CS 491G Combinatorial Optimization
Lecture Notes

David Owen

July 30, August 1

1 Matchings

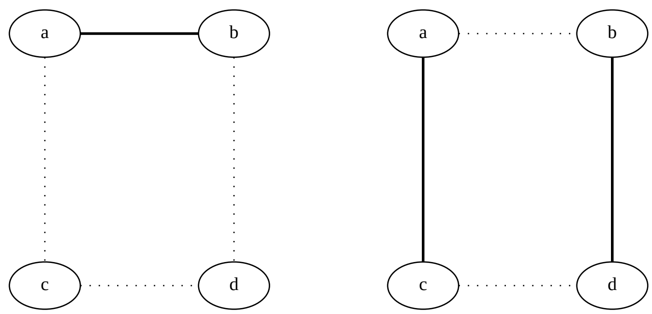

Figure 1: two possible matchings in a simple graph.

Definition 1 *Given a graph $G = \langle V, E\rangle$, a* **matching** *is a collection of edges M such that $e_i, e_j \in M \Longrightarrow e_i, e_j$ are vertex disjoint.*

Definition 2 *A* **maximal matching** *is a matching that cannot be improved.*

By "vertex disjoint," we mean that no two edges in the set may be incident with the same vertex. Figure 1 shows two possible matchings in a simple graph. A matching is "improved" when one or more edges is added to it.

Definition 3 *A* **maximum matching** *is a matching of maximum cardinality (over all possible matchings in a graph G); $\nu(G)$ denotes the cardinality of a maximum matching in G.*

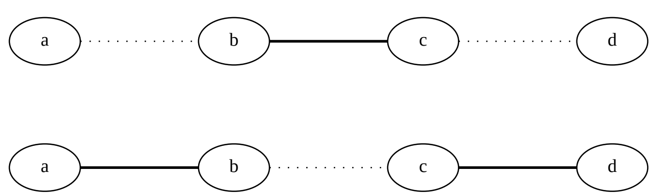

Figure 2: a maximal but not maximum matching; a maximum matching.

A maximal matching is not necessarily a maximum matching. Figure 2 shows a maximal matching (that is not a maximum matching) and a maximum matching. All maximum matchings are maximal matchings, but not all maximal matchings are maximum matchings. In general, it is harder to find a maximum matching than to find a maximal matching.

Definition 4 *Given a matching M, an M-***exposed** *vertex is a vertex not incident with any edge in M; an M-***covered** *vertex is a vertex incident with some edge in M.*

Given a maximum matching M the number of M-covered vertices $= 2 \cdot \nu(G)$; the number of M-exposed vertices $= |V| - 2 \cdot \nu(G)$. The number of M-exposed vertices is also called the "deficiency," or $def(G)$. In other words, $def(G)$ is the minimum number of nodes that will be exposed in any matching.

Definition 5 *Given a graph $G = \langle V, E\rangle$, a* **perfect matching** *is a matching with deficiency $def(G) = |V| - 2 \cdot \nu(G) = 0$.*

Interesting problems concerning matchings include: does G have a large matching? does G have a perfect matching? what is a maximum matching of G? etc. Up to now we have considered only unweighted graphs. If we consider weighted graphs and sum the weights of edges included in matchings, we might also try to find the minimum-weight perfect matching of a particular graph (if one exists).

2 Techniques for Finding a Maximum Matching

Definition 6 *Given a matching M in a graph G, a path P composed of edges that alternately belong to and do not belong to M is called an M-***alternating path.**

Definition 7 *An M-alternating path P is an M-***augmenting path** *if the first and last vertices are M-exposed.*

Figure 3 shows a simple graph with a matching $M = \{e_{ab}\}$. The path from c to a to b to d is an M-alternating path, because the edges in the path alternately belong to the matching ($e_{ca} \notin M$, $e_{ab} \in M$, $e_{bd} \notin M$). The same path is also an M-augmenting path because its endpoints, c and d are M-exposed; that is, they are not incident with any edge in M.

Figure 3: a matching $M = \{e_{ab}\}$.

Theorem 1 *A matching M in a graph $G = \langle V, E\rangle$ is maximum if and only if there is no M-augmenting path.*

In the textbook, Theorem 1 is called the "Augmenting Path Theorem of Matchings" [CO98]. For the proof of Theorem 1, we need the set operation "symmetric difference." We define the symmetric difference between two sets S and T as $S \triangle T = \{$ *elements* $\in S$ *or* T *but not both* $\}$ (this is like the "exclusive or" operation $\oplus$).

