Reinhard Diestel

Graph Theory

4th Electronic Edition 2010

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 $\mathcal{H}$-graphs 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 on G? In *packing* *covering*

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.

alternating path

augmenting path

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.

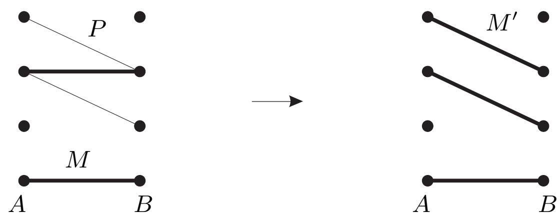

Fig. 2.1.1. Augmenting the matching M by the alternating path P

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

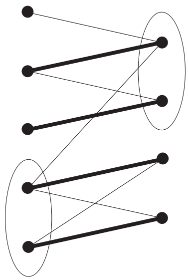

Fig. 2.1.2. The vertex cover U

Let $ab \in E$ be an edge; we show that either a or b lies in U. If $ab \in M$, this holds by definition of U, so we assume that $ab \notin M$. Since M is a maximal matching, it contains an edge $a'b'$ with $a = a'$ or $b = b'$. In fact, we may assume that $a = a'$: for if a is unmatched (and $b = b'$), then ab is an alternating path, and so the end of $a'b' \in M$ chosen for U was the vertex $b' = b$. Now if $a' = a$ is not in U, then $b' \in U$, and some alternating path P ends in b'. But then there is also an alternating path P' ending in b: either $P' := Pb$ (if $b \in P$) or $P' := Pb'a'b$. By the maximality of M, however, P' is not an augmenting path. So b must be matched, and was chosen for U from the edge of M containing it. $\square$

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

Let A' be the set of vertices in A that can be reached from a by a non-trivial alternating path, and $B' \subseteq B$ the set of all penultimate vertices of such paths. The last edges of these paths lie in M, so $|A'| = |B'|$. Hence by the marriage condition, there is an edge from a vertex v in $S = A' \cup \{a\}$ to a vertex b in $B \setminus B'$.

As $v \in A' \cup \{a\}$, there is an alternating path P from a to v. Then either Pvb or Pb (if $b \in P$) is an alternating path from a to b; call this path P'. If b was matched, by $a'b \in M$ say, then $P'ba'$ would be an alternating path putting a' in A' and b in B'. But $b \notin B'$, so b is unmatched, and P' is the desired augmenting path. □

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 $S \subsetneqq A$, we pick an edge $ab \in G$ and consider the graph $G' := G - \{a, b\}$. 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 $S = \{a\}$.

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

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.

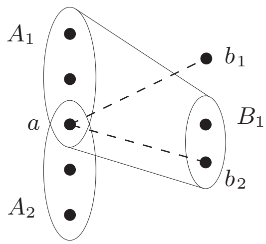

Fig. 2.1.3. B_1 contains b_2 but not b_1

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 $i = 1, 2$ 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. *If G is k-regular with $k \geqslant 1$, then G 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 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*:

stable matching

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

Proof (compare Exercise 15). 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 each 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 it is easy to check inductively that they all keep the vertices in A happy. So the sequence continues until it terminates with a matching M such that every unmatched vertex in A is inacceptable to all 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)

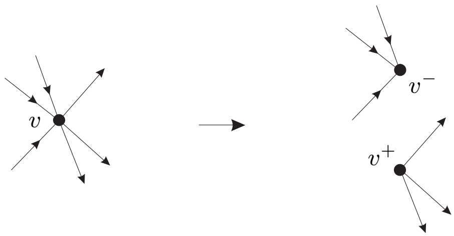

Fig. 2.1.4. Splitting vertices in the proof of Corollary 2.1.5

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 $\mathcal{C}_G$ $q(G)$

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

Tutte's condition

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

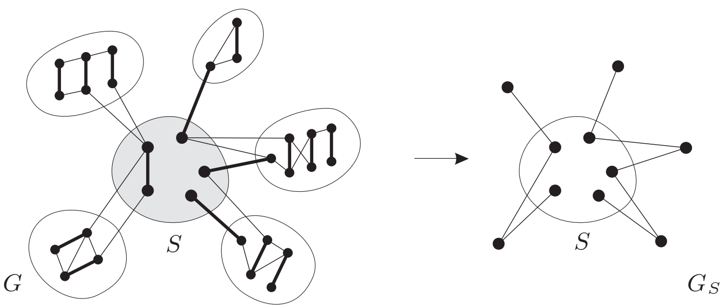

Fig. 2.2.1. Tutte's condition $q(G-S) \leqslant |S|$ for $q=3$, and the contracted graph G_S from Theorem 2.2.3.

