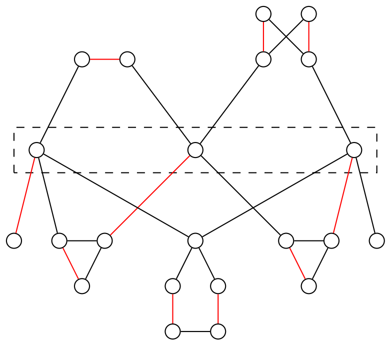
In this lecture, we take a look at the structure of maximum matchings in non-bipartite graphs and try to find a certain decomposition that helps constructing *all* maximum matchings.

1 Edmonds-Gallai Decomposition

Consider some graph $G = (V, E)$. We let $\nu(G)$ denote the cardinality of a maximum matching in G. For some vertex set $U \subseteq V$, let $G - U$ denote the subgraph of G obtained by deleting the vertices of U and all edges incident with them. Let $o(G - U)$ denote the number of connected components of $G - U$ that contain an odd number of vertices. Let M be a matching in $G - U$ and consider a component of $G - U$ with an odd number of vertices. There must be at least one unmatched vertex v in this component, since any matching necessarily covers an even number of vertices. Treating M as a matching in G, it is possible that we could increase the size of M by matching v with some vertex in U. However, we can add at most $|U|$ edges to M in this manner, since the vertices in U will eventually all be matched. Thus, any matching in G must have at least $o(G-U)-|U|$ unmatched vertices. This shows that the maximum size of a matching is upper-bounded by $(|V|+|U|-o(G-U))/2$, for any subset U. The figure illustrates this. Here, $|U| = 3$ and $|o(G - U)| = 5$.



The following theorem strengthens this result.

Theorem 1 (Tutte-Berge Formula)**.** *Let $G = (V, E)$ be a graph. Then*

$$\nu(G) = \max_{M} |M| = \min_{U \subset V} |V| + |U| - o(G - U))/2 ,$$

where the maximization is over all matching M in G.

Proof. We will consider the case that G is connected. Otherwise, the result follows by adding the formulas for the individual connected components. The proof proceeds by induction on the number of vertices in G. If G has at most one vertex then the result holds trivially. Otherwise, suppose that G has at least two vertices. We consider two cases:

Case 1: G contains a vertex v that is covered by *all* maximum matchings.

The subgraph $G - v$ cannot have a matching of size $\nu(G)$, otherwise that would give a maximum matching for G that leaves v unmatched. Thus $\nu(G - v) = \nu(G) - 1$. By induction, the result holds for the graph $G - v$, so there exists a set $U' \subset V - v$ that achieves equality in the Tutte-Berge Formula. Defining $U = U' \cup \{v\}$, we see that

$$\begin{aligned}
\nu(G) &= \nu(G - v) + 1 \\
&= \Big(|V - v| + |U'| - o(G - v - U')\Big)/2 + 1 \\
&= \Big((|V| - 1) + (|U| - 1) - o(G - U)\Big)/2 + 1 \\
&= \Big(|V| + |U| - o(G - U)\Big)/2 .
\end{aligned}$$

Case 2: For every vertex $v \in G$, there exists a maximum matching that does not cover v.

We will prove that each maximum matching leaves exactly one vertex uncovered. Suppose the contrary, that is, each maximum matching leaves at least two vertices uncovered. We choose a maximum matching M and its two uncovered vertices u and v such that we minimize $d(u, v)$, the distance between vertices u and v. If $d(u, v) = 1$, then the edge uv can be added to M to obtain a larger matching, a contradiction.

Otherwise, $d(u, v) \geq 2$ so we may fix an intermediate vertex t on some shortest u-v path. By the assumption of the present case, there is a maximum matching N that does not cover t. Furthermore, we may choose N such that its symmetric difference with M is minimal. If N does not cover u, then (N, u, t) contradicts our choice of (M, u, v). Thus, N covers u and, by symmetry, also v. Since N and M both leave at least two vertices uncovered, there exists a second vertex $x \neq t$ that is covered by M but not by N. Let xy be the edge in M that is incident with x. If y is also uncovered by N then $N \cup xy$ is a larger matching than N, a contradiction. So let yz be the edge in N that is incident with y, and note that $z \neq x$. Then $N \cup \{xy\} \setminus \{yz\}$ is a maximum matching that does not cover t and has smaller symmetric difference with M than N does. This contradicts our choice of N, a contradiction.

Hence, each maximum matching must leave exactly one vertex uncovered. Then $\nu(G) = (|V| - 1)/2$. The Tutte-Berge Formula then follows by choosing $U = \emptyset$. □

Now that we have established the correctness of the Tutte-Berge Formula, one might ask how we can compute such a set $U \subseteq V$ giving equality in the formula?