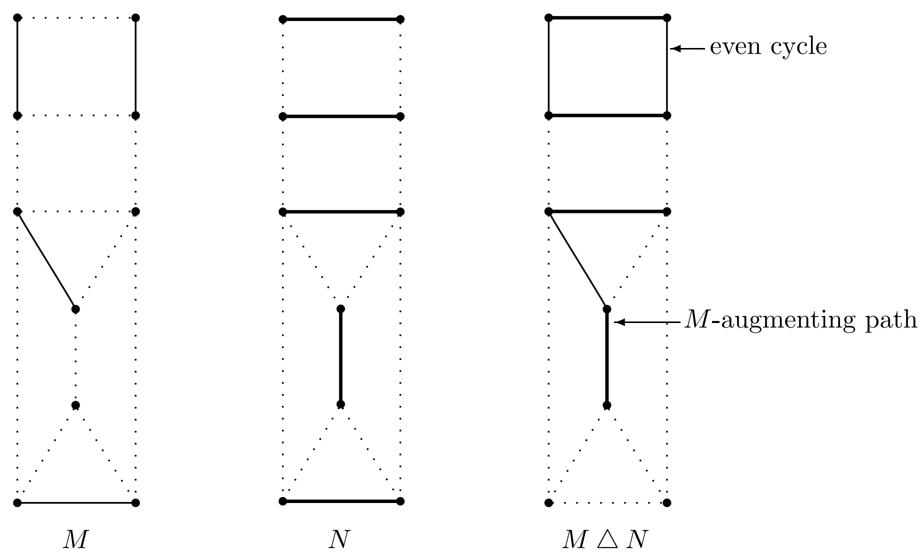

Figure 4: matchings M, N, and their symmetric difference $M \triangle N$, which includes an even cycle and an M-augmenting path.

To prove Theorem 1, we first show that if there is an M-augmenting path then M is not a maximum matching. Let P be an M-augmenting path. Consider the set $S = M \triangle E(P)$, which is the set of all elements from either the matching M or the path P but not both. Since the set of S-covered vertices includes all M-covered vertices, while also including the two $M exposed$ vertices, S improves on M; therefore M is not a maximum matching.

We now show that if M is not a maximum matching then there must be an M-augmenting path. Since M is not maximum, let N be another matching such that $|N| > |M|$; that is, N improves on M. Consider $S = M \triangle N$. Because M and N are matchings, vertices in the graph G may be incident to at most one edge from M and at most one edge from N. Therefore every vertex in the graph G is incident to at most two edges in S.

From this we can infer structural characteristics of S; specifically, S contains only even cycles and M-alternating paths. Edges in the cycles alternate between edges from N and edges from M. Since the cycles are even, they contain the same number of edges from N as from M. But $|N| > |M|$, so we know there must be at least one M-alternating path with more edges from N than from M. This path would begin and end with edges from N and would have M-exposed endpoints; therefore it would be an M-augmenting path. Figure 4 shows two matchings, M and N, for a graph. The symmetric difference $M \triangle N$ shown includes an even cycle with alternating edges (the bold edges come from N) and an M-augmenting path.

Theorem 1 gives us a general idea for designing an algorithm to find a maximum matching: as long as there is an M-augmenting path, we know that M is not a maximum matching.

3 Odd cycles and the Tutte-Berge Formula

Definition 8 *Given a graph $G = \langle V, E \rangle$, a* **cover** *is a set of vertices $A \subseteq V$ such that for every edge $e = vw \in E$, either $v \in A$ or $w \in A$.*

Suppose A is a cover of some graph for which we have a matching M. We know that all edges in M must have at least one endpoint in A, because all edges in the graph have at least one endpoint in A (A is a cover). Therefore the maximum size of the matching cannot exceed $|A|$ ($|M| \leq |A|$). In a bipartite graph a maximum matching and a minimum cover will be equal in size (there is a primal-dual relationship between the two problems when formulated as linear programs).

