of M lie in C, for every $C \in \mathcal{C}$. Finally, since these latter edges miss only one vertex in each C, the ends t of the edges st above lie in different components C for different s.

The seemingly technical Theorem 2.2.3 thus hides a wealth of structural information: it contains the essence of a detailed description of all maximum-cardinality matchings in all graphs. A reference to the full statement of this result, the *Gallai-Edmonds structure theorem*, is given in the notes at the end of this chapter.

2.3 The Erdős-Pósa theorem

Much of the charm of König's and Hall's theorems in Section 2.1 lies in the fact that they guarantee the existence of the desired matching as soon as some obvious obstruction does not occur. In König's theorem, we can find k independent edges in our graph unless we can cover all its edges by fewer than k vertices (in which case it is obviously impossible).

More generally, if G is an arbitrary graph, not necessarily bipartite, and $\mathcal{H}$ is any class of graphs, we might compare the largest number k of graphs from $\mathcal{H}$ (not necessarily distinct) that we can pack disjointly into G with the smallest number s of vertices of G that will cover all

Erdős-Pósa property its subgraphs in $\mathcal{H}$. If s can be bounded by a function of k, i.e. independently of G, we say that $\mathcal{H}$ has the *Erdős-Pósa property*. (Thus, formally, $\mathcal{H}$ has this property if there exists an $\mathbb{N} \to \mathbb{N}$ function $k \mapsto f(k)$ such that, for every k and G, either G contains k disjoint subgraphs each isomorphic to a graph in $\mathcal{H}$, or there is a set $U \subseteq V(G)$ of at most $f(k)$ vertices such that $G - U$ has no subgraph in $\mathcal{H}$.)

Our aim in this section is to prove the theorem of Erdős and Pósa that the class of all cycles has this property: we shall find a function f (about $4k \log k$) such that every graph contains either k disjoint cycles or a set of at most $f(k)$ vertices covering all its cycles.

We begin by proving a stronger assertion for cubic graphs. For $k \in \mathbb{N}$, put

r_k, s_k

$$s_k := \begin{cases} 4kr_k & \text{if } k \geqslant 2 \\ 1 & \text{if } k \leqslant 1 \end{cases} \quad \text{where} \quad r_k := \log k + \log\log k + 4\,.$$

Lemma 2.3.1. *Let $k \in \mathbb{N}$, and let H be a cubic multigraph. If $|H| \geqslant s_k$, then H contains k disjoint cycles.*

(1.3.5) *Proof.* We apply induction on k. For $k \leqslant 1$ the assertion is trivial, so let $k \geqslant 2$ be given for the induction step. Let C be a shortest cycle in H.

We first show that $H - C$ contains a subdivision of a cubic multigraph H' with $|H'| \geqslant |H| - 2|C|$. Let m be the number of edges between C and $H - C$. Since H is cubic and $d(C) = 2$, we have $m \leqslant |C|$. We now consider bipartitions $\{V_1, V_2\}$ of $V(H)$, beginning with $V_1 := V(C)$ and allowing $V_2 = \emptyset$. If $H[V_2]$ has a vertex of degree at most 1 we move this vertex to V_1, obtaining a new partition $\{V_1, V_2\}$ crossed by fewer edges. Suppose we can perform a sequence of n such moves, but no more. (Our assumptions imply $n \leqslant 3$, but we do not formally need this.) Then the resulting partition $\{V_1, V_2\}$ is crossed by at most $m - n$ edges. And $H[V_2]$ has at most $m - n$ vertices of degree less than 3, because each of these is incident with a crossing edge. These vertices have degree exactly 2 in $H[V_2]$, since we could not move them to V_1. Let H' be the cubic multigraph obtained from $H[V_2]$ by suppressing these vertices. Then

m n

$$|H'| \geqslant |H| - |C| - n - (m - n) \geqslant |H| - 2|C|\,,$$

as desired.

To complete the proof, it suffices to show that $|H'| \geqslant s_{k-1}$. Since $|C| \leqslant 2 \log |H|$ by Corollary 1.3.5 (or by $|H| \geqslant s_k$, if $|C| = g(H) \leqslant 2$), and $|H| \geqslant s_k \geqslant 6$, we have

$$|H'| \geqslant |H| - 2|C| \geqslant |H| - 4 \log |H| \geqslant s_k - 4 \log s_k\,.$$

(In the last inequality we use that the function $x \mapsto x - 4 \log x$ increases for $x \geqslant 6$.)

It thus remains to show that $s_k - 4\log s_k \geqslant s_{k-1}$. For $k = 2$ this is clear, so we assume that $k \geqslant 3$. Then $r_k \leqslant 4\log k$ (which is obvious for $k \geqslant 4$, while the case of $k = 3$ has to be calculated), and hence

$$\begin{aligned} s_k - 4\log s_k &= 4(k-1)r_k + 4\log k + 4\log\log k + 16 \\ &\qquad\qquad - \big(8 + 4\log k + 4\log r_k\big) \\ &\geqslant s_{k-1} + 4\log\log k + 8 - 4\log(4\log k) \\ &= s_{k-1}\,. \end{aligned}$$ □

Theorem 2.3.2. (Erdős & Pósa 1965)
There is a function $f\colon \mathbb{N} \to \mathbb{N}$ such that, given any $k \in \mathbb{N}$, every graph contains either k disjoint cycles or a set of at most $f(k)$ vertices meeting all its cycles.

