Applications of Hall's theorem

DEFINITION. Let $\mathcal{S} = \{S_1, S_2, \dots, S_m\}$ be a family of finite nonempty subsets of a set S. (S_i need not be distinct.) A **transversal** (or a **system of distinct representatives**) of $\mathcal{S}$ is a subset $T \subseteq S$ of m distinct elements one from each S_i.

As an example, let $S = [1,6]$, and let $S_1 = S_2 = \{1,2\}$, $S_3 = \{2,3\}$ and $S_4 = \{1,4,5,6\}$. For $\mathcal{S} = \{S_1, S_2, S_3, S_4\}$, the set $T = \{1,2,3,4\}$ is a transversal. If we add the set $S_5 = \{2,3\}$ to $\mathcal{S}$, then it is impossible to find a transversal for this new family.

The connection of transversals to the Marriage Theorem is as follows. Let $S = Y$ and $X = [1,m]$. Form an (X,Y)-bipartite graph G such that there is an edge (i,s) if and only if $s \in S_i$. The possible transversals T of $\mathcal{S}$ are then obtained from the matchings M saturating X in G by taking the ends in Y of the edges of M.

Corollary 3.2. *Let $\mathcal{S}$ be a family of finite nonempty sets. Then $\mathcal{S}$ has a transversal if and only if the union of any k of the subsets S_i of $\mathcal{S}$ contains at least k elements.*

Example 3.6. An $m \times n$ **latin rectangle** is an $m \times n$ integer matrix M with entries $M_{ij} \in [1,n]$ such that the entries in the same row and in the same column are different. Moreover, if $m = n$, then M is a **latin square**. Note that in a $m \times n$ latin rectangle M, we always have that $m \leq n$.

We show the following: *Let M be an $m \times n$ latin rectangle (with $m < n$). Then M can be extended to a latin square by the addition of $n - m$ new rows.*

The claim follows when we show that M can be extended to an $(m+1) \times n$ latin rectangle. Let $A_i \subseteq [1,n]$ be the set of those elements that do not occur in the i-th column of M. Clearly, $|A_i| = n - m$ for each i, and hence $\sum_{i \in I} |A_i| = |I|(n-m)$ for all subsets $I \subseteq [1,n]$. Now $|\cup_{i \in I} A_i| \geq |I|$, since otherwise at least one element from the union would be in more than $n - m$ of the sets A_i with $i \in I$. However, each row has all the n elements, and therefore each i is missing from exactly $n - m$ columns. By Marriage Theorem, the family $\{A_1, A_2, \dots, A_n\}$ has a transversal, and this transversal can be added as a new row to M. This proves the claim.

Tutte's theorem

The next theorem is a classic characterization of perfect matchings.

DEFINITION. A connected component of a graph G is said to be **odd** (**even**), if it has an odd (even) number of vertices. Denote by $c_{\text{odd}}(G)$ the number of odd connected components in G.

Denote by $m(G)$ be the number of edges in a maximum matching of a graph G.

Theorem 3.10 (Tutte-Berge Formula). *Each maximum matching of a graph G has*

$$m(G) = \min_{S \subseteq V_G} \frac{\nu_G + |S| - c_{\text{odd}}(G{-}S)}{2} \tag{3.3}$$

elements.

Note that the condition in (ii) includes the case, where $S = \emptyset$.

Proof. We prove the result for connected graphs. The result then follows for disconnected graphs by adding the formulas for the connected components.

We observe first that $\leq$ holds in (3.3), since, for all $S \subseteq V_G$,

$$m(G) \leq |S| + m(G{-}S) \leq |S| + \frac{|V_G \setminus S| - c_{\text{odd}}(G{-}S)}{2} = \frac{\nu_G + |S| - c_{\text{odd}}(G{-}S)}{2}.$$

Indeed, each odd component of $G{-}S$ must have at least one unsaturated vertex.

The proof proceeds by induction on ν_G. If $\nu_G = 1$, then the claim is trivial. Suppose that $\nu_G \geq 2$.