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,

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

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 $G + bd$.

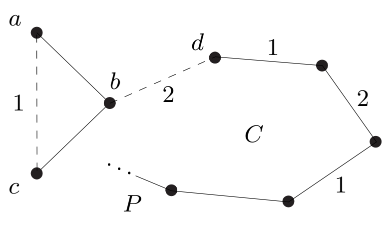

Fig. 2.2.2. Deriving a contradiction if S does not satisfy $(*)$

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 be the cycle $dPvbd$. In each 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 contained in E, a contradiction. □

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.

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.)

factor-critical

matchable

G_S

Theorem 2.2.3. *Every graph $G = (V, E)$ contains a vertex set S 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.

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

(2.1.2) **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 $|S| = |\mathcal{C}_{G-S}|$, 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 $T \subseteq V$ 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 $G - (S \setminus S')$, the set $T = S \setminus S'$ 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. 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 $\mathcal{C}$ $k_S, k_{\mathcal{C}}$

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

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 structural result, known as the *Gallai-Edmonds matching theorem*, is given in the notes at the end of this chapter.

2.3 Packing and covering

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 its

Erdős-Pósa property 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 multi-

m graph 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.

n Suppose we can perform a sequence of n such moves, but no more. (Our assumptions imply $n \leqslant 2$, 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 cut 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

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

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

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

□

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 $z \in Z$ 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.

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 X 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.

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. □

We shall 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 graph? And how few trees in it, not necessarily edge-disjoint, suffice to cover all its edges?

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 (ii)) 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, for any given pair of endvertices.

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

cross-edges 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:

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.

Before we prove Theorem 2.4.1, let us note a surprising corollary: to ensure the existence of k edge-disjoint spanning trees, it suffices to raise the edge-connectivity to just $2k$:

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

Proof. Every set in a vertex partition of G is joined to other partition sets 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. □

For a proof of the non-trivial direction of Theorem 2.4.1, let a multigraph $G = (V, E)$ and $k \in \mathbb{N}$ be given. Our approach will be as follows. Since we do not know yet whether the desired trees exist, we start with a family of k edge-disjoint spanning *forests*. These certainly exist; let $F_1, \dots, F_k$ be a choice whose total set of edges, $E(F_1 \cup \dots \cup F_k)$, is maximal. Next, we look for a set U of at least two vertices that is connected in every F_i. If $U = V$, then our forests F_i are in fact trees and we are done. If $U \subsetneqq V$, we contract U and apply induction to the multigraph G/U. The resulting k spanning trees of G/U can then be turned into spanning trees of G by inserting the trees $F_i[U]$. $G = (V, E)$ k $F_1, \dots, F_k$

How shall we construct such a set U? For reasons that will become clear later, we shall start by finding a set $U_0 = \{x^*, y^*\}$ of two vertices that are adjacent in G but not in any F_i. We want its vertices to be linked in $F_i[U]$ for every i, but they are not linked in $F_i[U_0]$. So we have to add some paths: let H_1 be the union of the paths $x^* F_i y^*$, one for each i, and $U_1 := V(H_1)$. (We shall have to show that these paths exist.) Now x^* and y^* are linked in $F_i[U_1]$ for every i. But we have many new pairs of vertices, and these may not yet be linked in every $F_i[U_1]$. To link these vertices too, we add some more paths—and so on. As our graph G is finite, this process will eventually stabilize: with a set U_n of vertices that is connected in every F_i. U_0

Lemma 2.4.3. *For every edge $e^* = x^* y^*$ in $E \smallsetminus E(F_1 \cup \dots \cup F_k)$ there exists a set $U \subseteq V$ that is connected in $F_i[U]$ for every $i = 1, \dots, k$ and contains both x^* and y^*.*