Graphs that are not bipartite may have odd cycles. Figure 5 shows a graph of an odd cycle. Clearly, the size of a maximum matching will be strictly less than the size of a minimum cover for odd cycles. Also, while

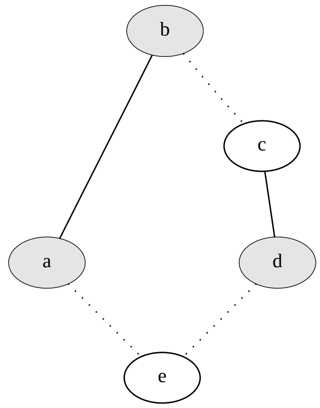

Figure 5: an odd cycle; darkened edges $\in$ matching M and gray vertices $\in$ cover A; $|M|$ (strictly) $< |A|$.

even cycles admit only two possible maximum matchings, odd cycles of size > 3 will have more (*many* more, as the size of the cycle increases). So it is odd cycles that cause a problem—we need to find a technique of handling odd cycles.

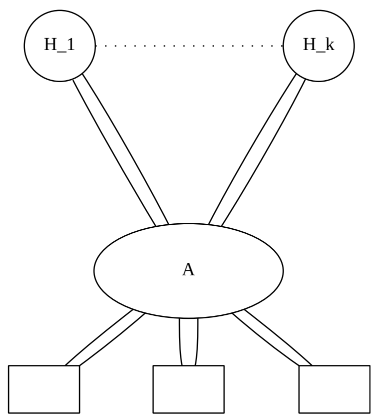

Figure 6: $H_1 \dots H_k$ are k odd components extracted from a graph G; A is the part of the graph left after the H's have been extracted.

We divide a graph G into sections as shown in Figure 6 ($G \setminus A$ is divided into odd components $H_i \dots H_k$). Suppose M is a matching from G. For each H_i, if there are no M-exposed vertices $\in H_i$, there must be an edge of M with one end in H_i and the other end in A. So the total number of M-exposed vertices in the graph will be at least $k - |A|$. Let the number of odd components (k until now) taken from $G \setminus A = oc(G \setminus A)$. So the number of M-exposed vertices is $\geq oc(G \setminus A) - |A|$. From this we conclude that, for any $A \in V$:

$$\nu(G) \leq \frac{1}{2}(|V| - oc(G \setminus A) + |A|)$$

That is, the size of a maximum matching is at most half of the total number of vertices minus the number of M-exposed vertices. If the set A chosen is a cover of G, there will be $|V| - |A|$ odd components in $G \setminus A$ (each one will be a single node). In that case the right side of the inequality will reduce to $|A|$, which gives us the same bound stated above, that the size of the maximum matching is at most the size of the minimum cover. This means our new bound is at least as good as the cover bound above. For the odd-cycle graph shown in Figure 5, the size of a maximum matching $= 2$. We choose the smallest possible A, $A = \emptyset$, so that $|V| = 5$, $|A| = 0$, $G \setminus A = G$, and (therefore) $oc(G \setminus A) = 1$; for this case we have:

$$\nu(G) \leq \frac{1}{2}(5 - 1 + 0)$$
$$2 \leq 2$$

In this case the inequality is strictly equal. In fact, the inequality will be strictly equal every time a maximum matching and minimum right side are used—this is the Tutte-Berge Formula, stated more exactly below in Theorem 2.

Theorem 2 *(Tutte-Berge Formula) given a graph $G = \langle V, E\rangle$,*

$$\max\{\, |M| : M \textit{ is a matching } \} = \min\{\, \tfrac{1}{2}(|V| - oc(G \setminus A) + |A|) : A \subseteq V \,\}.$$

Theorem 2 could be used to find the size of a maximum matching and thereby a stopping condition for some maximum matching algorithm. We could simply try every possible A and take the minimum right side. But there will be 2^n A's, so this would be very inefficient.

Tutte's original proof of Theorem 2 was extremely complex. Here we will present some of the background for Berge's simpler proof, published years after Tutte's. Theorem 3 follows from the Tutte-Berge Formula above (Theorem 2).

Theorem 3 *A graph $G = \langle V, E\rangle$ has a perfect matching if and only if*

$$\forall A \in V,\ oc(G \setminus A) \leq |A|$$

The main idea of Berge's proof involves a "shrinking" operation done on odd cycles, in which the odd cycle becomes a single vertex. Formally, we let C be an odd cycle in a graph G. We define $G' = G \times C$ as the graph formed from G after shrinking the odd cycle C. The set of vertices in $G' = (V \setminus V(C)) \cup \{C\}$, where $V(C)$ is the set of vertices in the cycle C, and the new vertex created by shrinking the cycle C is called C also. The set of edges in $G' = E \setminus \gamma(V(C))$, where $\gamma(V(C))$ is the set of edges in the odd cycle C. For each edge $e \in G'$, with vertex v one of e's ends, if $v \in V(C)$ then vertex C replaces v as the end of e in G'. All other edges remain the same in G' as they were in G. This process can be thought of as a series of "node identifications," as described on page 73 of the text [CO98] in reference to the *Minimum Cut Problem* for undirected graphs. An illustration of the "shrinking" operation appears on page 131 of the text [CO98].

