2 Trees

What is a tree?

Forests and Trees: A *forest* is a graph with no cycles, a *tree* is a connected forest.

Theorem 2.1 *If G is a forest, then $comp(G) = |V(G)| - |E(G)|$.*

Proof: We proceed by induction on $|E(G)|$. As a base, if $|E(G)| = 0$, then every component is an isolated vertex, so $comp(G) = |V(G)|$ as required. For the inductive step, we may assume $|E(G)| > 0$ and choose an edge $e \in E(G)$. Now, by Theorem 2.5 and induction on $G - e$, we have

$$\begin{aligned} comp(G) &= comp(G-e) - 1 \\ &= |V(G-e)| - |E(G-e)| - 1 \\ &= |V(G)| - |E(G)|. \end{aligned}$$

□

Corollary 2.2 *If G is a tree, then $|V(T)| = |E(T)| + 1$.*

Leaf: A *leaf* is a vertex of degree 1.

Proposition 2.3 *Let T be a tree with $|V(T)| \geq 2$. Then T has ≥ 2 leaf vertices. Further, if T has exactly 2 leaf vertices, then T is a path.*

Proof: By the above corollary and Theorem 1.1, we have

$$\begin{aligned} 2 &= 2|V(T)| - 2|E(T)| \\ &= 2|V(T)| - \sum_{v \in V(T)} deg(v) \\ &= \sum_{v \in V(T)} (2 - deg(v)). \end{aligned}$$

Since $|V(T)| \geq 2$, every vertex has degree > 0. It follows immediately from this and the above equation that T has ≥ 2 leaf vertices. Further, if T has exactly two leaf vertices, then every other vertex of T has degree 2, and it follows that T is a path. □

Lemma 2.4 *If T is a tree and v is a leaf of T, then $T - v$ is a tree.*

Proof: It is immediate that $T - v$ has no cycle. If $u, w \in V(T - v)$, then there is a path P from u to w in T and this path cannot contain v, so P is also a path in $T - v$. Thus $T - v$ is connected. $\square$

Note: The above lemma gives us a powerful inductive tool for proving properties of trees.

Proposition 2.5 *If T is a tree and $u, v \in V(T)$, then there is a unique path from u to v.*

Proof: We proceed by induction on $V(T)$. If there is a leaf $w \neq u, v$, then the result follows by applying induction to $T - w$. Otherwise, the result follows from Proposition 4.3. $\square$

Spanning Tree: If $T \subseteq G$ is a tree and $V(T) = V(G)$, we call T is a *spanning tree of* G.

If G is a graph, $H \subseteq G$ and $e \in E(G)$, we let $H + e$ be the subgraph of G obtained from H by adding the edge e and the endpoints of e.

Theorem 2.6 *Let G be a connected graph with $|V(G)| > 1$. If H is a subgraph of G chosen according to one of the following conditions, then H is a spanning tree.*

(i) *$H \subseteq G$ is minimal so that H is connected and $V(H) = V(G)$.*

(ii) *$H \subseteq G$ is maximal so that H has no cycles.*

Proof: For (i), note that if H has a cycle C and $e \in E(C)$, then $H - e$ is connected (by Theorem 2.5), which contradicts the minimality of H. Thus H has no cycle, and it is a spanning tree.

For (ii), note first that $V(H) = V(G)$ by the maximality of H. If X is the vertex set of a component of H and $X \neq V(G)$, then it follows from the connectivity of G that there exists an edge e of G with one end in X and one end in $V(G) \setminus X$. Now, e is a cut-edge of $H + e$ (by Theorem 2.5) so $H + e$ has no cycle, contradicting the maximality of H. Thus, H has only one component, and it is a spanning tree. $\square$

Proposition 2.7 *Let G be a graph with $|E(G)| = |V(G)| - 1$.*

(i) *If G has no cycle, then it is a tree.*

(ii) *If G is connected, it is a tree.*

Proof: Part (i) follows immediately from Theorem 4.1. For (ii), we have a connected graph G, so we may choose a spanning tree $T \subseteq G$. But then $|E(G)| = |V(G)| - 1 = |V(T)| - 1 = |E(T)|$ so $T = G$ and G is a tree. $\square$

Kruskal's Algorithm

Fundamental Cycles: Let T be a spanning tree of G and let $f \in E(G) \setminus E(T)$. A cycle $C \subseteq G$ with $f \in E(C)$ and $C - f$ a path of T is called a *fundamental cycle of f with respect to T*.

Proposition 2.8 *If T is a spanning tree of G and $f \in E(G) \setminus E(T)$, then there is exactly one fundamental cycle of f with respect to T.*

Proof: This follows immediately from Proposition 4.5. $\square$

Proposition 2.9 *Let T be a spanning tree of G, let $e \in E(T)$ and $f \in E(G) \setminus E(T)$.*

(i) *if e is in the fundamental cycle of f, then $T - e + f$ is a tree*

(ii) *if f has one end in each component of $T - e$, then $T - e + f$ is a tree.*

Proof: For (i), note that if e is in the fundamental cycle of f, then $T + f - e$ is a connected graph with one fewer edge than vertex, so it is a tree by Proposition 4.7.

For (ii), observe that if f has one end in each component of $T - e$, then $T - e + f$ is a forest (since f is a cut-edge of $T - e + f$ and $T - e$ is a forest) with one fewer edge than vertex, so it is a tree by Proposition 4.7. $\square$

Weighted graphs and min-cost trees: A *weighted graph* is a graph G together with a *weight* function on the edges $w : E(G) \to \mathbb{R}$. If $T \subseteq G$ is a spanning tree for which $\sum_{e \in E(T)} w(e)$ is minimum, we call T a *min-cost tree.*

Kruskal's Algorithm:

input:	A weighted graph G
output:	A subgraph $T \subseteq G$
procedure:	Choose a sequence of edges $e_1, e_2, \ldots, e_m$ according to the rule that e_i is an edge of minimum weight in $E(G) \setminus \{e_1, \ldots, e_{i-1}\}$ so that $\{e_1, \ldots, e_i\}$ does not contain the edge set of a cycle. When no such edge exists, stop and return the subgraph T consisting of all the vertices, and all chosen edges $\{e_1, \ldots, e_m\}$.

Theorem 2.10 *Let G be a connected weighted graph with weight function w. If w is one-to-one, then Kruskal's algorithm returns the unique min-cost tree for G.*

Proof: Let $e_1, \ldots, e_m$ be the sequence of edges chosen by Kruskal's Algorithm, and let T be the subgraph returned by it. It follows from the connectivity of G and Theorem 4.6 that T is a spanning tree. Suppose (for a contradiction) that there is a min-cost tree $T' \neq T$ and let f be the edge of minimum weight in the set $E(T') \setminus E(T)$. Let C be the fundamental cycle of f with respect to T and let $e_i \in E(C)$. Now, part (i) of Proposition 5.2 shows that $T - e_i + f$ is a tree, so, in particular, there is no cycle with edge set included in $\{e_1, \ldots, e_{i-1}, f\}$, and by Kruskal's Algorithm, we must have $w(f) > w(e_i)$. It follows that every edge in $C - f$ has smaller weight than f. However, $C - f$ is a path with the same ends as f, so there must exist an edge in $C - f$ with one end in each component of $T' - f$. But then, part (ii) of Proposition 5.2 gives us a contradiction to the assumption that T' is a min-cost tree. This contradiction proves that T is the unique min-cost tree of G, as required. $\square$

Distance and Dijkstra's Algorithm

Distance: If $u, v \in V(G)$, the *distance from u to v* is the length of the shortest path from u to v, or ∞ if no such path exists. If G is a weighted graph, then the *distance from u to v* is the minimum of $\sum_{e \in E(P)} w(e)$ over all paths from u to v, or ∞ if no such path exists. In either case, we denote this by $dist_G(u, v)$, or just $dist(u, v)$ if G is clear from context.

Observation 2.11 *The distance function obeys the triangle inequality. So, for all $u, v, w \in V(G)$ we have:*

$$dist(u, v) + dist(v, w) \geq dist(u, w).$$

Shortest Path Tree: If G is a weighted graph and $r \in V(G)$, a tree $T \subseteq G$ (not nec. spanning) is a *shortest path tree* for r if $dist_T(r,v) = dist_G(r,v)$ for every $v \in V(T)$.

Dijkstra's Algorithm:

input: A connected weighted graph G with $w : E(G) \to \mathbb{R}^+$ and a vertex r.

output: A tree $T \subseteq G$.

procedure: Start with T_1 the tree consisting of the vertex r. At step i, we have $T_i \subseteq G$. If $V(T_i) = V(G)$ stop and return $T = T_i$. Otherwise, choose an edge uv with $u \in V(T_i)$ and $v \in V(G) \setminus V(T_i)$ so that $w(uv) + dist_T(r,u)$ is minimum and set $T_{i+1} = T_i + uv$.