Proof. Consider the (unique) maximal sequence $\emptyset = \mathcal{P}_0 \subseteq \mathcal{P}_1 \subseteq \dots \subseteq \mathcal{P}_n$ of sets of paths such that $\mathcal{P}_\ell \smallsetminus \mathcal{P}_{\ell-1} \neq \emptyset$ and n

$$\mathcal{P}_\ell = \mathcal{P}_{\ell-1} \cup \bigcup \{\, xF_i y \mid xy \in E_{\ell-1};\ i = 1 \dots, k \,\} \qquad (1)$$

for all $\ell \geqslant 1$, where $E_0 := \{e^*\}$ and $E_\ell := E\big(\bigcup \mathcal{P}_\ell\big) \smallsetminus \bigcup_{i<\ell} E_i$ for all $\ell \geqslant 1$. (We shall prove in a moment that those paths $xF_i y$ exist, i.e. that x and y lie in the same component of F_i.) Thus for $\ell \geqslant 1$, the set E_ℓ consists of the edges 'added in step n' (induction on ℓ). Every path $P \subseteq F_j$ added at that time will be emulated by walks added to the other F_i in the next step, when we link the ends x, y of any edge of P also in F_i by adding $xF_i y$ if necessary. We finally define $\mathcal{P}_{n+1}$ by (1) for $\ell = n+1$, but note that $\mathcal{P}_{n+1} \smallsetminus \mathcal{P}_n = \emptyset$ by the maximality of n. E_ℓ

For all $1 \leqslant \ell \leqslant n+1$ write $H_\ell := \bigcup \mathcal{P}_\ell$ and $U_\ell := V(H_\ell)$. For every $P \in \mathcal{P}_n$ let $\ell(P)$ be the unique ℓ such that $P \in \mathcal{P}_\ell \smallsetminus \mathcal{P}_{\ell-1}$. For every $e \in \bigcup_{\ell=0}^{n} E_\ell$ let $\ell(e)$ be the unique ℓ such that $e \in E_\ell$. For every $1 \leqslant \ell \leqslant n$ and $e \in E_\ell$ choose a path $P(e) \in \mathcal{P}_\ell$ containing e. For every $1 \leqslant \ell \leqslant n$ and $P \in \mathcal{P}_\ell \smallsetminus \mathcal{P}_{\ell-1}$ choose an edge $e(P) \in E_{\ell-1}$ joining its ends. (Thus if $P = xF_i y$, then $e(P)$ is one of the—possibly several parallel—edges $H_\ell,\ U_\ell$ $\ell(P),\ \ell(e)$ $P(e),\ e(P)$

$xy \in E_{\ell-1}$ that gave rise to P in (1).) Then, for every e and P,

$$\ell(e) = \ell(P(e)) \quad \text{and} \quad \ell(P) > \ell(e(P))\,. \tag{2}$$

Let us show that the paths xF_iy in (1) always exist, i.e., that for every edge $xy \in \bigcup_{\ell=0}^{n} E_\ell$ and every i the vertices x, y lie in the same component of F_i. Suppose not, and add the edge xy to F_i; then F_i remains a forest. Now consider the maximal sequence $e_1, P_1, e_2, P_2, e_3, \dots$ such that $e_1 = xy$, and $P_q = P(e_q)$ and $e_{q+1} = e(P_q)$ for all q. This sequence can end only with the edge e^*, since for all edges $e \in \bigcup_{\ell=1}^{n} E_\ell$ the path $P(e)$ is defined, and for every path $P \in \mathcal{P}_n$ the edge $e(P)$ is defined. And since $\ell(e_q) = \ell(P_q) > \ell(e_{q+1})$ for all q by (2), the sequence does end. Recall that we added the edge e_1 to F_i. For $q = 1, 2, \dots$ inductively, we now delete e_q from the forest F_j containing P_q and add e_{q+1} to that forest. Since $P_q + e_{q+1}$ is a cycle, this operation preserves F_j as a forest: the addition of e_{q+1} to $F_j - e_q$ does not create a cycle in $F_j - e_q$. Each of our forests F_j may change several times in this process, but eventually, after adding e^* to the last of these forests, we shall have a new family $(F_1, \dots, F_k)$ of edge-disjoint spanning forests with one more edge in total than before; this contradicts our initial choice of $F_1, \dots, F_k$.