Theorem 4 *Let C be an odd circuit of a graph G, let $G' = G \times C$, and let M' be a matching of G'. There is a matching M of G such that $M \subseteq M' \cup E(C)$, and the number of M-exposed vertices of G is the same as the number of M'-exposed vertices of G'.*

Theorem 4 says that any matching on the reduced graph $G \times C$ can be extended into a matching in the original graph G.

For the proof of Theorem 4, we let M' = some matching in the graph $G' = G \times C$. Either C is M'-covered or M'-exposed in the graph G'. We go backwards from G' to modify the original graph G in the following way: if C is M'-covered in G', we remove from G the vertex through which C was M'-covered; if C is M'-exposed in G', we remove any (arbitrary) vertex from the odd cycle in G. So what we have done, in either case, is to remove two edges from an odd cycle. If two edges are removed from an odd cycle, the result is a odd-length path, which must have a perfect matching (choose the odd edges of the path: $1, 3, 5 \ldots$); let M'' be this perfect matching. The union of the matching M' from G' and this M'' gives us a matching matching M in the original graph G ($M' \cup M'' = M$), satisfying the properties of the theorem. Since M'' is a perfect mathing of an odd-length path, it contributes no M''-exposed nodes. Therefore the number of M'-exposed nodes must be equal to the number of M-exposed nodes.

The size of a maximum matching from a graph G is at least the sum of the size of the maximum matching in $G \times C$ (after shrinking an odd cycle C) plus the number of edges in a maximum matching of the odd cycle C (which will always be $\frac{1}{2}(|V(C)| - 1)$), or:

$$\nu(G) \geq \nu(G \times C) + \frac{1}{2}(|V(C)| - 1)$$

When this relationship holds with equality, it means there is a maximum matching in G that uses all $\frac{1}{2}(|V(C)|-1)$ edges (the edges in a maximum matching of the odd cycle C). In that case we say the odd cycle C is "tight," and the problem of finding a maximum matching in G reduces to the problem of finding a maximum matching in $G \times C$. Figure 7 below shows a graph with an odd cycle that is not tight. The maximum matching size $\nu(G)$ for the graph in Figure 7 is 3 (the maximum matching is the three edges that are not bold). But if we shrink the odd cycle, we have a graph in which the maximum matching size $\nu(G \times C)$ is only 1. Since the size of the odd cycle $|V(C)| = 3$ (if we substitute into the equation above), we have:

$$3 \geq 1 + \tfrac{1}{2}(3-1)$$
$$3 \geq 2$$

Since the relationship doesn't hold with equality, the odd cycle is not tight. It turns out that checking for "tightness" of odd cycles is a very difficult problem, but the concept of tightness is important for the proof of the Tutte-Berge Formula (Theorem 2).

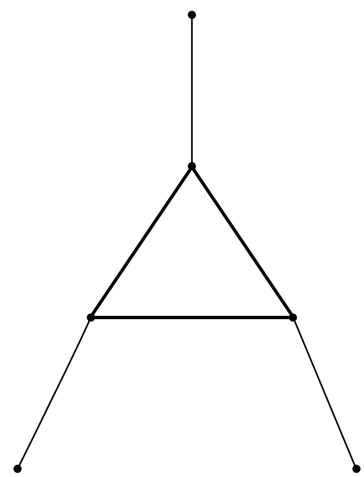

Figure 7: the odd cycle C in this graph (shown in bold) is not tight.

Definition 9 *Given a graph $G = \langle V, E\rangle$, a vertex $v \in V$ is* **essential** *if every maximum matching of G must include v; v is* **inessential** *if v is not essential.*

We return to the inequality shown below:

$$\nu(G) \leq \frac{1}{2}(|V| - oc(G \setminus A) + |A|) \tag{1}$$

Suppose we choose a set A for which the relationship above holds with equality. If we kick out a vertex $v \in A$ from G to get some modified graph G', what is the number of odd components $oc(G' \setminus A \setminus \{v\})$? It is the same number of odd components as there are in $G \setminus A$, or $oc(G \setminus A)$. Therefore $\nu(G') < \nu(G)$, so v is essential—as are all the vertices v' from the set A (A chosen so that the relationship (1) holds with equality).