Proof. We show the result for $f(k) := \lfloor s_k + k - 1 \rfloor$. Let k be given, and let G be any graph. We may assume that G contains a cycle, and so it has a maximal subgraph H in which every vertex has degree 2 or 3. Let U be its set of degree 3 vertices. U

Let $\mathcal{C}$ be the set of all cycles in G that avoid U and meet H in exactly one vertex. Let $Z \subseteq V(H) \setminus U$ be the set of those vertices. For each $z \in Z$ pick a cycle $C_z \in \mathcal{C}$ that meets H in z, and put $\mathcal{C}' := \{\, C_z \mid z \in Z \,\}$. By the maximality of H, the cycles in $\mathcal{C}'$ are disjoint. Z

Let $\mathcal{D}$ be the set of the 2-regular components of H that avoid Z. Then $\mathcal{C}' \cup \mathcal{D}$ is another set of disjoint cycles. If $|\mathcal{C}' \cup \mathcal{D}| \geqslant k$, we are done. Otherwise we can add to Z one vertex from each cycle in $\mathcal{D}$ to obtain a set X of at most $k-1$ vertices that meets all the cycles in $\mathcal{C}$ and all the 2-regular components of H. Now consider any cycle of G that avoids X. By the maximality of H it meets H. But it is not a component of H, it does not lie in $\mathcal{C}$, and it does not contain an H-path between distinct vertices outside U (by the maximality of H). So this cycle meets U. X

We have shown that every cycle in G meets $X \cup U$. As $|X| \leqslant k-1$, it thus suffices to show that $|U| < s_k$ unless H contains k disjoint cycles. But this follows from Lemma 2.3.1 applied to the multigraph obtained from H by suppressing its vertices of degree 2. □

A very short proof of Theorem 2.3.2 as such, with a recursively defined (and much worse) bound, is indicated in Exercise 11 of Chapter 9.

Our proof of Theorem 2.3.2 can be adapted to give an analogous result for packing cycles edge-disjointly and covering them by edges; this is outlined in Exercise 22 of Chapter 7. A simpler proof of the edge version using Ramsey's theorem is indicated in Exercise 12 of Chapter 9.

We shall also meet the Erdős-Pósa property again in Chapter 12. There, a considerable extension of Theorem 2.3.2 will appear as an unexpected and easy corollary of the theory of graph minors.

2.4 Tree packing and arboricity

In this section we consider packing and covering in terms of edges rather than vertices. How many edge-disjoint spanning trees can we find in a given connected graph? And how few trees, not necessarily edge-disjoint, suffice to cover all its edges? These two questions have two classical theorems answering them. But rather than proving these theorems directly, we shall obtain them both as corollaries of a beautiful recent unification due to Bowler and Carmesin: the *packing-covering* theorem.

To motivate the tree packing problem, assume for a moment that our graph represents a communication network, and that for every choice of two vertices we want to be able to find k edge-disjoint paths between them. Menger's theorem (3.3.6) in the next chapter will tell us that such paths exist as soon as our graph is k-edge-connected, which is clearly also necessary. This is a good theorem, but it does not tell us how to find those paths; in particular, having found them for one pair of endvertices we are not necessarily better placed to find them for another pair. If our graph has k edge-disjoint spanning trees, however, there will always be k canonical such paths, one in each tree. Once we have stored those trees in our computer, we shall always be able to find the k paths quickly, between any given pair of vertices.

When does a graph G have k edge-disjoint spanning trees? If it does, it clearly must be k-edge-connected. The converse, however, is easily seen to be false (try $k = 2$); indeed it is not even clear that any edge-connectivity will imply the existence of k edge-disjoint spanning trees. (But see Corollary 2.4.2 below.)

Here is another necessary condition. If G has k edge-disjoint spanning trees, then with respect to any partition of $V(G)$ into r sets, every spanning tree of G has at least $r-1$ *cross-edges*, edges whose ends lie in different partition sets. (Why?) Thus if G has k edge-disjoint spanning trees, it has at least $k\,(r-1)$ cross-edges. This condition is also sufficient:

cross-edges

tree packing theorem

[8.6.9]

Theorem 2.4.1. (Nash-Williams 1961; Tutte 1961)
A multigraph contains k edge-disjoint spanning trees if and only if for every partition P of its vertex set it has at least $k\,(|P|-1)$ cross-edges.

Theorem 2.4.1 has a striking corollary: $2k$-edge-connectedness is enough to ensure the existence of k edge-disjoint spanning trees.

[6.4.4]

Corollary 2.4.2. *Every $2k$-edge-connected multigraph G has k edge-disjoint spanning trees.*

Proof. Every class in a vertex partition of G is joined to other partition classes by at least $2k$ edges. Hence, for any partition into r sets, G has at least $\frac{1}{2}\sum_{i=1}^{r} 2k = kr$ cross-edges. The assertion thus follows from Theorem 2.4.1. $\square$

Note that the quantitative condition on cross-edges in Theorem 2.4.1 is equivalent to asking the same only for partitions into connected vertex sets: any other partition is refined by such a partition, and if the latter has enough cross-edges (even though it has more classes) then clearly so does the former. The tree packing theorem thus says that a multigraph has k edge-disjoint spanning trees as soon as all its contraction minors have enough edges to support k edge-disjoint spanning trees.

