5.1 The Tutte-Berge Formula

In Chapter 3, we have discussed the matching problem in bipartite graphs. As it turns out matchings still have a nice structural properties in general graphs, but the extensions of theorems from bipartite graphs can be highly non-trivial. Recall that in an undirected graph $G = (V, E)$, a matching is a set of edges $M \subseteq E$ that are not incident to each other. We did denote $\nu(G)$ as the maximum cardinality of a matching in $G$ and $\tau(G)$ is the minimum size of a vertex cover. We want to begin with an extension of König’s Theorem. Suppose we have a general graph $G = (V, E)$. How could we certify that there is no perfect matching? A trivial reason would be if $|V|$ is odd. Surprisingly this idea can be extended to an exact min-max formula.

Let us call the connected component in a graph odd if it has an odd number of vertices. We define $\text{odd}(G)$ as the number of odd components in graph $G$.

Consider a matching $M$ in a graph $G$ and fix some subset $U \subseteq V$. Suppose that $C_1, \ldots, C_k$ with $k := \text{odd}(G \setminus U)$ are the odd components in $G \setminus U$. Then for each $i = 1, \ldots, k$, the matching $M$ either leaves a node in $C_i$ exposed, or it contains an edge between a node in $C_i$ and $U$.

Either way, $M$ must have $k - |U|$ many exposed nodes (just that this number does not need to be positive). We obtain that the number of covered nodes is

$$2|M| \leq |V| - (k - |U|)$$

and hence

$$|M| \leq \frac{1}{2} \cdot (|V| + |U| - k)$$
Quite surprisingly it turns out that there is always a set $U$ that provides a tight bound.

**Theorem 29 (Tutte-Berge Formula).** For every graph $G = (V, E)$ one has

$$
\nu(G) = \min_{U \subseteq V} \left\{ \frac{1}{2} (|V| + |U| - \text{odd}(G \setminus U)) \right\}
$$

Let us abbreviate $\text{exposed}(G) := \min \{ \# \text{ of } M\text{-exposed nodes} \mid M \text{ matching} \}$. Note that in any graph, $2\nu(G) + \text{exposed}(G) = |V|$. Let us call a node $v$ critical if it is covered by every maximum matching.

**Lemma 30.** Let $G = (V, E)$ be a connected graph with $|V| \geq 2$ and no critical node. Then $\text{exposed}(G) = 1$.

**Proof.** If $\text{exposed}(G) = 0$, then every node is critical. Hence suppose for the sake of contradiction that $\text{exposed}(G) \geq 2$. For two nodes $u, v \in V$, let $d(u, v)$ be the distance in $G$ (in terms of the number of edges; note that here we use connectedness). Fix a maximal matching that minimizes $d(u, v)$ for a pair of $M$-exposed nodes. If $d(u, v) = 1$, then $M \cup \{u, v\}$ is a bigger matching and we have a contradiction. Hence suppose that $d(u, v) \geq 2$. Select any node $t \in V$ with $d(u, t), d(v, t) < d(u, v)$ (for example by taking a node $t$ on the shortest path between $u$ and $v$).

Consider a maximal matching $N$ that leaves $t$ exposed. If there are several, choose the matching that maximizes the number $|M \cap N|$ of joint edges.

Next, consider the symmetric difference $M \Delta N$. Each of the nodes $u, v, t$ is exposed in either $M$ or $N$, so they are all endpoints of some paths in $M \Delta N$. Since $M$ and $N$ are maximal and we maximized $|M \cap N|$ we know that $M \Delta N$ consists only of even length paths. Consider the even length path $P$ containing $u$ as an endpoint. If the other endpoint is $t$, then $M \Delta P$ has the exposed nodes $v$ and $t$ which contradicts the choice of $M$. That means $P$ has $u$ as one endpoint and the other one is neither $v$ nor $t$. In fact, the situation looks like this:

$$
P : \quad u \quad \in N \quad \in M
\quad \quad v \quad \quad \quad \quad t
$$