Such a set is provided by the Edmonds-Gallai Decomposition of G. This decomposition partitions $V(G)$ intro three sets: $D(G)$ is the set of all vertices v such that there is some maximum matching that leaves v uncovered, $A(G)$ is the neighbor set of $D(G)$, and $C(G)$ is the set of all remaining vertices.

Theorem 2 (Edmonds-Gallai Decomposition)**.** *Given a graph G, let*

$$D(G) := \{v : \textit{there exists a maximum size matching missing } v\} ,$$
$$A(G) := \{v : v \textit{is a neighbor of some } u \in D(G) \textit{ but } v \notin D(G)\} ,$$
$$C(G) := V(G) \setminus (D(G) \cup A(G)) .$$

Then:

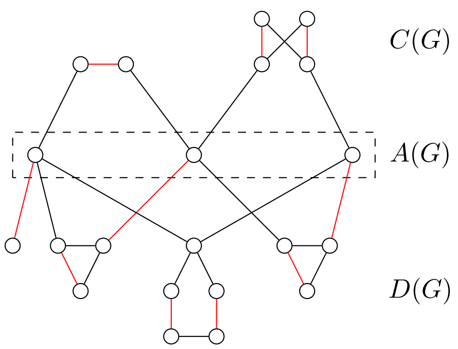
(i) $U = A(G)$ achieves the minimum on the right side of the Tutte-Berge formula,

(ii) $C(G)$ is the union of the even-sized connected components of $G \setminus A(G)$,

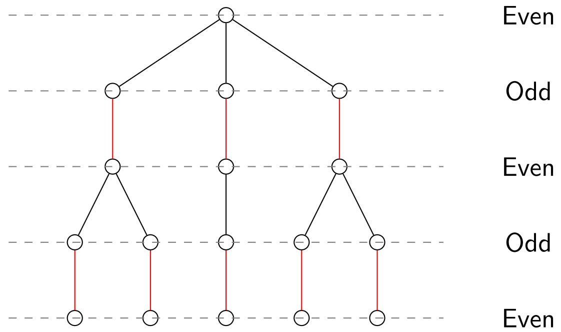
(iii) $D(G)$ is the union of the odd-sized components of $G \setminus A(G)$,

(iv) Every odd-sized component of $G \setminus A(G)$ is factor-critical. (A graph H is factor-critical if for every vertex v, there is a matching in H whose only unmatched vertex is v.)

Let $G[D(G)]$ be the subgraph of G induced by $D(G)$. The last condition says that every connected component H of $G[D(G)]$ is not only of odd cardinality but we can actually choose in it any particular vertex to be left uncovered. The Edmonds-Gallai Decomposition can be found as a byproduct of Edmonds' Matching Algorithm of the previous section.



Let us consider the maximum-size matching M that is returned by Edmonds' Matching Algorithm in any graph G. Let X be the set of vertices not matched by M. Consider all the vertices that can be reached by an alternating path from $x \in X$. The first edge on such a path must lie outside of M, the second edge must lie in M, and so on, leading to the following Figure.



Motivated by this, we define the following three subsets of $V(G)$:

$$\mathsf{Even} := \{v : \exists \text{ an alternating path of even length from } X \text{ to } v\} \quad ,$$

$$\mathsf{Odd} := \{v : \exists \text{ an alternating path from } X \text{ to } v\} \setminus \mathsf{Even} \quad ,$$

$$\mathsf{Free} := \{v : \nexists \text{ an alternating path from } X \text{ to } v\} \quad .$$

We will sometimes refer to a vertex as being even, odd, or free, according to which of these sets it belongs to. Note that in the above definitions, we are interested in alternating *paths*, i.e., alternating walks in which all the vertices are distinct.

We start with the following claim.

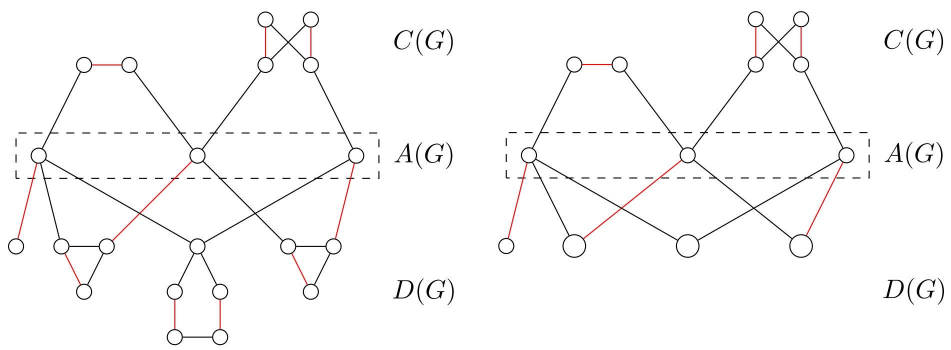
Claim 1. *If there is an edge from a vertex $u \in$ Even to some v, then there is an alternating walk of odd length from X to v, and there is an alternating path from X to v.*