We shall meet Theorem 2.4.1 again in Chapter 8.6, where we prove an infinite analouge. This is based not on ordinary spanning trees (for which the result is false) but on 'topological spanning trees': the analogous structures in a topological space formed by the graph together with its ends, points at infinity that make it compact.

Let us now turn to the covering problem. To bring out its duality to the packing problem, we begin by rephrasing the latter. Let us say that some given subgraphs of a multigraph G form an *edge-decomposition* of G if their edge sets partition $E(G)$. Our spanning tree problem can now be recast as follows: into how *many connected* spanning subgraphs can we edge-decompose G? Since a spanning subgraph is connected if and only if it has an edge in every bond, the packing problem in this new guise has a 'dual' reminiscent of Theorems 1.5.1 and 1.9.4: into how *few acyclic* subgraphs – those whose complement meets all their circuits – can we edge-decompose G?

Let us say that some given graphs, not necessarily subgraphs of G, *cover its edges* if every edge of G lies in at least one of them. Our dual problem, then, is for which multigraphs G can we cover their edges by at most k trees. *cover*

An obvious necessary condition is that every set $U \subseteq V(G)$ induces at most $k\,(|U|-1)$ edges, no more than $|U|-1$ for each tree. Or, to phrase it dually to the tree packing condition, that no 'deletion minor' (subgraph) of G has too many edges to be covered by k trees.

Once more, this condition turns out to be sufficient too:

Theorem 2.4.3. (Nash-Williams 1964) *tree covering theorem*
The edges of a multigraph $G = (V, E)$ can be covered by at most k trees if and only if $\|G[U]\| \leqslant k\,(|U|-1)$ for every non-empty set $U \subseteq V$.

The least number of trees that can cover the edges of a graph is its *arboricity*. By Theorem 2.4.3, the arboricity of a graph is a measure for its maximum local density: it has small arboricity if and only if it is 'nowhere dense' in the sense that it has no subgraph H with $\varepsilon(H)$ large. *arboricity*

We finally come to the packing-covering theorem. Recall from Chapter 1.10 that when we form a contraction minor G/P of a multigraph G, we keep all the edges of G between different partition classes: edges between the same two classes $U, U' \in P$ become parallel edges of G/P.

packing-covering theorem

Theorem 2.4.4. (Bowler & Carmesin 2015)
For every connected multigraph $G = (V, E)$ and every $k \in \mathbb{N}$ there is a partition P of V such that every $G[U]$ with $U \in P$ has k edge-disjoint spanning trees and the edges of G/P can be covered by k spanning trees.

Before we prove the packing-covering theorem, let us deduce Theorems 2.4.1 and 2.4.3.

Proof of Theorem 2.4.1 (assuming Theorem 2.4.4).
Suppose a multigraph G has at least $k\,(|P|-1)$ cross-edges for every partition P of $V(G)$. Let P be the partition provided by Theorem 2.4.4. By the theorem, G/P has k spanning trees covering its edges. Since $\|G/P\| \geqslant k\,(|P|-1)$, they must be edge-disjoint. Combining them with the edge-disjoint spanning trees in the $G[U]$ that are also provided by Theorem 2.4.4, we obtain the desired k spanning trees of G. □

Proof of Theorem 2.4.3 (assuming Theorem 2.4.4).
Suppose every $U \subseteq V$ induces at most $k\,(|U|-1)$ edges in G. Let C be a component of G, and P the partition of $V(C)$ provided by Theorem 2.4.4. For each $U \in P$, each of the k edge-disjoint spanning trees of $G[U]$ that the theorem provides has $|U|-1$ edges, so all the edges of $G[U]$ lie in these trees. Combining these trees with the spanning trees of C/P that cover its edges, also provided by Theorem 2.4.4, we obtain k spanning trees of C covering its edges. These can be combined to k forests covering the edges of G. Add edges to turn these into the desired k trees. □

Given the power of the packing-covering theorem, its proof is strikingly short and elegant. To prepare some notation, consider a spanning tree T of G, a chord e, and an edge $f \in T$ on its fundamental cycle C_e. Then $T' = T + e - f$ is another spanning tree: this is immediate from Corollary 1.5.2, because T' is still connected and has the same number of edges as T. One says that T' is obtained from T by *exchanging* f for e.

$\mathcal{T}$, k

Now let $\mathcal{T} = (T_j)_{j=1,\dots,k}$ be a family of spanning trees of G. Call a sequence $(e_0, \dots, e_n)$ of edges of G an *exchange chain* with respect to $\mathcal{T}$, *started by* e_0, if e_n lies on none of these trees but for every $i < n$ there exists $j =: j(i)$ such that e_{i+1} is a chord of T_j whose fundamental cycle with respect to T_j contains e_i.

exchange chain

$i \mapsto j(i)$

$E(\mathcal{T})$

Let us write $E(\mathcal{T}) := \bigcup \{\, E(T) \mid T \in \mathcal{T} \}$ for any such family.