Theorem 2.12 *Dijkstra's algorithm returns a shortest path tree for r.*

Proof: We prove by induction on i that each T_i in the algorithm is a shortest path tree for r. As a base, note that this is true for T_1 (since all weights are nonnegative). For the inductive step, we shall show that T_{i+1} is a shortest path tree for r assuming this holds for T_i. Assume that $T_{i+1} = T_i + uv$. Let P be the path of minimum distance in G from r to v, let y be the first vertex of P which is not in the tree T, and let the previous edge be xy. Then we have

$$\begin{aligned} dist_G(r,v) &= \sum_{e \in E(P)} w(e) \\ &\geq dist_{T_i}(r,x) + w(xy) \\ &\geq dist_{T_i}(r,u) + w(uv) \\ &= dist_{T_{i+1}}(r,v). \end{aligned}$$

It follows that T_{i+1} is a shortest path tree for r, as required. $\square$

Prüfer codes

In this section, we consider two graphs with the same vertex set to be the same if they have the same adjacencies.

Prüfer Encoding

input: A tree T with vertex set $S \subseteq \mathbb{Z}$ where $|S| = n \geq 2$.

output: A sequence $\mathbf{a} = (a_1, \ldots, a_{n-2})$ with elements in S.

procedure: At step i, we delete from T the smallest leaf vertex v, and we set a_i to be the vertex adjacent to v.

Observation 2.13 *If* $\mathbf{a}$ *is the Prüfer Encoding of the tree* T*, then the set of vertices which appear in* $\mathbf{a}$ *is exactly the set of non-leaf nodes of* T*.*

Prüfer Decoding

input:	A sequence $\mathbf{a} = (a_1, \ldots, a_{n-2})$ with elements in $S \subseteq \mathbb{Z}$ where $\|S\| = n \geq 2$.
output:	An tree T with vertex set S.
procedure:	Start with the graph T where $V(T) = S$ and $E(T) = \emptyset$, and with all vertices unmarked. At step i, add an edge from the smallest unmarked vertex v which does not appear in $(a_i, \ldots, a_{n-2})$ to a_i and mark v. After step $n-2$ is complete, add an edge between the two remaining unmarked vertices and stop.

Theorem 2.14 *Let* $S \subseteq \mathbb{Z}$ *with* $|S| = n \geq 2$*. Prüfer Encoding and Decoding are inverse bijections between the set of trees with vertex set* S *and the set of sequences of length* $n-2$ *with elements in* S*.*

Proof: We proceed by induction on n. As a base, observe that when $n = 2$, there is a single sequence of length 0 and a single tree with vertex set S, and the encoding/decoding operations exchange these.

For the inductive step, we may assume $n \geq 3$. Let T be a tree with vertex set S, let $\mathbf{a} = (a_1, \ldots, a_{n-2})$ be the encoding of T, and let T' be the decoding of $\mathbf{a}$. Let v be the smallest leaf of T. Then, by the encoding process, v is adjacent to a_1, and v is smaller than any other vertex which does not appear in $\mathbf{a}$, since (by the above observation) all such vertices are leaves. It follows from the decoding rules that in the tree T', the vertex v is a leaf adjacent to a_1. Now, $T - v$ encodes to $(a_2, \ldots, a_n)$ which decodes to $T' - v$, so by induction $T - v$ and $T' - v$ are the same, and it follows that T and T' are the same.

To complete the inductive step, we still need to show that every sequence $\mathbf{a} = (a_1, \ldots, a_{n-2})$ is the encoding of some tree. To see this, let $v \in S$ be the smallest element which does not appear in $\mathbf{a}$. Then, by induction $(a_2, \ldots, a_{n-2})$ is the encoding of some tree T with vertex set $S \setminus \{v\}$, and we find that $\mathbf{a}$ is an encoding of the tree obtained from T by adding the vertex v and the edge va_1. This completes the proof. $\square$

Corollary 2.15 *For every* $n \geq 2$*, there are exactly* n^{n-2} *trees with vertex set* $\{1, 2, \ldots, n\}$*.*

Matrices

Incidence and Adjacency Matrices: If G is a graph, the *incidence matrix* of G is the matrix $\{M_{v,e}\}_{v \in V(G), e \in E(G)}$ given by the following rule:

$$M_{v,e} = \begin{cases} 0 & \text{if } e \text{ is not incident with } v \\ 1 & \text{if } e \text{ is a non-loop incident with } v \\ 2 & \text{if } e \text{ is a loop incident with } v \end{cases}$$

(so each row of the incidence matrix corresponds to a vertex, and each column corresponds to an edge). The *adjacency matrix* of G is the matrix $\{A_{u,v}\}_{u,v \in V(G)}$ given by the rule that $A_{u,v}$ is the number of edges with ends u, v (so, for this matrix, both the rows and columns are indexed by vertices).

Theorem 2.16 *If A is the adjacency matrix of the graph G, and k is a positive integer, then the u, v entry of A^k is the number of walks from u to v of length k*

Proof: For every $u, v \in V(G)$ and integer i, we let $w^i_{u,v}$ denote the number of walks from u to v of length i. We prove the theorem by induction on k. As a base, note that this is true for $k = 1$ by our definitions. For the inductive step, we may assume $k \geq 2$. Now, for every $u, v \in V(G)$ we have:

$$\begin{aligned} A^k_{u,v} &= \sum_{w \in V(G)} A^{k-1}_{u,w} A_{w,v} \\ &= \sum_{w \in V(G)} w^{k-1}_{u,v} A_{w,v} \\ &= w^k_{u,v} \end{aligned}$$

as required. $\square$.

Laplacian and Oriented Incidence Matrices: If G is a simple graph, the *Laplacian matrix* of G is the matrix $\{L_{u,v}\}_{u,v \in V(G)}$ given by the following rule:

$$L_{u,v} = \begin{cases} deg(u) & \text{if } u = v \\ -1 & \text{if } u, v \text{ are adjacent} \\ 0 & \text{otherwise} \end{cases}$$

Note that (since G is loopless), every column of the incidence matrix has two entries which are 1 and all the other entries are 0. An *oriented incidence matrix* of G is obtained from the incidence matrix by changing one of the 1 entries in each column to a -1 (arbitrarily).

Observation 2.17 *If L is the Laplacian matrix of G and M is an oriented incidence matrix of G, then $MM^\top = L$.*

Proof: For every $v \in V(G)$ we let x_v denote the row of M indexed v. Then the u, v entry of $MM^\top$ is the dot product of x_u and x_v, and this will be $deg(u)$ if $u = v$, -1 if u is adjacent to v, and 0 otherwise. $\square$

Submatrices: Let R be a matrix indexed by $I \times J$. If $I' \subseteq I$ and $J' \subseteq J$, we let $R|^{I'}$ denote the submatrix of R consisting of rows in I' and we let $R|_{J'}$ denote the submatrix of R consisting of columns in J'. Let G be a graph and let $r \in V(G)$. If M is an oriented incidence matrix of G, the r-truncation of M is $M|^{V(G)\setminus\{r\}}$. Similarly, if L is the Laplacian matrix of G, the r-truncation of L is defined to be $L|^{V(G)\setminus\{r\}}_{V(G)\setminus\{r\}}$.

Proposition 2.18 *Let G be a simple connected graph on n vertices, let $r \in V(G)$, and let M^* be the r-truncation of an oriented incidence matrix of G. Then for every $S \subseteq E(G)$ with $|S| = n - 1$ we have*

$$det(M^*|_S) = \begin{cases} \pm 1 & \textit{if } S \textit{ is the edge set of a spanning tree} \\ 0 & \textit{otherwise} \end{cases}$$

Proof: For every edge $e \in E(G)$, let y_e be the column vector of M^* corresponding to e. So, if $e = vr$, then y_e has ± 1 in the position corresponding to v and 0 everywhere else, and if $e = uv$ with $u, v \neq r$, then y_e has ± 1 in the position corresponding to v, the opposite value in the position corresponding to u, and 0 everywhere else.

First suppose that the subgraph with vertex set $V(G)$ and edge set S is not connected, let H be a component of this graph which does not contain r, and let $U = V(H)$. Now consider the (column) vector χ_U indexed by $V(G) \setminus \{r\}$ and given by the rule that for every $v \in V(G) \setminus \{r\}$, the vector χ_U has a 1 in the coordinate corresponding to v if $v \in U$ and otherwise has a 0 in this position. Since $r \notin U$ and no edge in S has one end in U and one in $V(G) \setminus U$, it follows that χ_U is orthogonal to y_e for every $e \in S$. Thus, the columns in S do not span χ_U and we find that $det(M^*|_S) = 0$.