Assume first that there exists a vertex $v \in G$ such that v is saturated by all maximum matchings. Then $m(G{-}v) = m(G) - 1$. For a subset $S' \subseteq G{-}v$, denote $S = S' \cup \{v\}$. By the induction hypothesis, for all $S' \subseteq G{-}v$,

$$\begin{aligned} m(G) - 1 &\geq \frac{1}{2}\left((\nu_G - 1) + |S'| - c_{\text{odd}}(G{-}(S' \cup \{v\}))\right) \\ &= \frac{1}{2}\left((\nu_G + |S| - c_{\text{odd}}(G{-}S))\right) - 1. \end{aligned}$$

The claim follows from this.

Suppose then that for each vertex v, there is a maximum matching that does not saturate v. We claim that $m(G) = (\nu_G - 1)/2$. Suppose to the contrary, and let M be a maximum matching having two different unsaturated vertices u and v, and choose M so that the distance $d_G(u,v)$ is as small as possible. Now $d_G(u,v) \geq 2$, since otherwise $uv \in G$ could be added to M, contradicting the maximality of M. Let w be an intermediate vertex on a shortest path $u \xrightarrow{\star} v$. By assumption, there exists a maximum matching N that does not saturate w. We can choose N such that the intersection $M \cap N$ is maximal. Since $d_G(u,w) < d_G(u,v)$ and $d_G(w,v) < d_G(u,v)$, N saturates both u and v. The (maximum) matchings N and M leave equally many vertices unsaturated, and hence there exists another vertex $x \neq w$ saturated by M but which is unsaturated by N. Let $e = xy \in M$. If y is also unsaturated by N, then $N \cup \{e\}$ is a matching, contradicting maximality of N. It also follows that $y \neq w$. Therefore there exists an edge $e' = yz$ in N, where $z \neq x$. But now $N' = N \cup \{e\} \setminus \{e'\}$ is a maximum matching that does not saturate w. However, $N \cap M \subset N' \cap M$ contradicts the choice of N. Therefore, every maximum matching leaves exactly one vertex unsaturated, i.e., $m(G) = (\nu_G - 1)/2$.

In this case, for $S = \emptyset$, the right hand side of (3.3) gets value $(\nu_G - 1)/2$, and hence, by the beginning of the proof, this must be the minimum of the right hand side. □

For perfect matchings we have the following corollary, since for a perfect matching we have $m(G) = (1/2)\nu_G$.

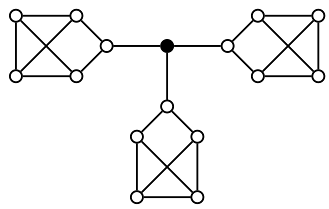
Theorem 3.11 (TUTTE (1947)). *Let G be a nontrivial graph. The following are equivalent.*

(i) G has a perfect matching.
(ii) For every proper subset $S \subset V_G$, $c_{\text{odd}}(G-S) \le |S|$.

Tutte's theorem does not provide a good algorithm for constructing a perfect matching, because the theorem requires 'too many cases'. Its applications are mainly in the proofs of other results that are related to matchings. There is a good algorithm due to EDMONDS (1965), which uses 'blossom shrinkings', but this algorithm is somewhat involved.

Example 3.7. The simplest connected graph that has no perfect matching is the path P_3. Here removing the middle vertex creates two odd components.

The next 3-regular graph (known as the **Sylvester graph**) does not have a perfect matching, because removing the black vertex results in a graph with three odd connected components. This graph is the smallest regular graph with an odd degree that has no perfect matching.



Using Theorem 3.11 we can give a short proof of PETERSEN's result for 3-regular graphs (1891).