Lemma 2.4.5. *If e_0 starts an exchange chain with respect to $\mathcal{T}$ and lies in two of its trees, then there is a family $\mathcal{T}'$ of k spanning trees of G such that $E(\mathcal{T}) \subsetneq E(\mathcal{T}')$.*

Proof. Choose $(e_0, \dots, e_n)$ of minimum length among the exchange chains with respect to $\mathcal{T}$ that start with e_0. Then its elements e_i are distinct. And no e_i lies on the fundamental cycle with respect to any tree in $\mathcal{T}$ of any e_ℓ with $\ell > i+1$: if it did, we could omit $e_{i+1}, \dots, e_{\ell-1}$ from our sequence, contradicting its minimality.

For $i = 0, \dots, n$ inductively, we shall define families $\mathcal{T}^i = (T^i_j)_{j=1,\dots,k}$ T^i_j
of spanning trees of G such that, for all $i \leqslant \ell < n$,

(i) $E(\mathcal{T}^i) = E(\mathcal{T})$;

(ii) e_i lies in two of the trees in the family $\mathcal{T}^i$;

(iii) $e_{\ell+1}$ is a chord of $T^i_{j(\ell)}$ and has the same fundamental cycle there as with respect to $T_{j(\ell)}$.

Note that these families have the same index set as $\mathcal{T}$, and we continue to use the map $i \mapsto j(i)$ from the indices of our exchange chain to the $j(i)$
indices of $\mathcal{T}^i$ that we defined together with $\mathcal{T}$.

For $i = 0$ the family $\mathcal{T}^0 := \mathcal{T}$ is as required. Given $i < n$, let $\mathcal{T}^{i+1}$ $\mathcal{T}^i$
be obtained from $\mathcal{T}^i$ by replacing $T^i_{j(i)}$ with $T^i_{j(i)} + e_{i+1} - e_i =: T^{i+1}_{j(i)}$ and letting $T^{i+1}_j := T^i_j$ for all other j.

By (iii) for i with $\ell = i$, the modified tree $T^{i+1}_{j(i)}$ is again a spanning tree of G: recall that e_i lies on the fundamental cycle of e_{i+1} with respect to $T_{j(i)}$, by definition of $j(i)$.

Assertion (i) for $i+1 < n$ follows from (i)–(iii) for i: we are not losing e_i by (ii), and we are not gaining e_{i+1}, since it lies on the fundamental cycle of e_{i+2} with respect to $T_{j(i+1)}$, and hence is already in $T^i_{j(i+1)} \in \mathcal{T}^i$ by (iii).

To verify (ii) for $i+1 < n$, note first that $j(i+1) \neq j(i)$, because e_{i+1} is a chord of $T_{j(i)}$ but lies on $T_{j(i+1)}$. Now $e_{i+1} \in T^i_{j(i+1)} = T^{i+1}_{j(i+1)}$ as noted above, as well as $e_{i+1} \in T^{i+1}_{j(i)}$, by definition of this tree.

For a proof of (iii) for $i+1 < n$ consider any $i+1 \leqslant \ell < n$. As long as $j(\ell) \neq j(i)$, we have $T^i_{j(\ell)} = T^{i+1}_{j(\ell)}$, and the assertion follows from (iii) for i. Now consider the case that $j(\ell) = j(i) =: j$. Then $e_{\ell+1}$ is a chord of T^i_j by (iii) for i, and $T^{i+1}_j = T^i_j + e_{i+1} - e_i$. As $e_{i+1} \neq e_{\ell+1}$ since $i \neq \ell$, our edge $e_{\ell+1}$ is still a chord of T^{i+1}_j. As noted at the start of the proof, its fundamental cycle with respect to T_j does not contain e_i. Hence (iii) for i implies (iii) for $i+1$.

It remains to check that $\mathcal{T}' := \mathcal{T}^n$ satisfies $E(\mathcal{T}) \subsetneq E(\mathcal{T}')$. As $e_n \in E(\mathcal{T}')$ by definition of $\mathcal{T}' = \mathcal{T}^n$, and $e_n \notin E(\mathcal{T})$ by assumption, all we have to check is that $E(\mathcal{T}) \subseteq E(\mathcal{T}')$. Now $E(\mathcal{T}) = E(\mathcal{T}^{n-1})$, by (i) for $i = n-1$. But the only edge we deleted from any tree in $\mathcal{T}^{n-1}$ when we turned it into $\mathcal{T}^n$ was e_{n-1}. By (ii) for $i = n-1$, this edge also lay on another tree in $\mathcal{T}^{n-1}$, which we left unchanged for $\mathcal{T}^n$. $\square$

Proof of Theorem 2.4.4. Let $\mathcal{T} = (T_1, \dots, T_k)$ be a family of k spanning trees of G, chosen with $E(\mathcal{T})$ maximal. Let D be the set of all edges of G that start an exchange chain with respect to $\mathcal{T}$. These include all edges not in $E(\mathcal{T})$, since they form singleton exchange chains. Let P be the partition of V into the vertex sets of the components of (V, D).

For the theorem's packing assertion, let $U \in P$ be given. For all $j = 1, \ldots, k$ let S_j be the subgraph of T_j induced on U by its edges in D. These forests S_j are edge-disjoint, since by the maximality of $E(\mathcal{T})$ and Lemma 2.4.5 no edge in D lies in more than one T_j. Let us show that the S_j are connected.