If the subgraph with vertex set $V(G)$ and edge set S is connected, then it must be the edge set of a spanning tree since $|S| = n-1 = |V(G)| - 1$ (Prop. 4.7). In this case, we prove by induction on n that $det(M^*|_S) = \pm 1$. As a base, note that when $n = 1$ the matrix M^* is empty, and if $S = \emptyset$ we have $det(M^*|_S) = det([]) = 1$. For the inductive step, we may then assume that $n \geq 2$, so S is the edge set of a spanning tree on ≥ 2 vertices. Choose a leaf vertex v of this spanning tree, and let $e \in S$ be the edge incident with v. Now, the only column of $M^*|_S$ with a nonzero entry in the coordinate corresponding to v is y_e, and the entry of y_v in this coordinate is ± 1. It follows that $det(M^*|_S)$ is equal to $\pm$ the determinant of the matrix obtained from $M^*|_S$ by removing the row v and the column e. By induction, this smaller matrix has determinant ± 1, and this completes the proof. $\square$

Theorem 2.19 (Binet-Cauchy Formula) *Let A be an $n \times m$ matrix with columns indexed by J and B be an $m \times n$ matrix with rows indexed by J, then*

$$det(AB) = \sum_{S \subseteq J : |S| = n} det(A|_S \, B|^S)$$

Proof: Then this theorem follows from (computing determinants) in the equation below (here we let I denote an identity matrix and 0 denote a zero matrix of the appropriate size).

$$\begin{bmatrix} I & 0 \\ A & I \end{bmatrix} \begin{bmatrix} -I & B \\ A & 0 \end{bmatrix} = \begin{bmatrix} -I & B \\ 0 & AB \end{bmatrix}$$

$\square$

Theorem 2.20 (The Matrix Tree Theorem) *If G is a connected graph on n vertices, $r \in V(G)$, and L^* is the r-truncated Laplacian matrix for G, then $det(L^*)$ is the number of spanning trees of G.*

Proof: Let M^* be the r-truncation of an oriented incidence matrix for G. Then, by the argument in Observation 8.2 and the Binet-Cauchy formula, we find

$$\begin{aligned} det(L^*) &= det(M^*(M^*)^\top) \\ &= \sum_{S \subseteq E(G) : |S| = n-1} det(M^*|_S) det((M^*)^\top|^S) \\ &= |\{T \subseteq G : T \text{ is a spanning tree}\}| \end{aligned}$$

thus completing the proof. $\square$

3 Matchings

Hall's Theorem

Matching: A *matching* in G is a subset $M \subseteq E(G)$ so that no edge in M is a loop, and no two edges in M are incident with a common vertex. A matching M is *maximal* if there is no matching M' with $M \subset M'$ and *maximum* if there is no matching M'' with $|M| < |M''|$.

Alternating & Augmenting Paths: If M is a matching in G, a path $P \subseteq G$ is M-*alternating* if the edges of P belong alternately to M and to $E(G) \setminus M$ (in other words, for every $v \in V(P)$ with degree 2 in P, some edge of P incident with v is in M). The path P is M-*augmenting* if it is M-alternating, has distinct ends, say u, v, and no edge of M is incident with u or v in G (not just in P).

Theorem 3.1 (Berge) *A matching M in G is maximum if and only if there is no M-augmenting path.*

Proof: For the "only if" direction we prove the contrapositive. Assuming G contains an M-augmenting path P, the set $(M \setminus E(P)) \cup (E(P) \setminus M)$ is a matching with larger cardinality than M, so M is not maximum.

For the "if" direction, we also prove the contrapositive, so we shall assume that M is not maximum, and show there is an augmenting path. Since M is not maximum, there exists a matching M' with $|M'| > |M|$. Consider the subgraph $H \subseteq G$ with $V(H) = V(G)$ and $E(H) = M \cup M'$. Every component of this graph is either a cycle of even length with edges alternately in M and M', a path with edges alternately in M and M', or a path consisting of one edge e with $e \in M \cap M'$. Since $|M'| > |M|$, there is a component of H which is a path with more edges in M' than M. Then P is an M-augmenting path. $\square$

Neighbors: If $X \subseteq V(G)$, the *neighbors* of X, is the set

$$N(X) = \{v \in V(G) \setminus X : v \text{ is adjacent to some point in } X\}.$$

For $x \in X$, we define $N(x) = N(\{x\})$.

Cover: We say that a set of edges $S \subseteq E(G)$ *covers* a set of vertices X if every $x \in X$ is incident with some edge in S. Similarly, a set of vertices $X \subseteq V(G)$ *covers* a set of edges $S \subseteq E(G)$ if every edge in S is incident with some point in X.

Theorem 3.2 (Hall's Marriage Theorem) *Let G be a bipartite graph with bipartition (A, B). Then, there is a matching $M \subseteq G$ which covers A if and only if $|N(X)| \geq |X|$ for every $X \subseteq A$.*

Proof: The "only if" condition is obvious: if there exists $X \subseteq A$ with $|N(X)| < |X|$, then no matching can cover A.

For the "if" direction, let M be a maximum matching, and suppose that M does not cover A. Choose a vertex $u \in A$ not covered by M, and define the sets X, Y as follows:

$$\begin{aligned} X &= \{x \in A : \text{there is an } M\text{-alternating path from } u \text{ to } x\} \\ Y &= \{y \in B : \text{there is an } M\text{-alternating path from } u \text{ to } y\} \end{aligned}$$

Let $M' \subseteq M$ be the set of edges in M which are incident with a point in $X \cup Y$. By parity, every M-alternating path which begins at u and ends at a point in $X \setminus \{u\}$ must have its last edge in M, so every point in $X \setminus \{u\}$ is incident with an edge in M'. If there is a point $y \in Y$ not incident with an edge in M', then there is an M-alternating path from u to y which is M-augmenting, contradicting the previous theorem. Thus, every point in Y is incident with some edge in M'. It follows from this that $|X \setminus \{u\}| = |M'| = |Y|$.

Let $x \in X$ and let $y \in N(x)$. Since $x \in A$, we may choose an M-alternating path from u to x. Note (as before) that the last edge of this path is in M'. If y appears in this path, then $y \in Y$. Otherwise, we may extend this path by the edge xy to a new M-alternating path. Thus, in either case, we find that $y \in Y$. It follows from this that $N(X) \subseteq Y$. But then, $|N(X)| \leq |Y| = |M'| < |X|$. This completes the proof. $\square$

Regular A graph G is *k-regular* if every vertex of G has degree k. We say that G is *regular* if it is k-regular for some k.

Perfect Matchings: A matching M is *perfect* if it covers every vertex.

Corollary 3.3 *Every regular bipartite graph has a perfect matching.*

Proof: Let G be a k-regular bipartite graph with bipartition (A, B). Let $X \subseteq A$ and let t be the number of edges with one end in X. Since every vertex in X has degree k, it follows that $k|X| = t$. Similarly, every vertex in $N(X)$ has degree k, so t is less than or equal to $k|N(X)|$. It follows that $|X|$ is at most $|N(X)|$. Thus, by Hall's Theorem, there is a matching covering A, or equivalently, every maximum matching covers A. By a similar argument, we find that every maximum matching covers B, and this completes the proof. $\square$

Stable Marriages

System of Preferences: If G is a graph, a *system of preferences* for G is a family $\{>_v\}_{v \in V(G)}$ so that each $>_v$ is a linear ordering of $N(v)$. If $u, u' \in N(v)$ and $u >_v u'$, we say that v *prefers* u to u'.

Marriage Systems and Stable Marriages: A *Marriage System* consists of a complete bipartite graph $K_{n,n}$ with bipartition $(men, women)$ which is equipped with a system of preferences. We say that a matching M is *stable* if there do not exist edges $mw, m'w' \in M$ with $m, m' \in men$ and $w, w' \in women$ so that m prefers w' to w and w' prefers m to m'. A matching which covers every vertex and is stable is called a *stable marriage.*

Gale-Shapley Algorithm:

input: A marriage system.

output: A stable marriage.

procedure: At each step, every man proposes to the woman he prefers most among those who have not yet rejected him. If every woman receives at most one proposal, stop and output the corresponding matching. Otherwise, every woman who receives more than one proposal says "maybe" to the man who proposes to her whom she most prefers, and rejects the others who proposed.

Theorem 3.4 *The Gale-Shapley Algorithm outputs a stable marriage (as claimed).*

Proof: Note first that this algorithm must terminate, since some man is rejected at each non-final step (and the total number of rejections is no more than n^2). Let M be the marriage resulting from this algorithm, and say that a man m and woman w are *married* if $mw \in M$.

Suppose that the woman w receives proposals from some nonempty set $X \subseteq men$ at some step. Then w says "maybe" to the man m who she prefers most among X, and at the next step, m will again propose to w (since the set of women who have rejected him has not changed). This immediately implies the following claim.

Claim: Every woman w is married to the man m she most prefers among those who propose to her during the algorithm. In particular, if w has at least one proposal, then w is married to some man.