For our proof that U_n is connected in every F_i, let us show inductively for $\ell = 1, \dots, n+1$ that

$$(\forall\, v \in U_{\ell-1})(\forall\, i \in \{1, \dots, k\})\big(F_i \cap H_\ell \text{ contains a } v\text{–}x^* \text{ path}\big) \tag{3}$$

$$(\forall\, v \in U_\ell)(\exists\, j \in \{1, \dots, k\})\big(F_j \cap H_\ell \text{ contains a } v\text{–}x^* \text{ path}\big) \tag{4}$$

Both these are clear for $\ell = 1$. To prove (3) for $\ell = m > 1$, let $v \in U_{m-1}$ and $i \in \{1, \dots, k\}$ be given. Since (4) holds for $\ell = m-1$ by the induction hypothesis, there is a v–x^* path $Q \subseteq F_j \cap H_{m-1}$ for some j. Every edge uv of Q lies in $E_{\ell-1}$ for some $\ell \leqslant m$; then $\mathcal{P}_\ell$ contains the path uF_iv. Replacing the edges of Q with these paths we may turn Q into a v–x^* walk in $F_i \cap H_m$, which contains the desired v–x^* path.

To prove (4) for $\ell = m > 1$, let $v \in U_m$ be given. As (4) holds for all $\ell < m$ by the induction hypothesis, we may assume that $v \in U_m \smallsetminus U_{m-1}$. Then v lies on a path $P = xF_jy \subseteq H_m$ for some j and $x, y \in U_{m-1}$. By (3) for $\ell = m$ (as proved above), there is an x–x^* path P' in $F_j \cap H_m$. Then $vPxP'x^*$ is a walk in $F_j \cap H_m$ that contains the desired v–x^* path.

Finally, let us show that the graphs $F_i \cap H_n$ are all connected; then $U = U_n$ proves the lemma. Recall that $\mathcal{P}_{n+1} = \mathcal{P}_n$, and hence $H_{n+1} = H_n$. By (3) for $\ell = n+1$, the graphs $F_i \cap H_{n+1} = F_i \cap H_n$ have only one component, the component containing x^*. □

Proof of Theorem 2.4.1. We prove the backward implication by induction on $|G|$. For $|G| = 2$ the assertion holds. For the induction step we assume that for every partition P of V there are at least $k\,(|P|-1)$ cross-edges, and find k edge-disjoint spanning trees in G. (1.5.3)

Let us start with the k edge-disjoint spanning forests $F_1, \dots, F_k$ defined earlier, those whose total set of edges was maximal. If these are not all trees, then

$$\sum_{i=1}^{k} \|F_i\| < k\,(|G|-1)$$

by Corollary 1.5.3. On the other hand, we have $\|G\| \geqslant k\,(|G|-1)$ by assumption: consider the partition of V into single vertices. So there exists an edge $e^* \in E \setminus E(F_1 \cup \ldots \cup F_k)$. e^*

By Lemma 2.4.3, there exists a set $U \subseteq V$ that is connected in every F_i and contains the ends of e^*; in particular, $|U| \geqslant 2$. Since every partition of the contracted multigraph G/U induces a partition of G with the same cross-edges,[3] G/U has at least $k\,(|P|-1)$ cross-edges with respect to any partition P of its vertex set. By the induction hypothesis, therefore, G/U has k edge-disjoint spanning trees, $T_1, \dots, T_k$ say. Replacing in each T_i the vertex v_U contracted from U by the spanning tree $F_i[U]$ of $G[U]$, we obtain k edge-disjoint spanning trees in G. □ U

Let us say that subgraphs $G_1, \dots, G_k$ of a graph G *partition* G if their edge sets form a partition of $E(G)$. Our spanning tree problem may then be recast as follows: into how *many connected* spanning subgraphs can we partition a given graph? The excuse for rephrasing our simple tree problem in this more complicated way is that it now has an obvious dual (cf. Theorem 1.5.1): into how *few acyclic* (spanning) subgraphs can we partition a given graph? Or for given k: which graphs can be partitioned into at most k forests? *graph partition*

An obvious necessary condition now is that every set $U \subseteq V(G)$ induces at most $k\,(|U|-1)$ edges, no more than $|U|-1$ for each forest. Once more, this condition turns out to be sufficient too. And surprisingly, this can be shown with the help of Lemma 2.4.3, which was designed for the proof of our theorem on edge-disjoint spanning trees:

Theorem 2.4.4. (Nash-Williams 1964)
A multigraph $G = (V, E)$ can be partitioned into at most k forests if and only if $\|G[U]\| \leqslant k\,(|U|-1)$ for every non-empty set $U \subseteq V$.