Since the edges from D form a connected submultigraph on U, it suffices to show that for every edge $uu' \in D$ with $u, u' \in U$ there is a u–u' path in S_j. This is clear if uu' lies in T_j, and hence in S_j. If it does not, then the path uT_ju' still has all its edges e in D, and hence lies in S_j: if $e_0, \ldots, e_n$ is an exchange chain witnessing that $e_0 = uu' \in D$, then $e, e_0, \ldots e_n$ is an exchange chain putting e in D, because e lies on the fundamental cycle of e_0 with respect to T_j.

As every T_j induces connected subgraphs S_j on the partition classes of P, contracting these S_j turns the T_j into spanning trees T'_j of G/P. These T'_j cover all the edges of G/P, since $E \setminus E(\mathcal{T}) \subseteq D$. □

The packing-covering theorem differs from both the tree packing and the tree covering theorem in a fundamental way. The non-trivial directions of the latter two theorems each obtain a structural assertion about a graph, the existence of a packing or covering, as a consequence of quantitative assumptions about all their minors of a certain type: contraction minors for the packing theorem, and 'deletion minors' – i.e., subgraphs – for the covering theorem. This format makes them interesting: they offer valuable structural information for one graph in exchange for less valuable quantitative information about many smaller graphs.

The packing-covering theorem, by contrast, makes a structural assertion about every graph: with no need for any assumptions at all, neither quantitative nor qualitative.

The packing-covering theorem extends to infinite graphs in two interestingly different ways; see Exercises 20 and 131 in Chapter 8.

2.5 Path covers

Let us return once more to König's duality theorem for bipartite graphs, Theorem 2.1.1. If we orient every edge of G from A to B, the theorem tells us how many disjoint directed paths we need in order to cover all the vertices of G: every directed path has length 0 or 1, and clearly the number of paths in such a 'path cover' is smallest when it contains as many paths of length 1 as possible – in other words, when it contains a maximum-cardinality matching.

In this section we put the above question more generally: how many paths in a given directed graph will suffice to cover its entire vertex set? Of course, this could be asked just as well for undirected graphs. As it turns out, however, the result we shall prove is rather more trivial in the undirected case (exercise), and the directed case will also have an interesting corollary.

A *directed path* is a directed graph $P \neq \emptyset$ with distinct vertices $x_0, \dots, x_k$ and edges $e_0, \dots, e_{k-1}$ such that e_i is an edge directed from x_i to x_{i+1}, for all $i < k$. In this section, *path* will always mean 'directed path'. The vertex x_k above is the *last vertex* of the path P, and when $\mathcal{P}$ is a set of paths we write $\text{ter}(\mathcal{P})$ for the set of their last vertices. A *path cover* of a directed graph G is a set of disjoint paths in G which together contain all the vertices of G. *path* $\text{ter}(\mathcal{P})$ *path cover*

Theorem 2.5.1. (Gallai & Milgram 1960)
Every directed graph G has a path cover $\mathcal{P}$ and an independent set $\{ v_P \mid P \in \mathcal{P} \}$ of vertices such that $v_P \in P$ for every $P \in \mathcal{P}$.

Proof. Clearly, G has a path cover, e.g. by trivial paths. We prove by induction on $|G|$ that for every path cover $\mathcal{P} = \{P_1, \dots, P_m\}$ with $\text{ter}(\mathcal{P})$ minimal there is a set $\{ v_P \mid P \in \mathcal{P} \}$ as claimed. For each i, let v_i denote the last vertex of P_i. $\mathcal{P},\ P_i$ v_i

If $\text{ter}(\mathcal{P}) = \{v_1, \dots, v_m\}$ is independent there is nothing more to show, so we assume that G has an edge from v_2 to v_1. Since $P_2 v_2 v_1$ is again a path, the minimality of $\text{ter}(\mathcal{P})$ implies that v_1 is not the only vertex of P_1; let v be the vertex preceding v_1 on P_1. Then $\mathcal{P}' := \{P_1 v, P_2, \dots, P_m\}$ is a path cover of $G' := G - v_1$ (Fig. 2.5.1). Clearly, any independent set of representatives for $\mathcal{P}'$ in G' will also work for $\mathcal{P}$ in G, so all we have to check is that we may apply the induction hypothesis to $\mathcal{P}'$. It thus remains to show that $\text{ter}(\mathcal{P}') = \{v, v_2, \dots, v_m\}$ is minimal among the sets of last vertices of path covers of G'. v $\mathcal{P}', G'$

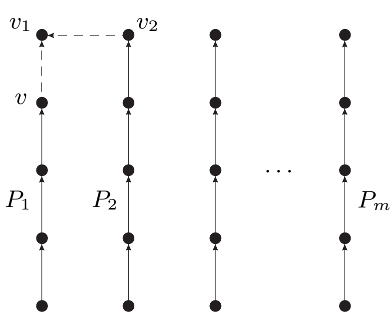

Fig. 2.5.1. Path covers of G and G'

Suppose then that G' has a path cover $\mathcal{P}''$ with $\text{ter}(\mathcal{P}'') \subsetneq \text{ter}(\mathcal{P}')$. If a path $P \in \mathcal{P}''$ ends in v, we may replace P in $\mathcal{P}''$ by $P v v_1$ to obtain a path cover of G whose set of last vertices is a proper subset of $\text{ter}(\mathcal{P})$, contradicting the choice of $\mathcal{P}$. If a path $P \in \mathcal{P}''$ ends in v_2 (but none in v), we similarly replace P in $\mathcal{P}''$ by $P v_2 v_1$ to obtain a contradiction to the minimality of $\mathrm{ter}(\mathcal{P})$. Hence $\mathrm{ter}(\mathcal{P}'') \subseteq \{v_3, \ldots, v_m\}$. But now $\mathcal{P}''$ and the trivial path $\{v_1\}$ together form a path cover of G that contradicts the minimality of $\mathrm{ter}(\mathcal{P})$. □

chain
antichain

As a corollary to Theorem 2.5.1 we obtain a classical result from the theory of partial orders. Recall that a subset of a partially ordered set $(P, \leqslant)$ is a *chain* in P if its elements are pairwise comparable; it is an *antichain* if they are pairwise incomparable.