With this claim, we now show that M covers every vertex. Suppose (for a contradiction) that M does not cover some man m. Then m must have been rejected by every woman. But then, by the claim, every woman must be married. Since $|men| = |women|$, this is contradictory.

Next let us show that M is stable. Suppose (for a contradiction) that it is not, and choose $mw, m'w' \in M$ so that m prefers w' to w and w' prefers m to m'. It follows from the definition of the algorithm that m must have proposed to w' at some step (since m will propose to w' before w). Applying the claim to w', we see that w' must be married to m or a person she prefers to m, thus contradicting our assumptions.

It follows that M is a stable marriage, as claimed. $\square$

Fact: Let M be the stable marriage output by the above algorithm and let M' be another stable marriage. Then, for every man m, if $mw \in M$ and $mw' \in M'$, then either $w = w'$ or m prefers w to w'. Similarly, for every woman w, if $wm \in M$ and $wm' \in M'$, then either $m = m'$, or w prefers m' to m. So, among all stable marriages, the Gale-Shapley algorithm produces one which is best possible for every man, and worst possible for every woman.

Covers

Covers: A *vertex cover* of G is a set of vertices $X \subseteq V(G)$ so that every edge is incident with some vertex in X. Simlarly, an *edge cover* of G is a set of edges $S \subseteq V(G)$ so that every vertex is incident with some edge in S.

Independent Set: A subset of vertices $X \subseteq V(G)$ is *independent* if there is no loop with endpoint in X and there is no non-loop with both ends in X.

Matching & Cover Parameters: For every graph G, we define the following parameters

$\alpha(G)$	maximum size of an independent set
$\alpha'(G)$	maximum size of a matching
$\beta(G)$	minimum size of a vertex cover
$\beta'(G)$	minimum size of an edge cover

Observation 3.5 $\alpha(G) + \beta(G) = |V(G)|$ *for every simple graph* G.

Proof: A set $X \subseteq V(G)$ is independent if and only if $V(G) \setminus X$ is a vertex cover. Thus, the complement of an independent set of maximum size is a vertex cover of minimum size. $\square$

Theorem 3.6 (König, Egerváry) *If G is bipartite, then $\alpha'(G) = \beta(G)$.*

Proof: It is immediate that $\beta(G) \geq \alpha'(G)$ since for a maximum matching M, any vertex cover must contain at least one endpoint of each edge in M.

Next we shall show that $\beta(G) \leq \alpha'(G)$. Let (A, B) be a bipartition of G, let X be a vertex cover of minimum size, and define two bipartite subgraphs H_1 and H_2 so that H_1 has bipartition $(A \cap X, B \setminus X)$, H_2 has bipartition $(A \setminus X, B \cap X)$, and both H_1 and H_2 have all edges with both ends in their vertex sets.

Suppose (for a contradiction) that there does not exist a matching in H_1 which covers $A \cap X$. Then, by Hall's theorem, there is a subset $Y \subseteq A \cap X$ so that $|N_{H_1}(Y)| < |Y|$. Now, we claim that the set $X' = (X \setminus Y) \cup N_{H_1}(Y)$ is a vertex cover. Let $e \in E(G)$. If e has one end in $B \cap X$, then e is covered by X'. If e has no end in $B \cap X$, then (since X is a vertex cover) e must have one end in $A \cap X$ and the other in $B \setminus X$, so $e \in E(H_1)$. If e does not have an end in Y, then e is covered by $X \setminus Y \subseteq X'$. Otherwise, e is an edge in H_1 with one end in Y, so its other end is in $N_{H_1}(Y)$ and we again find that e is covered. But then X' is a vertex cover with $|X'| = |X| - |Y| + |N_{H_1}(Y)| < |X|$, giving us a contradiction.

Thus, H_1 has a matching M_1, which covers $A \cap X$. By a similar argument, H_2 has a matching, M_2, which covers $B \cap X$. Since these subgraphs have disjoint vertex sets, $M = M_1 \cup M_2$ is a matching of G. Furthermore, $\alpha'(G) \geq |M| = |X| = \beta(G)$. This completes the proof. $\square$

Theorem 3.7 (Gallai) *If G is a simple connected graph with at least two vertices, then $\alpha'(G) + \beta'(G) = |V(G)|$.*

Proof: First, let M be a maximum matching (so $|M| = \alpha'(G)$). Now, we form an edge cover L from M as follows: For every vertex v not covered by M, choose an edge e incident with v and add e to L. Then L is an edge cover, so $\beta'(G) \leq |L| = |M| + |V(G)| - 2|M| = |V(G)| - \alpha'(G)$.

Next, let L be a minimum edge cover (so $|L| = \beta'(G)$) and consider the subgraph H consisting of all the vertices, and those edges in L. Since L is a minimum edge cover, it follows that $L \setminus \{e\}$ is not an edge cover for every $e \in L$. Thus, every edge $e \in E(H)$

must have one endpoint of degree 1 in H. It follows from this that every component of H is isomorphic to a star (a graph of the form $K_{1,m}$ for some positive integer m). Choose a matching $M \subseteq L$ by selecting one edge from each component of H. Then we have $\alpha'(G) \geq |M| = comp(H) = |V(G)| - |L| = |V(G)| - \beta'(G)$.

Combining the two inequalities yields $\alpha'(G) + \beta'(G) = |V(G)|$, as required. □

Corollary 3.8 *If G is a connected bipartite graph with at least two vertices, then $\alpha(G) = \beta'(G)$.*

Proof: By Observation 11.1 and Theorem 11.3 we have $\alpha(G) + \beta(G) = \alpha'(G) + \beta'(G)$. Now, subtracting the relation $\beta(G) = \alpha'(G)$ proved in 11.2 we have $\alpha(G) = \beta'(G)$ as desired. □

Tutte's Theorem

Odd Components: For every graph G, we let $odd(G)$ denote the number of components of G which have an odd number of vertices.

Identification: If $X \subseteq V(G)$, we may form a new graph from G by merging all vertices in X to a single new vertex. If an edge has an endpoint in X, then this edge will have the new vertex as its new endpoint. We say this graph is obtained from G by *identifying* X.

Theorem 3.9 (Tutte) *G has a perfect matching if and only if $odd(G-X) \leq |X|$ for every $X \subseteq V(G)$*

Proof: The "only if" is immediate: if G has a set $X \subseteq V(G)$ with $odd(G - X) > |X|$, then G cannot have a perfect matching.

We prove the "if" direction by induction on $|V(G)|$. As a base, observe that this is trivial when $|V(G)| \leq 2$. For the inductive step, let G be a graph for which $odd(G - X) \leq |X|$ for every $X \subseteq V(G)$ and assume the theorem holds for all graphs with fewer vertices. Call a set $X \subseteq V(G)$ *critical* if $odd(G - X) \geq |X| - 1$. We shall establish the theorem in steps.

(1) $|V(G)|$ is even

This follows from $odd(G - \emptyset) \leq |\emptyset| = 0$.

(2) If X is critical, then $odd(G - X) = |X|$.

This follows from the observation that $|X| + odd(G - X) \cong |V(G)|$ (modulo 2).

(3) There is a critical set.

For instance, $\emptyset$ is critical.

Based on (3), we may now choose a maximal critical set X. Let $|X| = k$ and let the odd components of $G - X$ be $G_1, \ldots, G_k$.

(4) $G - X$ has no even components.

If $G - X$ has an even component G', then choose $v \in V(G')$. Now $X \cup \{v\}$ is critical, contradicting the choice of X.

(5) For every $1 \leq i \leq k$ and $v \in V(G_i)$, the graph $G_i - v$ has a perfect matching.

If not, then by induction there exists $Y \subseteq V(G_i - v)$ so that $(G_i - v) - Y$ has $> |Y|$ odd components. But then $G \setminus (X \cup Y \cup \{v\})$ has $\geq |X| + |Y|$ odd components, so it is critical - again contradicting the maximality of X.

(6) G has a matching M with $|M| = k$ so that M covers X and every G_i has exactly one vertex covered by M.

Construct a graph H from G by identifying $V(G_i)$ to a new vertex y_i for every $1 \leq i \leq k$ and then deleting every loop and every edge with both ends in X. Now, H is bipartite with bipartition (X, Y) where $Y = \{y_1, \ldots, y_k\}$. Suppose (for a contradiction) that H does not have a perfect matching. Then by Hall's Theorem there exists $Y' \subseteq Y$ with $|N_H(Y')| < |Y'|$. Let $X' = N_H(Y')$. Now the graph $G - X'$ has $\geq |Y'| > |X'|$ odd components, giving us a contradiction. So, H has a perfect matching, which proves (6).