Theorem 3.12 (PETERSEN (1891)). *If G is a bridgeless 3-regular graph, then it has a perfect matching.*

Proof. Let S be a proper subset of V_G, and let G_i, $i \in [1,t]$, be the odd connected components of $G-S$. Denote by m_i the number of edges with one end in G_i and the other in S. Since G is 3-regular,

$$\sum_{v \in G_i} d_G(v) = 3 \cdot \nu_{G_i} \quad \text{and} \quad \sum_{v \in S} d_G(v) = 3 \cdot |S|\,.$$

The first of these implies that

$$m_i = \sum_{v \in G_i} d_G(v) - 2 \cdot \varepsilon_{G_i}$$

is odd. Furthermore, $m_i \neq 1$, because G has no bridges, and therefore $m_i \geq 3$. Hence the number of odd connected components of $G-S$ satisfies

$$t \le \frac{1}{3}\sum_{i=1}^{t} m_i \le \frac{1}{3}\sum_{v \in S} d_G(v) = |S|\,,$$

and so, by Theorem 3.11, G has a perfect matching. □

Stable Marriages

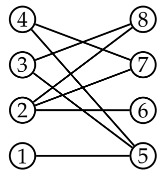
DEFINITION. Consider a bipartite graph G with a bipartition (X, Y) of the vertex set. In addition, each vertex $x \in G$ supplies an order of preferences of the vertices of $N_G(x)$. We write $u <_x v$, if x prefers v to u. (Here $u, v \in Y$, if $x \in X$, and $u, v \in X$, if $x \in Y$.) A matching M of G is said to be **stable**, if for each unmatched pair $xy \notin M$ (with $x \in X$ and $y \in Y$), it is not the case that x and y prefer each other better than their matched companions:

$$xv \in M \text{ and } y <_x v, \quad \text{or} \quad uy \in M \text{ and } x <_y u.$$

We omit the proof of the next theorem.

Theorem 3.13. *For bipartite graphs G, a stable matching exists for all lists of preferences.*

Example 3.8. That was the good news. There is a catch, of course. A stable matching need not saturate X and Y. For instance, the graph on the right does have a perfect matching (of 4 edges).



Suppose the preferences are the following:

$1\colon 5$	$2\colon 6 < 8 < 7$	$3\colon 8 < 5$	$4\colon 7 < 5$
$5\colon 4 < 1 < 3$	$6\colon 2$	$7\colon 2 < 4$	$8\colon 3 < 2$

Then there is no stable matchings of four edges. A stable matching of G is the following: $M = \{28, 35, 47\}$, which leaves 1 and 6 unmatched. (You should check that there is no stable matching containing the edges 15 and 26.)

Theorem 3.14. *Let $G = K_{n,n}$ be a complete bipartite graph. Then G has a perfect and stable matching for all lists of preferences.*

Proof. Let the bipartition be (X, Y). The algorithm by GALE AND SHAPLEY (1962) works as follows.

Procedure.
Set $M_0 = \emptyset$, and $P(x) = \emptyset$ for all $x \in X$.
Then iterate the following process until all vertices are saturated:
Choose a vertex $x \in X$ that is unsaturated in M_{i-1}. Let $y \in Y$ be the most preferred vertex for x such that $y \notin P(x)$.

(1) Add y to $P(x)$.
(2) If y is not saturated, then set $M_i = M_{i-1} \cup \{xy\}$.
(3) If $zy \in M_{i-1}$ and $z <_y x$, then set $M_i = (M_{i-1} \setminus \{zy\}) \cup \{xy\}$.

First of all, the procedure terminates, since a vertex $x \in X$ takes part in the iteration at most n times (once for each $y \in Y$). The final outcome, say $M = M_t$, is a perfect matching, since the iteration continues until there are no unsaturated vertices $x \in X$.

Also, the matching $M = M_t$ is stable. Note first that, by (3), if $xy \in M_i$ and $zy \in M_j$ for some $x \neq z$ and $i < j$, then $x <_y z$. Assume the that $xy \in M$, but $y <_x z$ for some $z \in Y$. Then xy is added to the matching at some step, $xy \in M_i$, which means that $z \in P(x)$ at this step (otherwise x would have 'proposed' z). Hence x took part in the iteration at an earlier step M_k, $k < i$ (where z was put to the list $P(x)$, but xz was not added). Thus, for some $u \in X$, $uz \in M_{k-1}$ and $x <_z u$, and so in M the vertex z is matched to some w with $x <_z w$.

Similarly, if $x <_y v$ for some $v \in X$, then $y <_v z$ for the vertex $z \in Y$ such that $vz \in M$. □