Corollary 2.5.2. (Dilworth 1950)
In every finite partially ordered set $(P, \leqslant)$, the minimum number of chains with union P is equal to the maximum cardinality of an antichain in P.

Proof. If A is an antichain in P of maximum cardinality, then clearly P cannot be covered by fewer than $|A|$ chains. The fact that $|A|$ chains will suffice follows from Theorem 2.5.1 applied to the directed graph on P with the edge set $\{\, (x, y) \mid x < y \,\}$. □

Exercises

1. Let M be a matching in a bipartite graph G. Show that if M is sub-optimal, i.e. contains fewer edges than some other matching in G, then G contains an augmenting path with respect to M. Does this fact generalize to matchings in non-bipartite graphs?

2. (continued)

   Describe an algorithm that finds, as efficiently as possible, a matching of maximum cardinality in any bipartite graph.

3. Show that if there exist injective functions $A \to B$ and $B \to A$ between two infinite sets A and B then there exists a bijection $A \to B$.

4. Moving alternately, two players jointly construct a path in some fixed graph G. If $v_1 \ldots v_n$ is the path constructed so far, the player to move next has to find a vertex v_{n+1} such that $v_1 \ldots v_{n+1}$ is again a path. Whichever player cannot move loses. For which graphs G does the first player have a winning strategy, for which the second?

5.$^-$ Derive the marriage theorem from König's theorem.

6. Let G and H be defined as for the third proof of Hall's theorem. Show that $d_H(b) \leqslant 1$ for every $b \in B$, and deduce the marriage theorem.

7. Does our first proof of the marriage theorem use the assumption that the graph is finite? If so, can it be adapted so that it works for infinite graphs too?

8. Let k be an integer. Show that any two partitions of a finite set into k-sets admit a common choice of representatives.

9. Let A be a finite set with subsets $A_1, \dots, A_n$, and let $d_1, \dots, d_n \in \mathbb{N}$. Show that there are disjoint subsets $D_k \subseteq A_k$, with $|D_k| = d_k$ for all $k \leqslant n$, if and only if

$$\Big|\bigcup_{i \in I} A_i\Big| \geqslant \sum_{i \in I} d_i$$

for all $I \subseteq \{1, \dots, n\}$.

10.$^+$ Prove that in an n-set X there are never more than $\binom{n}{\lfloor n/2 \rfloor}$ subsets such that none of these contains another.

(Hint. Construct $\binom{n}{\lfloor n/2 \rfloor}$ chains covering the power set lattice of X.)

11. Let G be a bipartite graph with bipartition $\{A, B\}$. Assume that $\delta(G) \geqslant 1$, and that $d(a) \geqslant d(b)$ for every edge ab with $a \in A$. Show that G contains a matching of A.

12. Find a bipartite graph with a set of preferences such that no matching of maximum size is stable and no stable matching has maximum size. Find a non-bipartite graph with a set of preferences that has no stable matching.

13.$^-$ Consider the algorithm described in the proof of the stable marriage theorem. Observe that once a vertex of B is matched, she remains matched and gets happier with every change of her matching edge. On the other hand, show that the sequence of matching edges incident with a given vertex of A makes this vertex unhappier with every change (disregarding the interim periods when he is unmatched).

14. Show that all stable matchings of a given graph cover the same vertices. (In particular, they have the same size.)

15.$^+$ Show that the algorithm in our proof of Theorem 2.1.4 produces a matching M such that no other stable matching makes any vertex in A happier or any vertex in B unhappier than he or she is in M. Consider only matched vertices for happiness.

16.$^+$ Show that the following 'obvious' algorithm need not produce a stable matching in a bipartite graph. Start with any matching. If the current matching is not maximal, add an edge. If it is maximal but not stable, insert an edge that creates instability, deleting any current matching edges at its ends.

17. Show that the union of two partial orderings $\leqslant_1, \leqslant_2$ of a finite set P has a 'dominating antichain', a set $A \subseteq P$ such that no two elements of A are related in either $\leqslant_1$ or $\leqslant_2$ and for every $x \in P$ there exists an $a \in A$ such that $x \leqslant_1 a$ or $x \leqslant_2 a$. Deduce Theorem 2.1.4.

18. Find a set S for Theorem 2.2.3 when G is a forest.

19. A graph G is called (vertex-) *transitive* if, for any two vertices $v, w \in G$, there is an automorphism of G mapping v to w. Using the observations following the proof of Theorem 2.2.3, show that every transitive connected graph of even order contains a 1-factor.

20.$^+$ Show that a graph G contains k independent edges if and only if $q(G - S) \leqslant |S| + |G| - 2k$ for all sets $S \subseteq V(G)$.