Proof. If $e = uv$ is the edge in question, and P is an alternating path of even length from X to u, then an alternating walk of odd length from X to v is constructed as follows. If $e \in M$, then we take P and delete the final edge, which is necessarily e. If $e \notin M$, then we append e to P. If this alternating walk is not a path, it can only be because v lies on P, in which case P contains a sub-path which is an alternating path from X to v. □

As a result, any vertex adjacent to a vertex from Even has to belong to either Even or Odd. This gives the following corollary.

Corollary 1. *In G there is no edge between* Even *and* Free.

Let us now define the *shrunk graph* G_k to be the graph obtained in the final iteration of the execution of Edmonds' algorithm on G. More precisely, G_k is the final graph obtained from G by repeated shrinking of blossoms performed during the course of the algorithm. Let M_k be the maximum size matching in G_k computed by the algorithm - M_k is just the matching M from which the edges of the blossoms shrunk in G_k have been removed. Note that the set of the vertices of G_k that are unmatched in M_k is still X. Notice also that all vertices of a blossom become even whenever we expand them, since the stem is an even-length alternating path from X to the base v of the blossom, and all other vertices of the blossom are reachable from v by an even-length alternating path which goes around the blossom in one of the directions (as it is odd). An example of G with matching M and the shrunk graph G_k is given below.



Also, we claim that the vertices in $V(G_k)$ have the same classification (as even, odd, or free) no matter whether we classify them with respect to G_k and M_k, or G and M. Indeed, first consider an alternating path (of even or odd parity) from X to v in G_k. As we expand blossoms, if our alternating path went through the shrunk blossom then we can easily update the alternating path into the expanded graph without modifying the parity of its length as the alternating path will be entering the blossom through its base. Conversely, if we have an alternating path P in G from X to a vertex v which intersects a blossom B then consider the first time P visits a vertex of the corresponding flower. We can now replace the this prefix of P with part of the flower in such a way that we still have an alternating path and the parity of the length of the path has not changed. By properties of the algorithm, G_k has no flowers, and M_k is a maximum matching in G_k. Therefore, G_k has no alternating walk from X to X – if such walk existed then from the previous lecture we would know that there is either an augmenting path or a flower in G_k. This fact implies the following.

Claim 2. *In G_k there is no edge between two even vertices.*

Proof. If such an edge $e = uv$ existed, then by Claim 1, G_k contained an alternating walk P of odd length from X to v. But v is even, so there would also be an alternating path P' of even length from X to v. Concatenating P with the reverse of P', we would obtain an alternating walk from X to X, contradicting the definition of G_k. □

Claim 3. $\mathsf{Even} = D(G) = \{v : \exists$ *a maximum-size matching missing* $v\}$.

Proof. ($\subseteq$) Certainly, if v is even then there is a maximum-size matching M' missing v. Such a matching is obtained by taking an even-length alternating path P from X to v and putting $M' = M \Delta P$.

($\supseteq$) Conversely, if there exists a maximum-size matching M' missing v, then $M \Delta M'$ is a union of even-length cycles and even-length paths, and v is an endpoint of one of these paths, because it does not belong to an edge of M' The other endpoint of this path P does not belong to an edge of M, i.e., it is an element of X. This confirms that P is an even-length alternating path from X to v. □

Claim 4. $\mathsf{Odd} = A(G) = \{v : v$ *is a neighbor of some* $u \in D(G)$*, but* $v \notin D(G)\}$.

Proof. ($\subseteq$) If v is odd, then there is an alternating path of odd length from X to v. The vertex preceding v on this path must be even, thus v is a neighbor of some vertex from Even. Moreover, since it is odd then it is not in Even. But by Claim 3, Even $= D(G)$, so indeed $v \in A(G)$.

($\supseteq$) The reverse inclusion follows from Claim 1, which ensures that every vertex adjacent to Even belongs to Even $\cup$ Odd, which in conjunction with Claim 3 gives us that $v \in$ Odd. □

Claim 5. $\mathsf{Free} = C(G) = V(G) \setminus (D(G) \cup A(G))$.