Lemma 1 *Given a graph $G = \langle V, E\rangle$, let edge $vw \in E$. If both v and w are inessential, then there exists a tight odd cycle C that includes vw, and C is an inessential vertex of $G \times C$.*

Proof: *Since v and w are both inessential, there exist maximum matchings M_1 and M_2 under which v and w are exposed respectively. Note that M_1 must cover w and M_2 must cover v; otherwise the edge vw can be added to M_1 (or M_2) to increase the size of that matching. Consider the subgraph $H =< V, M_1 \triangle M_2 >$. The component containing v consists of a path P originating from v; since v is M_1-exposed. If P ends at another M_1-exposed node, then, we have an M_1-augmenting path, contradicting the maximum cardinality of M_1. Likewise, if it ends at an M_2-exposed node other than w, then adding the edge vw to this path, we get an M_2-exposed node. Thus, P must end at w; it thus forms an circuit C with edge vw. Note that P starts with an edge $vx \in M_2$ and ends with an edge $yw \in M_1$, for some vertices $x, y \in V$. Thus the circuit formed with edge vw is necessarily odd. Further, M_1 is a maximum matching of G, that contains $\frac{1}{2}(|V(C)| - 1)$ edges from $E(C)$ (delete vertex v from C!) and thus C is tight. Finally, $M_1 \setminus E(C)$ is a maximum matching of $G \times C$ not covering C (check?!), proving that C is an inessential node of $G \times C$.* $\square$

Proof: *(of Tutte-Berge Formula): Let us restate the formula:*

$$\max\{\, |M| : M \textit{ is a matching } \} = \min\{\, \frac{1}{2}(|V| - oc(G \setminus A) + |A|) : A \subseteq V \,\} \tag{2}$$

We have already shown that the LHS of (2) is less than or equal to the RHS for all $A \subseteq V$. Hence, all that is required is to produce a matching M and a set A, such that the number of $M-$exposed nodes is exactly $oc(G \setminus A) - |A|$. We provide an inductive proof. The result is certainly true, if G has no edges. Simply choose $A = \phi$; the maximum matching has cardinality zero and so is $|V| - oc(G \setminus A) + |A|$, proving that the LHS and RHS of (2) are equal.

Now consider a graph having at least one edge, say vw. We need to consider two cases:

1. *Let v be an essential node, i.e. $\nu(G \setminus v) = \nu(g) - 1$. Set $G' = G \setminus v$. Using the inductive hypothesis on G', we note that there exists a set A' and a matching M' of G' such that there are exactly $l_1 = oc(G' \setminus A') - |A'|$ $M'-$exposed nodes. Set $A = A' \cup \{v\}$. From the previous statement and the fact that v is an essential node, we know that there must exist a matching M of G, such that the number of $M-$exposed nodes is $l_1 + 1 = oc(G \setminus A' \cup \{v\}) - |A' \cup \{v\}|$. Setting $A = A' \cup \{v\}$, we are done. The proof for the case in which w is an essential node is identical to the above proof.*

2. *We now consider the case in which both v and w are both inessential. From Lemma (1), we know that G has a tight circuit C. We apply induction to $G' = G \times C$ to get a matching M' and a set A' of nodes, such that the number of $M'-$exposed nodes is exactly $oc(G' \setminus A') - |A'|$. The node C of G' cannot be in A', since it is an inessential node and if A' is a subset for which (2) is met with equality, then all nodes are essential. We also know from Theorem (4), that M' can be extended into a matching M of G, having the same number of exposed nodes. Observe that deleting A' from G results in the same number of odd components as deleting A' from G'. If C belongs to an odd component of $G' \setminus A'$, then it belongs to the same component in G; however it is replaced by $|V(C)|$ nodes. Since $|V(C)|$ is odd, the component in G is also odd. The same argument works when C does not belong to an odd component of $G' \setminus A'$. Thus, we get a matching M in G and a set A', such that the number of $M'-$exposed nodes is exactly $oc(G \setminus A') - |A'|$ as required. Note that in this case, the set A' of $G \times C$ is used as the set A of G.*

$\square$

References

[CO98] William Cook, William H. Cunningham, William Pulleyblank, and Alexander Schrijver. *Combinatorial Optimization.* John Wiley & Sons, 1998.