21.$^-$ Find a cubic graph without a 1-factor.

22.$^+$ Derive the marriage theorem from Tutte's theorem.

23.$^-$ Disprove the analogue of König's theorem (2.1.1) for non-bipartite graphs, but show that $\mathcal{H} = \{K^2\}$ has the Erdős-Pósa property.

24. Let T be a tree and $\mathcal{T}$ a set of subtrees of T. Show that the maximum number of disjoint trees in $\mathcal{T}$ equals the least cardinality of a set X of vertices such that $T - X$ contains no tree from $\mathcal{T}$.

25. For cubic graphs, Lemma 2.3.1 is considerably stronger than the Erdős-Pósa theorem. Extend the lemma to arbitrary multigraphs of minimum degree $\geqslant 3$, by finding a function $g\colon \mathbb{N} \to \mathbb{N}$ such that every multigraph of minimum degree $\geqslant 3$ and order at least $g(k)$ contains k disjoint cycles, for all $k \in \mathbb{N}$. Alternatively, show that no such function g exists.

26. Given a graph G, let $\alpha(G)$ denote the largest size of a set of independent vertices in G. Prove that the vertices of G can be covered by at most $\alpha(G)$ disjoint subgraphs each isomorphic to a cycle or a K^2 or K^1.

27. Show that if G has two edge-disjoint spanning trees, it has a connected spanning subgraph all whose degrees are even.

28. In the proofs of Theorems 2.4.1, 2.4.3 and 2.4.4, there is exactly one place where we use that we are working with multigraphs. Where is it?

29. Find the error in the following short 'proof' of Theorem 2.4.1. Call a partition *non-trivial* if it has at least two classes and at least one of the classes has more than one element. We show by induction on $|V| + |E|$ that $G = (V, E)$ has k edge-disjoint spanning trees if every non-trivial partition of V into r sets (say) has at least $k(r-1)$ cross-edges. The induction starts trivially with $G = K^1$ if we allow k copies of K^1 as a family of k edge-disjoint spanning trees of K^1. We now consider the induction step. If every non-trivial partition of V into r sets (say) has more than $k(r-1)$ cross-edges, we delete any edge of G and are done by induction. So V has a non-trivial partition $\{V_1, \dots, V_r\}$ with exactly $k(r-1)$ cross-edges. Assume that $|V_1| \geqslant 2$. If $G' := G[V_1]$ has k disjoint spanning trees, we may combine these with k disjoint spanning trees that exist in G/V_1 by induction. We may thus assume that G' has no k disjoint spanning trees. Then by induction it has a non-trivial vertex partition $\{V'_1, \dots, V'_s\}$ with fewer than $k(s-1)$ cross-edges. Then $\{V'_1, \dots, V'_s, V_2, \dots, V_r\}$ is a non-trivial vertex partition of G into $r + s - 1$ sets with fewer than $k(r-1) + k(s-1) = k((r+s-1)-1)$ cross-edges, a contradiction.

30. A graph G is called *balanced* if $\varepsilon(H) \leqslant \varepsilon(G)$ for every subgraph $H \subseteq G$.
    (i) Find a few natural classes of balanced graphs.
    (ii) Show that the arboricity of a balanced graph is bounded above by its average degree. Is it even bounded by ε? Or by $\varepsilon + 1$?
    (iii) Characterize, in terms of the balanced graphs or otherwise, the graphs G such that $\varepsilon(H) \geqslant \varepsilon(G)$ for every induced subgraph $H \subseteq G$.

31. Rephrase König's and Dilworth's theorems as pure existence statements without any inequalities.

32.$^-$ Prove the undirected version of the theorem of Gallai & Milgram (without using the directed version).

33. Derive the marriage theorem from the theorem of Gallai & Milgram.

34.$^-$ Show that a partially ordered set of at least $rs + 1$ elements contains either a chain of size $r + 1$ or an antichain of size $s + 1$.

35. Prove the following dual version of Dilworth's theorem: in every finite partially ordered set $(P, \leqslant)$, the minimum number of antichains with union P is equal to the maximum cardinality of a chain in P.

36. Derive König's theorem from Dilworth's theorem.

37. Find a partially ordered set that has no infinite antichain but is not a union of finitely many chains.

Notes

There is a very readable and comprehensive monograph about matching in finite graphs: L. Lovász & M.D. Plummer, *Matching Theory*, Annals of Discrete Math. **29**, North Holland 1986. Two other very comprehensive sources are A. Schrijver, *Combinatorial optimization*, Springer 2003, and A. Frank, *Connections in combinatorial optimization*, Oxford University Press 2011. All the references for the results in this chapter can be found in these books.

All the main theorems in this chapter are of a particularly attractive type, which they share with several other fundamental theorems in graph theory. They each assert that in some obvious inequality we can attain equality, or that some obvious potential obstruction to some desirable property must necessarily occur in all graphs that fail to have this property. Phrased in yet another way, these theorems provide easily checked 'certificates' for the absence of some desirable property, which comes with its own obvious certificate when it does hold. The theorems of König, Hall, Tutte, Dilworth, and Nash-Williams proved in this chapter can all be viewed in this way.