Then the matching $N \Delta E(P)$ still has $t$ exposed but has more edges in common with $M$, which is a contradiction. \qed

An equivalent, but perhaps more intuitive form of the Tutte-Berge Formula is that in any graph $G = (V, E)$ one has

$$
\text{exposed}(G) = \max \{ \text{odd}(G \setminus U) - |U| \mid U \subseteq V \}
$$

**Proof of the Tutte-Berge Formula.** We already argued the direction “≤”. We prove the other direction by induction.
• **Case: G is not connected.** Let $G_1, \ldots, G_k$ be the connected components of $G$ (it does not matter whether these are even or odd) with $k \geq 2$. We apply induction to find $U_i \subseteq V(G_i)$ so that $\text{exposed}(G_i) = \text{odd}(G_i \setminus U_i) - |U_i|$. Then returning $U := \bigcup_{i=1}^{k} U_i$ will satisfy the claim as

$$\text{exposed}(G) = \sum_{i=1}^{k} \text{exposed}(G_i) = \sum_{i=1}^{k} (\text{odd}(G_i \setminus U_i) - |U_i|) = \text{odd}(G \setminus U) - |U|.$$

• **Case: There is a critical node $u \in V$.** Then $\nu(G \setminus u) = \nu(G) - 1$ and $\text{exposed}(G \setminus u) = \text{exposed}(G) + 1$. We apply induction to $G \setminus u$ and obtain a subset $U \subseteq V \setminus \{u\}$ with

$$\text{exposed}(G) + 1 = \text{exposed}(G \setminus u) = \text{odd}((V/u)/U) - |U| = \text{odd}(V \setminus (U \cup \{u\})) - |U \cup \{u\}| + 1$$

which means that $U \cup \{u\}$ satisfies the claim.

• **Case: G is connected and there is no critical node.** Then Lemma 30 applies and $\text{exposed}(G) = 1 = \text{odd}(G)$ using that $|V|$ has to be odd. Then $U = \emptyset$ satisfies the claim.

\[\square\]

**Corollary 31** (Tutte’s 1-factor theorem). A graph $G = (V, E)$ has a perfect matching if and only if $\text{odd}(G \setminus U) \leq |U|$ for all $U \subseteq V$.

**Proof.** Clear since there is a perfect matching if and only if $0 = \text{exposed}(G) = \max\{\text{odd}(G \setminus U) - |U| : U \subseteq V\}$.

\[\square\]

### 5.2 Cardinality matching algorithm

In this section, we want to design a polynomial time algorithm for the **cardinality matching problem**: given a graph $G = (V, E)$, find a matching $M \subseteq E$ maximizing $|M|$. In the special case of bipartite graphs, we saw that the concept of $M$-augmenting paths is the right one to solve the problem. Unfortunately, $M$-augmenting paths are harder to find in general graphs. Recall that a **walk** in a graph $G = (V, E)$ is a sequence $v_0, v_1, \ldots, v_t$ so that $\{v_i, v_{i+1}\} \in E$ for all $i = 0, \ldots, t - 1$. In particular in a walk one is allowed to revisit nodes and edges. A **path** is a walk where all visited nodes are distinct. We say that a walk $(v_0, v_1, \ldots, v_t)$ is **$M$-alternating** if for each node $v_i$ with $i \in \{1, \ldots, t - 1\}$ exactly one of the edges $\{v_{i-1}, v_i\}, \{v_i, v_{i+1}\}$ lies in $M$ and the other one not. As the name suggests, edges on the walk are alternatingly in $M$ and not in $M$. We have proven in a previous chapter that a matching is of maximum cardinality if and only if there is no $M$-augmenting path. In particular an $M$-augmenting path is also an $M$-alternating walk. For a subset $W \subseteq V$, we call a walk a **$W$-$W$ walk** if start and endpoint are in $W$.