It follows from (5) and (6) that G has a perfect matching, as desired. $\square$

Theorem 3.10 (Tutte-Berge Formula)

$$\alpha'(G) = \frac{1}{2}\left(|V(G)| - \max_{X \subseteq V(G)} (odd(G - X) - |X|)\right)$$

Proof: Let $k = \max_{X \subseteq V(G)}(odd(G - X) - |X|)$ and choose $X \subseteq V(G)$ so that $k = odd(G - X) - |X|$. Note that $k = odd(G - X) - |X| \cong odd(G - X) + |X| \cong |V(G)|$ (modulo 2). By considering X and $odd(G - X)$ we find that every matching of G must not cover $\geq k$ vertices, so $\alpha'(G) \leq \frac{1}{2}(|V(G)| - k)$.

To prove the other inequality, we construct a new graph G' from G by adding a set Y of k new vertices to G each adjacent to every other vertex. Let $Z' \subseteq V(G')$. We claim that $odd(G' - Z') \leq |Z'|$. If $Z' = \emptyset$, then this follows from the observation that $k \cong |V(G)|$ (modulo 2). If $Y \not\subseteq Z'$, then $G' - Z'$ is connected, so $|Z'| \geq 1 \geq odd(G' - Z')$. Finally, if $Y \subseteq Z'$, then we have

$$odd(G' - Z') = odd(G - Z) \leq k + |Z| = |Y| + |Z| = |Z'|.$$

Since Z' was arbitrary, Tutte's Theorem shows that G' has a perfect matching, and it follows that G has a matching covering all but k vertices, so $\alpha'(G) \geq \frac{1}{2}(|V(G)| - k)$ as required. $\square$

Theorem 3.11 (Petersen) *If every vertex of G has degree 3 and G has no cut-edge, then G has a perfect matching.*

Proof: We shall show that G satisfies the condition for Tutte's Theorem. Let $X \subseteq V(G)$, let $G_1, \ldots, G_k$ be the odd components of $G - X$, and for every $1 \leq i \leq k$ let S_i be the set of edges with one end in X and the other in $V(G_i)$. Now, for every $1 \leq i \leq G_i$, we have $3|V(G_i)| = \sum_{v \in V(G_i)} deg_G(v) = |S_i| + 2|E(G_i)|$. Since $|V(G_i)|$ is odd, it follows that $|S_i|$ must also be odd. By our assumptions, $|S_i| \neq 1$, so we conclude that $|S_i| \geq 3$.

Now, form a new graph H from G by deleting every vertex in every even component of $G - X$, then identifying every G_i to a single new vertex y_i, and then deleting every loop and every edge with both ends in X. This graph H is bipartite with bipartition (X, Y) where $Y = \{y_1, \ldots, y_k\}$. Furthermore, by our assumptions, every vertex in Y has degree ≥ 3 and every vertex in X has degree ≤ 3. Thus, we have

$$3|X| \geq \sum_{x \in X} deg_H(x) = |E(H)| = \sum_{y \in Y} deg_H(y) \geq 3|Y|.$$

So, $|X| \geq |Y| = k = odd(G - X)$. Since X was arbitrary, it follows from Theorem 3.9 that G has a perfect matching. $\square$

4 Connectivity

2-connectivity

Separation: A *separation* of G of order k is a pair of subgraphs (H, K) with $H \cup K = G$ and $E(H \cap K) = \emptyset$ and $|V(H) \cap V(K)| = k$. Such a separation is *proper* if $V(H) \setminus V(K)$ and $V(K) \setminus V(H)$ are nonempty.

Observation 4.1 *G has a proper separation of order 0 if and only if G is disconnected.*

Cut-vertex: A vertex v is a *cut-vertex* if $comp(G - v) > comp(G)$.

Observation 4.2 *If G is connected, then v is a cut-vertex of G if and only if there exists a proper 1-separation (H, K) of G with $V(H) \cap V(K) = \{v\}$.*

Proposition 4.3 *Let e, f be distinct non-loop edges of the graph G. Then exactly one of the following holds:*

(i) *There exists a cycle C with $e, f \in E(C)$*

(ii) *There is a separation (H, K) of order ≤ 1 with $e \in E(H)$ and $f \in E(K)$.*

Proof: It is clear that (i) and (ii) are mutually exclusive, so it suffices to show that (i) or (ii) holds. For this, we may assume that G is connected, and set k to be the length of the shortest walk containing e, f. We proceed by induction k. For the base case, if $k = 2$, then we may assume $e = uv$ and $f = vw$. If u, w are in the same component of the graph $G - v$, then (i) holds. Otherwise, v is a cut-vertex and (ii) holds.

For the inductive step, we may assume $k \geq 3$. Let $f = uv$ and choose an edge $f' = vw$ so that there is a walk containing e, f' of length $k - 1$. First suppose that there is a cycle C containing e, f'. If $C - v$ and u are in distinct components of $G - v$, then v is a cut-vertex and (ii) holds. Otherwise, we may choose a path $P \subseteq G - v$ from u to a vertex of $V(C) \setminus \{v\}$. Now $P \cup C + f$ has a cycle which contains e, f, so (i) holds. If there is no cycle containing e, f', then it follows from our inductive hypothesis that there is a 1-separation (H, K) with $e \in E(H)$ and $f' \in E(K)$. Suppose (for a contradiction) that $f \in E(H)$. Then $V(H) \cap V(K) = \{v\}$, the shortest walk containing e, f has length k and the shortest walk containing e, f' has length $< k$ which is contradictory. Thus, $f \in E(K)$ and (H, K) satisfy (ii). This completes the proof. $\square$

2-connected: A graph G is 2-connected if it is connected, $|V(G)| \geq 3$, and G has no cut-vertex.

Theorem 4.4 *Let G be a graph with at least three vertices. Then the following are equivalent:*

(i) *G is 2-connected*

(ii) *For all $x, y \in V(G)$ there exists a cycle C with $x, y \in V(C)$.*

(iii) *G has no vertex of degree 0, and for all $e, f \in E(G)$ there exists a cycle C with $e, f \in E(C)$.*

Proof: It is easy to see that (iii) implies (ii): to find a cycle C with $x, y \in V(C)$ just choose an edge e incident with x and an edge f incident with y and apply (iii) to e, f. Trivially, (ii) implies (i). So, to complete the argument, we need only show that (i) implies (iii), but this is an immediate consequence of the previous proposition. $\square$

Block: A *block* of G is a maximal connected subgraph $H \subseteq G$ so that H does not have a cut-vertex. Note that if H is a block, then either H is 2-connected, or $|V(H)| \leq 2$.

Proposition 4.5 *If H_1, H_2 are distinct blocks in G, then $|V(H_1) \cap V(H_2)| \leq 1$.*

Proof: Suppose (for a contradiction) that $|V(H_1) \cap V(H_2)| \geq 2$. Let $H' = H_1 \cup H_2$, let $x \in H'$ and consider $H' - x$. By assumption, $H_1 - x$ is connected, and $H_2 - x$ is connected. Since these graphs share a vertex, $H' - x = (H_1 - x) \cup (H_2 - x)$ is connected. Thus, H' has no cut-vertex. This contradicts the maximality of H_1, thus completing the proof. $\square$

Block-Cutpoint graph: If G is a graph, the *block-cutpoint* graph of G, denoted $BC(G)$ is the simple bipartite graph with bipartition (A, B) where A is the set of cut-vertices of G, and B is the set of blocks of G, and $a \in A$ and $b \in B$ adjacent if the block b contains the cut-vertex a.

Observation 4.6 *If G is connected, then $BC(G)$ is a tree.*

Proof: Let (A, B) be the bipartition of $BC(G)$ as above. It follows from the connectivity of G that $BC(G)$ is connected. If there is a cycle $C \subseteq BC(G)$, then set H to be the union of all blocks in $B \cap V(C)$. It follows that H is a 2-connected subgraph of G (as in the proof of the previous proposition). This contradicts the maximality of the blocks in $B \cap V(C)$. $\square$

Ears: An *ear* of a graph G is a path $P \subseteq G$ which is maximal subject to the constraint that all interior vertices of P have degree 2 in G. An *ear decomposition* of G is a decomposition of G into $C, P_1, \ldots, P_k$ so that C is a cycle of length ≥ 3, and for every $1 \leq i \leq k$, the subgraph P_i is an ear of $C \cup P_1 \cup \ldots P_i$.

Theorem 4.7 *A graph G is 2-connected if and only if it has an ear decomposition.*

Proof: For the "if" direction, let $C, P_1, \ldots, P_k$ be an ear decomposition of G. We shall prove that G is 2-connected by induction on k. As a base, if $k = 0$, then $G = C$ is 2-connected. For the inductive step, we may assume that $k \geq 1$ and that $C \cup P_1 \cup \ldots P_{k-1}$ is 2-connected. It then follows easily that $G = C \cup P_1 \cup \ldots P_k$ is also 2-connected.