As we shall see in Chapter 3, König's Theorem of 1931 is no more than the bipartite case of a more general theorem that is also of this type, which is generally attributed to Menger (1927). However, Menger missed the bipartite case in his original proof. When Menger showed König his theorem and proof during a visit to Budapest in 1930, they seem to have noticed this gap. König published is proof in two papers of 1931 and 1933, and quotes Menger as claiming to have settled this case independently.

At the time, neither of these results was nearly as well known as Hall's marriage theorem, which he proved even later, in 1935. To this day, Hall's theorem remains one of the most applied graph-theoretic results. The first two of our proofs are folklore. The edge-minimal subgraph approach of our third proof can be traced back to a paper of Rado (1967); our version and its dual, Exercise 6, are due to Kriesell.

More on the stable marriage theorem can be found in D. Gusfield & R.W. Irving, *The Stable Marriage Problem: Structure and Algorithms*, MIT Press 1989, and in A. Tamura, Transformation from arbitrary matchings to stable matchings, *J. Comb. Theory, Ser. A* **62** (1993), 310–323. How the world outside mathematics sees the stable marriage theorem can be gleaned from https://www.nobelprize.org/prizes/economic-sciences/2012/shapley/facts/.

Our proof of Tutte's 1-factor theorem is based on a proof by Lovász (1975). Our extension of Tutte's theorem, Theorem 2.2.3 (including the informal discussion following it) is a lean version of a comprehensive structure theorem for matchings, due to Gallai (1964) and Edmonds (1965). See Lovász & Plummer for a detailed statement and discussion of this theorem.

Theorem 2.3.2 is due to P. Erdős & L. Pósa, On independent circuits contained in a graph, *Canad. J. Math.* **17** (1965), 347–352. Our proof is essentially due to M. Simonovits, A new proof and generalization of a theorem of Erdős and Pósa on graphs without $k+1$ independent circuits, *Acta Sci. Hungar* **18** (1967), 191–206. Calculations such as in Lemma 2.3.1 are standard for proofs where one aims to bound one numerical invariant in terms of another. This book does not emphasize this aspect of graph theory, but it is not atypical.

There is also an analogue of the Erdős-Pósa theorem for directed graphs, due to B. Reed, N. Robertson, P.D. Seymour and R. Thomas, Packing directed circuits, *Combinatorica* **16** (1996), 535–554. Its proof is more difficult than the undirected case; see Chapter 12.6, and in particular Theorem 12.6.5, for a glimpse of the techniques used.

The tree packing theorem, Theorem 2.4.1, was proved independently by Nash-Williams and Tutte; both papers are contained in *J. Lond. Math. Soc.* **36** (1961). The tree covering theorem, Theorem 2.4.3, is due to C.St.J.A. Nash-Williams, Decompositions of finite graphs into forests, *J. Lond. Math. Soc.* **39** (1964), 12. The partitions whose existence is asserted by the packing-covering theorem, Theorem 2.4.4, were first constructed explicitly by B. Jackson and T. Jordán, Brick partitions of graphs, *Discrete Math.* **310** (2010), 270–275. They may not be unique, and are interesting in their own right; see the paper, and Frank's monograph cited earlier, for more.

The packing-covering theorem itself, together with its direct proof that does not rely on the classical tree packing and covering theorems but implies them, is from N. Bowler and J. Carmesin, Matroid intersection, base packing and base covering for infinite matroids, *Combinatorica* **35** (2015), 153–180, arXiv:1202.3409.

It has long been known that the tree packing and covering theorems can be naturally expressed in terms of matroids; see Schrijver's book cited earlier. However it was only recently when infinite matroids were axiomatized and thus made accessible to systematic study, forcing the translation of quantitative assertions about finite matroids into structural ones to make them meaningful also for infinite matroids, that Bowler and Carmesin found the packing-covering theorem. The main focus of their paper is to show how the unproved infinite version of the packing-covering theorem for matroids, the *packing-covering conjecture*, plays a central role in infinite matroid theory. The conjecture implies, among other things, the Aharoni-Berger theorem for infinite graphs (Theorem 8.4.4), one of the deepest theorems in graph theory.

The packing-covering theorem extends to infinite graphs in two ways: with ordinary spanning trees (Exercise 20, Ch. 8), and with 'topological' spanning trees (Exercise 131, Ch. 8). These infinite versions also follow from two cases of the infinite packing-covering conjecture that Bowler and Carmesin prove in their paper, those for finitary and for cofinitary matroids.

An interesting vertex analogue of Corollary 2.4.2 is to ask which connectivity forces the existence of k spanning trees $T_1, \ldots, T_k$, all rooted at a given vertex r, such that for every vertex v the k paths vT_ir are independent. For example, if G is a cycle then deleting the edge left or right of r produces two such spanning trees. A. Itai and A. Zehavi, Three tree-paths, *J. Graph Theory* **13** (1989), 175–187, conjectured that $\kappa \geqslant k$ should suffice. This conjecture has been proved for $k \leqslant 4$; see S. Curran, O. Lee & X. Yu, Chain decompositions and independent trees in 4-connected graphs, *Proc. 14th Ann. ACM SIAM symposium on Discrete algorithms* (Baltimore 2003), 186–191.

Theorem 2.5.1 is due to T. Gallai & A.N. Milgram, Verallgemeinerung eines graphentheoretischen Satzes von Rédei, *Acta Sci. Math. (Szeged)* **21** (1960), 181–186.