**Lemma 32.** Let $G = (V, E)$ be a graph and $M \subseteq E$ be a matching leaving $W \subseteq V$ exposed. Then a shortest $M$-alternating $W$-$W$ walk $P$ can be found in time $O(|V| + |E|)$.

**Proof.** We define a directed graph $D = (\{s\} \cup V' \cup V'', A)$ that has nodes $v'$ and $v''$ for every original node $v$ in $G$. We insert an edge $(u', v'') \in A$ if $\{u, v\} \in E \setminus M$. We insert $(u'', v') \in A$ if $\{u, v\} \in M$. Moreover, we insert arcs $(s, v')$ for every $M$-exposed node $v$.

25
Then in linear time one can compute the set of nodes reachable from \( s \) using breadth first search. This also determines the shortest paths from \( s \).

Unfortunately, an \( M \)-alternating path between \( M \)-exposed nodes is not necessarily an \( M \)-augmenting path:

But if we are not finding an \( M \)-augmenting path, we can at least find a different structure: An \( M \)-blossom is an \( M \)-alternating walk \( v_0, \ldots, v_t \) in \( G \) with \( t \geq 3 \) so that (i) \( v_0, \ldots, v_{t-1} \) are distinct; (ii) \( v_0 = v_t \); (iii) and \( v_0 \) is \( M \)-exposed.

Actually, the example figure above does not contain an \( M \)-blossom. But \( M' := M \Delta (v_0, v_1, v_2) \) is a matching with \( |M'| = |M| \) and the graph contains an \( M' \)-blossom.

**Lemma 33.** Given a graph \( G = (V, E) \) and a matching \( M \), in time \( O(|V| + |E|) \) one can obtain one of the following:

1. Decide there is no alternating \( W \)-\( W \) walk where \( W \) are the \( M \)-exposed nodes.
2. Find an \( M \)-augmenting path.
3. Find a matching \( M' \subseteq E \) of size \( |M'| = |M| \) and an \( M' \)-blossom.

**Proof.** Use the auxiliary directed graph \( D = (\{s\} \cup V' \cup V'', A) \) from Lemma 32 to find a shortest \( M \)-alternating \( W \)-\( W \) walk. Note that for each \( u \in V \), the shortest walk is going to visit each of the copies \( v', v'' \) at most once. If such a walk does not exist, we are in (1). If the walk is also a path, we are in (2). Otherwise, let \( v_0, \ldots, v_k \) be the beginning of the walk until we revisit the first time a node in \( v_k = v_t \) with \( 0 \leq t < k \).
Then this part of the walk has 3 edges incident to \( v_k \). Exactly one of the edges \( \{v_{t-1}, v_t\}, \{v_t, v_{t-1}\} \) has to be in \( M \), the other one not. That implies \( \{v_{k-1}, v_k\} \notin M \). If \( \{v_t, v_{t+1}\} \in M \), then \( v_t \) would have been entered twice with a non-matching edge, which means that the node \( v''_t \) in \( D \) was visited twice, which is impossible. Hence \( \{v_{t-1}, v_t\} \in M \). That means the figure above is accurate and \( \{v_t, v_{t+1}, \ldots, v_k\} \) is an \( M' \)-flower where \( M' := M \Delta (v_0, \ldots, v_t) \).

It remains to be proven why finding such a blossom is useful. For a graph \( G = (V, E) \) and a subset \( C \subseteq V \) we define \( G \setminus C = ((V \setminus C) \cup \{v_C\}, E') \) as the contraction where every appearance of a node in \( C \) is replaced by \( v_C \).

**Lemma 34.** Let \( C = (v_0, \ldots, v_t) \) be an \( M \)-blossom in \( G \). Then \( M \) has maximum size in \( G \) if and only if \( M/C \) has maximal size in \( G \setminus C \).