Proof. immediate from the definition of Free, and from the preceeding two claims which identify Even, Odd with $D(G)$, $A(G)$, respectively. □

We proceed to proving the desired properties of the decomposition stated in Theorem 2. We start with Property (ii) which is directly implied by the following claim asserting that not only all the vertices of $C(G)$ are matched in M, but also the edges matching them are always connecting two free vertices.

Claim 6. $|M \cap C(G)| = |C(G)|/2$.

Proof. Consider some v from $C(G)$ (which is equal to Free by Claim 6). By Corollary 1 we know that v cannot be adjacent to any even vertex, so $C(G)$ is disconnected from $D(G)$ in $G \setminus A(G)$. Moreover, v has to be matched by some edge $e = vu$ in M, otherwise it would be even. However, u cannot be odd, since then we could augment the odd-length path from X to u by e which would imply that v is either odd or even. Therefore, we must have u being free as well. This implies that $M \cap C(G)$ matches all the vertices of $C(G)$ and thus has the desired size. □

To establish Properties (iii) and (iv) we prove the following claim. Recall that X is the set of exposed vertices of M.

Claim 7. *For every connected component H of $G \setminus A(G)) \cap D(G)$:*

(a) either $|X \cap H| = 1$ and $|M \cap \delta(H)| = 0$; or $|X \cap H| = 0$ and $|M \cap \delta(H)| = 1$, where $\delta(H)$ is the set of edges with exactly one endpoint in H.

(b) H is factor-critical.

Proof. The proof is by induction on the number of blossoms which are shrunk during the execution of Edmonds' Algorithm. If no blossoms are shrunk, then $G = G_k$ and the claim follows as a consequence of Corollary 1 and Claim 3 that assert that $(G \setminus A(G)) \cap D(G)$ is the union of isolated vertices (for which both ((a)) and ((b)) trivially hold). Now for the induction step, suppose B is a blossom in G and that the claim holds for G/B (in which B is shrunk). In this case, B corresponds to a vertex $b \in G/B$ which has to be even, since the stem of the flower containing B corresponds to an even-length alternating path from X to b in G/B. In fact, as it was already mentioned before, in G all vertices of B are even and they have all to be in the same connected component, say H_b, of $(G \setminus A(G)) \cap D(G)$. Clearly, since the vertices of $B \setminus \{b\}$ are all matched in M by edges inside B, neither the size of $M \cap \delta(H_b)$ nor the size of $X \cap H_b$ can increase as a result of expanding B in G/B. Thus, we see that ((a)) holds.

Now to prove ((b)), we note that, by inductive assumption, all connected components of $(G \setminus A(G)) \cap D(G)$ other than H_b are factor-critical. Thus, it remains to show that H_b is factor-critical as well. To this end, assume that some vertex $v \in H_b$ was removed. If $v \notin B$ then, by inductive assumption, we know that there exists a matching M' in H_b that matches all vertices of H_b except v and $B \setminus \{b\}$. But then M' can be straight-forwardly augmented inside B to match all vertices in $B \setminus \{b\}$. Similarly, if $v \in B$ then we know that there is a matching M'' that matches all vertices of H_b except B – this correspond to situation in which we remove b in G/B. But, if we remove any vertex of a blossom the rest of them can be easily matched within B, thus once again giving raise to matching that matches the whole $H_b \setminus \{v\}$. This concludes the proof. □

Having proved Claim 7, Property (iii) follows since each factor-critical graph has to be odd-sized, and Property (iv) is implied by Claim 6 which shows that all odd-sized connected components of $G \setminus A(G)$ are in $D(G)$. Finally, we prove Property (i).

Claim 8. $|M| = \frac{1}{2}\Big(|V| + |A(G)| - o(G \setminus A(G))\Big)$.

Proof. We only need to show that $|M| \geq \frac{1}{2}\Big(|V| + |A(G)| - o(G \setminus A(G))\Big)$. Observe that

$$|M| \geq |M \cap C(G)| + |M \cap E(D(G))| + |M \cap \delta(A(G))| \,.$$

By Claim 6, the first term is $|C(G)|/2$. By Claim 7 part (a), the second term is $\frac{|D(G)| - o(G \setminus A(G))}{2}$ while the third term is $|A(G)|$ since every vertex of $A(G)$ is matched to a vertex of $D(G)$. Thus,

$$|M| \geq \frac{1}{2}\Big(|C(G)| + |D(G)| + 2|A(G)| - o(G \setminus A(G))\Big) = \frac{1}{2}\Big(|V| + |A(G)| - o(G \setminus A(G))\Big) \,,$$

proving the claim. □