We prove the "only if" direction by a simple process. First, choose a cycle $C \subseteq G$ of length ≥ 3 (this is possible by Theorem 4.4). Next we choose a sequence of paths $P_1, \ldots, P_k$ as follows. If $G' = C \cup P_1 \cup \ldots P_{i-1} \neq G$, then choose an edge $e \in E(G')$ and $f \in E(G) \setminus E(G')$, and then choose a cycle $D \subseteq G$ containing e, f (again using 4.4). Finally, let P_i be the maximal path in D which contains the edge f but does not contain any edge in $E(G')$. Then P_i is an ear of $C \cup P_1 \cup \ldots P_i$, and when this process terminates, we have an ear decomposition. $\square$

Menger's Theorem

Theorem 4.8 *Let G be a graph, let $A, B \subseteq V(G)$ and let $k \geq 0$ be an integer. Then exactly one of the following holds:*

(i) *There exist k pairwise (vertex) disjoint paths $P_1, \ldots, P_k$ from A to B.*

(ii) *There is a separation (H, K) of G of order $< k$ with $A \subseteq V(H)$ and $B \subseteq V(K)$.*

Proof: It is clear that (i) and (ii) are mutually exclusive, so it suffices to show that (i) or (ii) holds. We prove this by induction on $|E(G)|$. As a base, observe that the theorem holds trivially when $|E(G)| \leq 1$. For the inductive step, we may then assume $|E(G)| \geq 2$. Choose an edge e and consider the graph $G' = G - e$. If G' contains k disjoint paths from A to B, then so does G, and (i) holds. Otherwise, by induction, there is a separation (H, K) of G' of order $< k$ with $A \subseteq V(H)$ and $B \subseteq V(K)$.

Now consider the separations $(H + e, K)$ and $(H, K + e)$ of G. If one of these separations has order $< k$, then (ii) holds. Thus, we may assume that e has one end in $V(H) \setminus V(K)$, the other end in $V(K) \setminus V(H)$, and both $(H + e, K)$ and $(H, K + e)$ have order k. Choose (H', K') to be one of these two separations with $E(H'), E(K') \neq \emptyset$ (this is possible since $|E(G)| \geq 2$) and set $X = V(H') \cap V(K')$ (note that $|X| = k$). Now, we apply the theorem inductively to the graph H' for the sets A, X and to K' for the sets X, B. If there are k disjoint paths from A to X in H' and k disjoint paths from X to B in K', then (i) holds. Otherwise, by induction there is a separation of H' or K' in accordance with (ii), and it follows that (ii) is satsified. $\square$

Note: The above theorem implies Theorem 11.2 (König Egerváry). Simply apply the above theorem to the bipartite graph G with bipartition (A, B). Then (i) holds if and only if $\alpha'(G) \geq k$, and (ii) holds if and only if $\beta(G) < k$ (here $V(H) \cap V(K)$ is a vertex cover).

Internally Disjoint: The paths $P_1, \ldots, P_k$ are *internally disjoint* if they are pairwise vertex disjoint except for their ends.

Theorem 4.9 (Menger's Theorem) *Let u, v be distinct non-adjacent vertices of G, and let $k \geq 0$ be an integer. Then exactly one of the following holds:*

(i) *There exist k internally disjoint paths $P_1, \ldots, P_k$ from u to v.*

(ii) *There is a separation (H, K) of G of order $< k$ with $u \in V(H) \setminus V(K)$ and $v \in V(K) \setminus V(H)$.*

Proof: Let $A = N(u)$ and $B = N(v)$ and apply the above theorem to $G - \{u, v\}$. $\square$

k-Connected: A graph G is *k-connected* if $|V(G)| \geq k + 1$ and $G - X$ is connected for every $X \subseteq V(G)$ with $|X| < k$. Note that this generalizes the notion of 2-connected from Section 13. Also note that 1-connected is equivalent to connected.

Corollary 4.10 *A simple graph G with $|V(G)| \geq k+1$ is k-connected if and only if for every $u, v \in V(G)$ there exist k internally disjoint paths from u to v.*

Line Graph: If G is a graph, the *line graph* of G, denoted $L(G)$, is the simple graph with vertex set $E(G)$, and two vertices $e, f \in E(G)$ adjacent if e, f share an endpoint in G.

Edge cut: If $X \subseteq V(G)$, we let $\delta(X) = \{xy \in E(G) : x \in X \text{ and } y \notin X\}$, and we call any set of this form an *edge cut*. If $v \in V(G)$ we let $\delta(v) = \delta(\{v\})$.

Theorem 4.11 (Menger's Theorem - edge version) *Let u, v be distinct vertices of G and let $k \geq 0$ be an integer. Then exactly one of the following holds:*

(i) *There exist k edge disjoint paths $P_1, \ldots, P_k$ from u to v.*

(ii) *There exists $X \subseteq V(G)$ with $u \in X$ and $v \notin X$ so that $|\delta(X)| < k$.*

Proof: Apply Theorem 4.8 to the graph $L(G)$ for $\delta(u)$ and $\delta(v)$ and k. $\square$

k-edge-connected: A graph G is k-edge-connected if $G - S$ is connected for every $S \subseteq E(G)$ with $|S| < k$.

Corollary 4.12 *A graph G is k-edge-connected if and only if for every $u, v \in V(G)$ there exist k pairwise edge disjoint paths from u to v.*

Fans and Cycles

Subdivision: If $e = uv$ is an edge of the graph G, then we *subdivide* e by removing the edge e, adding a new vertex w, and two new edges uw and wv.

Observation 4.13

1. *Subdividing an edge of a 2-connected graph yields a 2-connected graph.*
2. *Adding an edge to a k-connected graph results in a k-connected graph.*
3. *If G is k-connected and $A \subseteq V(G)$ satisfies $|A| \geq k$, then adding a new vertex to G and an edge from this vertex to each point in A results in a k-connected graph.*

Fan: Let $v \in V(G)$ and let $A \subseteq V(G) \setminus \{v\}$. A *$(v, A)$-fan* of *size* k is a collection of k paths $\{P_1, \ldots, P_k\}$ so that each P_i is a path from v to a point in A, and any two such paths intersect only in the vertex v.

Lemma 4.14 *If G is k-connected, $v \in V(G)$ and $A \subseteq V(G) \setminus \{v\}$ satisfies $|A| \geq k$, then G contains a (v, A)-fan of size k.*

Proof: Construct a new graph G' from G by adding a new vertex u and then adding a new edge between u and each point of A. By the above observation, G' is k-connected, and $u, v \in V(G')$ are nonadjacent, so by Menger's theorem there exist k internally disjoint paths from u to v. Removing the vertex u from each of these paths yields a (v, A)-fan of size k in G. $\square$

Theorem 4.15 *Let G be a k-connected graph with $k \geq 2$ and let $X \subseteq V(G)$ satisfy $|X| = k$. Then there exists a cycle $C \subseteq G$ with $X \subseteq V(C)$.*

Proof: Choose a cycle $C \subseteq G$ so that $|V(C) \cap X|$ is maximum, and suppose (for a contradiction) that $X \not\subseteq V(C)$. Choose a vertex $v \in X \setminus V(C)$ and set $k' = \min\{k, |V(C)|\}$. It follows from the above lemma that G has a $(v, V(C))$-fan of size k', say $\{P_1, \ldots, P_{k'}\}$. Since $|X \cap V(C)| < k$, it follows that there exists a cycle $C' \subseteq C \cup P_1 \cup \ldots \cup P_{k'}$ so that $\{v\} \cup (X \cap V(C)) \subseteq V(C')$. This contradiction completes the proof. $\square$

5 Directed Graphs

What is a directed graph?

Directed Graph: A *directed graph*, or *digraph*, D, consists of a set of *vertices* $V(D)$, a set of *edges* $E(D)$, and a function which assigns each edge e an ordered pair of vertices (u, v). We call u the *tail* of e, v the *head* of e, and u, v the *ends* of e. If there is an edge with tail u and head v, then we let (u, v) denote such an edge, and we say that this edge is directed *from u to v*.

Loops, Parallel Edges, and Simple Digraphs: An edge $e = (u, v)$ in a digraph D is a *loop* if $u = v$. Two edges e, f are *parallel* if they have the same tails and the same heads. If D has no loops or parallel edges, then we say that D is *simple*.

Drawing: As with undirected graphs, it is helpful to represent them with drawings so that each vertex corresponds to a distinct point, and each edge from u to v is represented by a curve directed from the point corresponding to u to the point corresponding to v (usually we indicate this direction with an arrowhead).

Orientations: If D is a directed graph, then there is an ordinary (undirected) graph G with the same vertex and edge sets as D which is obtained from D by associating each edge (u, v) with the ends u, v (in other words, we just ignore the directions of the edges). We call G the *underlying (undirected)* graph, and we call D an *orientation* of G.