Proof. The forward implication was shown above. Conversely, we show that every family $F_1, \dots, F_k$ of edge-disjoint spanning forests with a maximal total set of edges partitions G. If not, pick $e \in E \setminus E(F_1 \cup \ldots \cup F_k)$. (1.5.3)

[3] see Chapter 1.10 on contraction in multigraphs

By Lemma 2.4.3, there exists a set $U \subseteq V$ that is connected in every F_i and contains the ends of e. Then $G[U]$ contains $|U|-1$ edges from each F_i, and in addition the edge e. Thus $\|G[U]\| > k\,(|U|-1)$, contrary to our assumption. □

arboricity

The least number of forests forming a partition of a graph G is called the *arboricity* of G. By Theorem 2.4.4, the arboricity is a measure for the maximum local density: a graph has small arboricity if and only if it is 'nowhere dense', i.e. if and only if it has no subgraph H with $\varepsilon(H)$ large.

We shall meet Theorem 2.4.1 again in Chapter 8.5, where we prove its infinite version. 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*.

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.

path

ter($\mathcal{P}$)

path cover

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 $\mathrm{ter}(\mathcal{P})$ 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.

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 $\mathrm{ter}(\mathcal{P})$ 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 $\mathrm{ter}(\mathcal{P}) = \{v_1, \dots, v_m\}$ 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 $\mathrm{ter}(\mathcal{P})$ 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 $\mathrm{ter}(\mathcal{P}') = \{v, v_2, \dots, v_m\}$ is minimal among the sets of last vertices of path covers of G'. v $\mathcal{P}', G'$

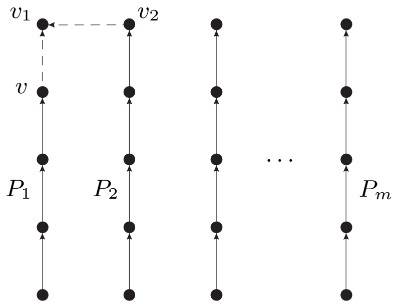

Fig. 2.5.1. Path covers of G and G'

Suppose then that G' has a path cover $\mathcal{P}''$ with $\mathrm{ter}(\mathcal{P}'') \subsetneq \mathrm{ter}(\mathcal{P}')$. 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 $\mathrm{ter}(\mathcal{P})$, 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 $\mathrm{ter}(\mathcal{P}'') \subseteq \{v_3, \dots, v_m\}$. But now $\mathcal{P}''$ and the trivial path $\{v_1\}$ together form a path cover of G that contradicts the minimality of $\mathrm{ter}(\mathcal{P})$. □

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. *chain* *antichain*

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. 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 $v_1 \dots v_n$ is the path constructed so far, the player to move next has to find a vertex v_{n+1} such that $v_1 \dots v_{n+1}$ 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.$^+$ Find an infinite counterexample to the statement of the marriage theorem.

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 *Sperner's theorem*: 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 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.

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

17. 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.

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

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

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

21.$^-$ 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.

22. 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.

23. 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.

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

25. 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.

26. 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 $H \subseteq G$.

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

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

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

30.$^-$ 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 $s+1$.

31. 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.

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

33. 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. Another very comprehensive source is A. Schrijver, *Combinatorial optimization*, Springer 2003. All the references for the results in this chapter can be found in these two books.

As we shall see in Chapter 3, König's Theorem of 1931 is no more than the bipartite case of a more general theorem due to Menger, of 1929. 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.

For background and applications of the stable marriage theorem, see D. Gusfield & R.W. Irving, *The Stable Marriage Problem: Structure and Algorithms*, MIT Press 1989, as well as A. Tamura, Transformation from arbitrary matchings to stable matchings, *J. Comb. Theory A* **62** (1993), 310–323.

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 (with directed cycles), which had long been conjectured but was only recently proved by B. Reed, N. Robertson, P.D. Seymour and R. Thomas, Packing directed circuits, *Combinatorica* **16** (1996), 535–554. Its proof is much more difficult than the undirected case; see Chapter 12.4, and in particular Corollary 12.4.10, for a glimpse of the techniques used.

Theorem 2.4.1 was proved independently by Nash-Williams and by Tutte; both papers are contained in *J. Lond. Math. Soc.* **36** (1961). Theorem 2.4.4 is due to C.St.J.A. Nash-Williams, Decompositions of finite graphs into forests, *J. Lond. Math. Soc.* **39** (1964), 12. Both results can be elegantly expressed and proved in the setting of matroids; see Schrijver's book.

An interesting vertex analogue of Corollary 2.4.2 is to ask which connectivity forces the existence of k spanning trees $T_1, \dots, T_k$, 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.