**Proof.** We will prove the equivalent negated statement

\[
M \text{ not of maximal size in } G \iff M/C \text{ not of maximal size in } G/C
\]

"\( \Rightarrow \)". If \( M \) is not of maximal size, then there is an \( M \)-augmenting path \( P = (u_0, \ldots, u_m) \). We may assume that \( P \) contains at least one node from \( C \), otherwise there is nothing to show. Let \( k \) be the first index with \( u_k \in C \). We may assume that \( u_0 \neq v_0 \) and \( \{u_{k-1}, u_k\} \notin E \) since not both endpoints of \( P \) can be in \( C \) and \( C \) is entered at most once with a matching edge (and in that case \( v_0 \) is not an endpoint of \( P \)). Then \( u_0, \ldots, u_k \) is an \( M/C \)-augmenting path in \( G/C \).
“⇐”. Suppose $M/C$ is not of maximal size. Let $P$ be an $M/C$ augmenting path in $G/C$. Note that $C$ is $M/C$-exposed, hence we may assume $C$ is an endpoint (or there is nothing to show). Then the path corresponds to $u_0, \ldots, u_k = v_i, v_i \pm 1, \ldots, v_0$ with $i \in \{0, \ldots, t\}$ is an $M$-augmenting path in $G$. Here we go either clockwise or counterclockwise, taking the incident matching edge (the figure from above applies again).

Finally, this gives us an algorithm to find maximum cardinality matchings.

<table>
<thead>
<tr>
<th>Edmonds Maximum Cardinality Matching Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Set $M := \emptyset$</td>
</tr>
<tr>
<td>(2) REPEAT</td>
</tr>
<tr>
<td>(3) Call AugmentMatching($G, M$) (either replacing $M$ by larger matching or terminating).</td>
</tr>
</tbody>
</table>

AugmentMatching($G, M$) :

(1) Compute a shortest $M$-alternating $W$-$W$ walk $P$ with $W := \{v \in V \mid v$ is $M$-exposed$\}$

(2) If there is none, return “$M$ is maximal”

(3) If $P$ is path RETURN $M \Delta P$

(4) Otherwise, extract $M'$ with $|M| = |M'|$ and $M'$-blossom $C$

(5) Call AugmentMatching($G/C, M'/C$). If returns $M'/C$ is maximal, then RETURN $M$ is maximal

(6) Otherwise, construct matching $M''$ in $G$ with $|M''| > |M|$ as in Lemma 34

5.3 Exercises

Exercise 5.1. Let $G = (V, E)$ be a graph and $M$ be a matching. Suppose that there is no $M$-augmenting path of length less than $2k + 1$. Show that $\nu(G) \leq (1 + \frac{1}{k}) \cdot |M|$. 

Exercise 5.2. Consider a graph $G = (V, E)$. Let $P := \text{conv}\{\chi^M \in \mathbb{R}^E \mid M \subseteq E \text{ is matching}\}$ be the convex hull of characteristic vectors of all matchings in $G$ and let $P_{=k} = \text{conv}\{\chi^M \mid M \subseteq E \text{ matching with } |M| = k\}$ be the convex hull for all matchings with $k$ edges. Prove that $P \cap \{x \in \mathbb{R}^E \mid \sum_{e \in E} x_e = k\} = P_{=k}$.

Hint. It may be useful to observe that $\chi^{M_1} + \chi^{M_2} = \chi^{M_1 \Delta P} + \chi^{M_2 \Delta P}$ where $M_1, M_2$ are matchings and $P$ is a component of $M_1 \Delta M_2$.

Exercise 5.3. Let $G = (V, E)$ be a graph. Define

$$\mathcal{I} := \{U \subseteq V \mid \exists \text{ matching } M \subseteq E \text{ with } U \subseteq V(M)\}$$

(where $V(M)$ are the nodes covered by $M$). Prove that $(V, \mathcal{I})$ is a matroid.