Standard Diraphs

Null digraph	the (unique) digraph with no vertices or edges.
Directed Path	a graph whose vertex set may be numbered $\{v_1, \ldots, v_n\}$ and edges may be numbered $\{e_1, \ldots, e_{n-1}\}$ so that $e_i = (v_i, v_{i+1})$ for every $1 \leq i \leq n - 1$.
Directed Cycle	a graph whose vertex set may be numbered $\{v_1, \ldots, v_n\}$ and edges may be numbered $\{e_1, \ldots, e_n\}$ so that $e_i = (v_i, v_{i+1})$ (*modulo* n) for every $1 \leq i \leq n$
Tournament	A digraph whose underlying graph is a complete graph.

Subgraphs and Isomorphism: These concepts are precisely analogous to those for undirected graphs.

Degrees: The *outdegree* of a vertex v, denoted $deg^+(v)$ is the number of edges with tail v, and the *indegree* of v, denoted $deg^-(v)$ is the number of edges with head v.

Theorem 5.1 *For every digraph D*

$$\sum_{v \in V(D)} deg^+(v) = |E(D)| = \sum_{v \in V(D)} deg^-(v)$$

Proof: Each edge contributes exactly 1 to the terms on the left and right. $\square$

Connectivity

Directed Walks & Paths: A *directed walk* in a digraph D is a sequence $v_0, e_1, v_1, \ldots, e_n v_n$ so that $v_i \in V(D)$ for every $0 \leq i \leq n$, and so that e_i is an edge from v_{i-1} to v_i for every

$1 \leq i \leq n$. We say that this is a walk from v_0 to v_n. If $v_0 = v_n$ we say the walk is *closed* and if $v_0, v_1, \ldots, v_n$ are distinct we call it a *directed path*.

Proposition 5.2 *If there is a directed walk from u to v, then there is a directed path from u to v.*

Proof: Every directed walk from u to v of minimum length is a directed path. $\square$

δ^+ and δ^-: If $X \subseteq V(D)$, we let $\delta^+(X)$ denote the set of edges with tail in X and head in $V(G) \setminus X$, and we let $\delta^-(X) = \delta^+(V(G) \setminus X)$.

Proposition 5.3 *Let D be a digraph and let $u, v \in V(D)$. Then exactly one of the following holds.*

(i) *There is a directed walk from u to v.*

(ii) *There exists $X \subseteq V(D)$ with $u \in X$ and $v \notin X$ so that $\delta^+(X) = \emptyset$.*

Proof: It is immediate that (i) and (ii) are mutually exclusive, so it suffices to show that at least one holds. Let $X = \{w \in V(D) : \text{there is a directed walk from } u \text{ to } w\}$. If $v \in X$ then (i) holds. Otherwise, $\delta^+(X) = \emptyset$, so (ii) holds. $\square$

Strongly Connected: We say that a digraph D is *strongly connected* if for every $u, v \in V(D)$ there is a directed walk from u to v.

Proposition 5.4 *Let D be a digraph and let $H_1, H_2 \subseteq D$ be strongly connected. If $V(H_1) \cap V(H_2) \neq \emptyset$, then $H_1 \cup H_2$ is strongly connected.*

Proof: If $v \in V(H_1) \cap V(H_2)$, then every vertex has a directed walk both to v and from v, so it follows that $H_1 \cup H_2$ is strongly connected. $\square$

Strong Component: A *strong component* of a digraph D is a maximal strongly connected subgraph of D.

Theorem 5.5 *Every vertex is in a unique strong component of D.*

Proof: This follows immediately from the previous proposition, and the observation that a one-vertex digraph is strongly connected.

Observation 5.6 *Let D be a digraph in which every vertex has outdegree ≥ 1. Then D contains a directed cycle.*

Proof: Construct a walk greedily by starting at an arbitrary vertex v_0, and at each step continue from the vertex v_i along an arbitrary edge with tail v_i (possible since each vertex has outdegree ≥ 1) until a vertex is repeated. At this point, we have a directed cycle. $\square$

Acyclic: A digraph D is *acyclic* if it has no directed cycle.

Proposition 5.7 *The digraph D is acyclic if and only if there is an ordering $v_1, v_2, \ldots, v_n$ of $V(D)$ so that every edge (v_i, v_j) satisfies $i < j$.*

Proof: The "if" direction is immediate. We prove the "only if" direction by induction on $|V(D)|$. As a base, observe that this is trivial when $|V(D)| = 1$. For the inductive step, we may assume that D is acyclic, $|V(D)| = n \geq 2$, and that the proposition holds for all digraphs with fewer vertices. Now, apply the Observation 5.6 to choose a vertex v_n with $deg^+(v_n) = 0$. The digraph $D - v_n$ is acyclic, so by induction we may choose an ordering $v_1, v_2, \ldots, v_{n-1}$ of $V(D - v_n)$ so that every edge (v_i, v_j) satisfies $i < j$. But then $v_1, \ldots, v_n$ is such an ordering of $V(D)$. $\square$

Proposition 5.8 *Let D be a digraph, and let D' be the digraph obtained from D by taking each strong component $H \subseteq D$, identifying $V(H)$ to a single new vertex, and then deleting any loops. Then D' is acyclic.*

Proof: If D' had a directed cycle, then there would exist a directed cycle in D not contained in any strong component, but this contradicts Theorem 5.5. $\square$

Theorem 5.9 *If G is a 2-connected graph, then there is an orientation D of G so that D is strongly connected.*

Proof: Let $C, P_1, \ldots, P_k$ be an ear decomposition of G. Now, orient the edges of C to form a directed cycle, and orient the edges of each path P_i to form a directed path. It now follows from the obvious inductive argument (on k) that the resulting digraph D is strongly connected. $\square$

Eulerian, Hamiltonian, & path partitions

Proposition 5.10 *Let D be a digraph and assume that $deg^+(v) = deg^-(v)$ for every vertex v. Then there exists a list of directed cycles $C_1, C_2, \ldots, C_k$ so that every edge appears in exactly one.*

Proof: Choose a maximal list of cycles $C_1, C_2, \ldots, C_k$ so that every edge appears in at most one. Suppose (for a contradiction) that there is an edge not included in any cycle C_i and let H be a component of $D \setminus \cup_{i=1}^{k} E(C_i)$ which contains an edge. Now, every vertex $v \in V(H)$ satisfies $deg_H^+(v) = deg_H^-(v) \neq 0$, so by Observation 17.5 there is a directed cycle $C \subseteq H$. But then C may be appended to the list of cycles $C_1, \ldots, C_k$. This contradiction completes the proof. $\square$

Eulerian: A closed directed walk in a digraph D is called *Eulerian* if it uses every edge exactly once. We say that D is *Eulerian* if it has such a walk.

Theorem 5.11 *Let D be a digraph D whose underlying graph is connected. Then D is Eulerian if and only if $deg^+(v) = deg^-(v)$ for every $v \in V(D)$.*

Proof: The "only if" direction is immediate. For the "if" direction, choose a closed walk $v_0, e_1, \ldots, v_n$ which uses each edge at most once and is maximum in length (subject to this constraint). Suppose (for a contradiction) that this walk is not Eulerian. Then, as in the undirected case, it follows from the fact that the underlying graph is connected that there exists an edge $e \in E(D)$ which does not appear in the walk so that e is incident with some vertex in the walk, say v_i. Let $H = D - \{e_1, e_2, \ldots, e_n\}$. Then every vertex of H has indegree equal to its outdegree, so by the previous proposition, there is a list of directed cycles in H containing every edge exactly once. In particular, there is a directed cycle $C \subseteq H$ with $e \in C$. But then, the walk obtained by following $v_0, e_1, \ldots, v_i$, then following the directed cycle C from v_i back to itself, and then following $e_{i+1}, v_i, \ldots, v_n$ is a longer closed walk which contradicts our choice. This completes the proof. $\square$

Hamiltonian: Let D be a directed graph. A cycle $C \subseteq D$ is *Hamiltonian* if $V(C) = V(D)$. Similarly, a path $P \subseteq D$ is *Hamiltonian* if $V(P) = V(D)$.

In & Out Neighbors: If $X \subseteq V(D)$, we define

$$\begin{aligned} N^+(X) &= \{y \in V(D) \setminus X : (x,y) \in E(D) \text{ for some } x \in X\} \\ N^-(X) &= \{y \in V(D) \setminus X : (y,x) \in E(D) \text{ for some } x \in X\} \end{aligned}$$

We call $N^+(X)$ the *out-neighbors* of X and $N^-(X)$ the *in-neighbors* of X. If $x \in X$ we let $N^+(x) = N^+(\{x\}$ and $N^-(x) = N^-(\{x\})$.

Theorem 5.12 (Rédei) *Every tournament has a Hamiltonian path.*

Proof: Let T be a tournament. We prove the result by induction on $|V(T)|$. As a base, if $|V(T)| = 1$, then the one vertex path suffices. For the inductive step, we may assume that $|V(T)| \geq 2$. Choose a vertex $v \in V(T)$ and let T^- (resp. T^+) be the subgraph of T consisting of all vertices in $N^-(v)$ (resp. $N^+(v)$) and all edges with both ends in this set. If both T^- and T^+ are not null, then each has a Hamiltonian path, say P^- and P^+ and we may form a Hamiltonian path in T by following P^- then going to the vertex v, then following P^+. A similar argument works if either T^- or T^+ is null. $\square$

Theorem 5.13 (Camion) *Every strongly connected tournament has a Hamiltonian cycle.*

Proof: Let T be a strongly connected tournament, and choose a cycle $C \subseteq T$ with $|V(C)|$ maximum. Suppose (for a contradiction) that $V(C) \neq V(T)$. If there is a vertex $v \in V(T) \setminus V(C)$ so that $N^+(v) \cap V(C) \neq \emptyset$ and $N^-(v) \cap V(C) \neq \emptyset$, then there must exist an edge $(w, x) \in E(C)$ so that $(w, v), (v, x) \in E(T)$. However, then we may use these edges to find a longer cycle. It follows that the vertices in $V(T) \setminus V(C)$ may be partitioned into $\{A, B\}$ so that every $x \in A$ has $V(C) \subseteq N^+(v)$ and every $y \in B$ has $V(C) \subseteq N^-(y)$. It follows from the strong connectivity of T that $A, B \neq \emptyset$ and that there exists an edge (y, z) with $y \in B$ and $z \in A$. However, then we may replace an edge $(w, x) \in E(C)$ with the path containing the edges $(w, y), (y, z), (z, x)$ to get a longer cycle. This contradiction completes the proof. $\square$

Path Partition: A *path partition* of a digraph D is a collection $\mathcal{P} = \{P_1, P_2, \ldots, P_k\}$ so that P_i is a directed path for $1 \leq i \leq k$ and $\{V(P_1), V(P_2), \ldots, V(P_k)\}$ is a partition of $V(D)$. We let $heads(\mathcal{P})$ ($tails(\mathcal{P})$) denote the set of vertices which are the initial (terminal) vertex in some P_i.

Lemma 5.14 (Bondy) *Let $\mathcal{P}$ be a path partition of the digraph D, and assume $|\mathcal{P}| > \alpha(D)$. Then there is a path partition $\mathcal{P}'$ of D so that $|\mathcal{P}'| = |\mathcal{P}| - 1$, and $tails(\mathcal{P}') \subseteq tails(\mathcal{P})$.*

Proof: We proceed by induction on $|V(D)|$. As a base, observe that the result is trivial when $|V(D)| = 1$. For the inductive step, note that since $\alpha(D) < |tails(\mathcal{P})|$ there must exist an edge (x, y) with $x, y \in \mathit{tails}(\mathcal{P})$. Choose i so that $y \in V(P_i)$. If $|V(P_i)| = 1$, then we may remove P_i from $\mathcal{P}$ and then append the edge (x, y) to the path containing x to get a suitable path partition. Thus, we may assume that $|V(P_i)| > 1$, and choose $w \in V(D)$ so that $(w, y) \in E(P_i)$. Now, $\mathcal{P}' = \{P_1, \ldots, P_{i-1}, P_i - y, P_{i+1}, \ldots, P_k\}$ is a path partition of $D - y$ and $\alpha(D - y) \leq \alpha(D) < |\mathcal{P}'|$, so by induction, there is a path partition $\mathcal{P}''$ of $D - y$ with $|\mathcal{P}''| = |\mathcal{P}'| - 1$ and $\mathit{tails}(\mathcal{P}'') \subseteq \mathit{tails}(\mathcal{P}')$. Since $x, w \in \mathit{tails}(\mathcal{P}')$, at least one of x, w is in $x, w \in \mathit{tails}(\mathcal{P}'')$. Since $(x, y), (w, y) \in E(D)$, we may extend $\mathcal{P}''$ to a suitable path partition of D by using one of these edges. $\square$

Theorem 5.15 (Gallai-Milgram) *Every digraph D has a path partition $\mathcal{P}$ with $|\mathcal{P}| = \alpha(D)$.*

Proof: This follows immediately from the observation that every digraph has a path partition (for instance, take each vertex as a one vertex path), and (repeated applications of) the above lemma. $\square$

Note: This is a generalization of Theorem 5.12.

Partially Ordered Set: A *partially ordered set* (or *poset*) consists of a set X and a binary relation $\prec$ which is reflexive ($x \prec x$ for every $x \in X$), antisymmetric ($x \prec y$ and $y \prec x$ imply $x = y$), and transitive ($x \prec y$ and $y \prec z$ imply $x \prec z$). We say that two points $x, y \in X$ are *comparable* if either $x \prec y$ or $y \prec x$.

Chains and Antichains: In a poset, a *chain* is a subset $A \subseteq X$ so that any two points in A are comparable. An *antichain* is a subset $B \subseteq X$ so that no two points in B are comparable.

Theorem 5.16 (Dilworth) *Let $(X, \prec)$ be a poset and let k be the size of the largest antichain. Then there is a partition of X into k chains.*

Proof: Form a digraph D with vertex set X by adding an edge from x to y whenever $x \neq y$ and $x \prec y$. Now $\alpha(D) = k$, so the Gallai-Milgram Theorem gives us a path partition of D of size k. However, the vertex set of a directed path is a chain in the poset, so this yields a partition of X into k chains. $\square$

The Ford-Fulkerson Theorem

Flows: If D is a digraph and $s, t \in V(D)$, then an (s,t)-*flow* is a map $\phi : E(D) \to \mathbb{R}$ with the property that for every $v \in V(D) \setminus \{s,t\}$ the following holds.

$$\sum_{e \in \delta^+(v)} \phi(e) = \sum_{e \in \delta^-(v)} \phi(e).$$

The *value* of ϕ is $\sum_{e \in \delta^+(s)} \phi(e) - \sum_{e \in \delta^-(s)} \phi(e)$.

Proposition 5.17 *If ϕ is an (s,t)-flow of value q, then every $X \subseteq V(D)$ with $s \in X$ and $t \notin X$ satisfies*

$$\sum_{e \in \delta^+(X)} \phi(e) - \sum_{e \in \delta^-(X)} \phi(e) = q.$$

Proof:

$$\begin{aligned} q &= \sum_{e \in \delta^+(s)} \phi(e) - \sum_{e \in \delta^-(s)} \phi(e) \\ &= \sum_{x \in X} \left(\sum_{e \in \delta^+(x)} \phi(e) - \sum_{e \in \delta^-(x)} \phi(e) \right) \\ &= \sum_{e \in \delta^+(X)} \phi(e) - \sum_{e \in \delta^-(X)} \phi(e) \end{aligned}$$

$\square$

Capacities: We shall call a weight function $c : E(D) \to \mathbb{R}^+ \cup \{\infty\}$ a *capacity function.* If $X \subseteq V(D)$, we say that $\delta^+(X)$ has *capacity* $\sum_{e \in \delta^+(X)} \phi(e)$.

Admissible Flows: An (s,t)-flow ϕ is *admissible* if $0 \leq \phi(e) \leq c(e)$ for every edge e.

Augmenting Paths: Let c be a capacity function and $\phi : E(D) \to \mathbb{R}$ an admissible (s,t)-flow. A path P from u to v is called *augmenting* if for every edge $e \in E(P)$, either e is traversed in the forward direction and $\phi(e) < c(e)$ or e is traversed in the backward direction and $\phi(e) > 0$.

Theorem 5.18 (Ford-Fulkerson) *Let D be a digraph, let $s, t \in V(D)$, and let c be a capacity function. Then the maximum value of an (s,t)-flow is equal to the minimum capacity of a cut $\delta^+(X)$ with $s \in X$ and $t \notin X$. Furthermore, if c is integer valued, then there exists a flow of maximum value ϕ which is also integer valued.*

Proof: It follows immediately from Proposition 19.1 that every admissible (s,t)-flow has value less than or equal to the capacity of any cut $\delta^+(X)$ with $s \in X$ and $t \notin X$.

We shall prove the other direction of this result only for capacity functions $c : E(D) \to \mathbb{Q}^+$ (although it holds in general). For every edge e, let $\frac{p_e}{q_e}$ be a reduced fraction equal to $c(e)$, and let n be the least common multiple of $\{q_e : e \in E(D)\}$. We shall prove that there exists a flow $\phi : E(D) \to \mathbb{Q}^+$ so that $\phi(e)$ can be expressed as a fraction with denominator n for every edge e. To do this, choose a flow ϕ with this property of maximum value. Define the set X as follows.

$$X = \{v \in V(D) : \text{there is an augmenting path from } s \text{ to } v\}$$

If $t \in X$, then there exists an augmenting path P from s to t. However, then we may modify the flow ϕ to produce a new admissible flow of greater value by increasing the flow by $\frac{1}{n}$ on every forward edge of P and decreasing the flow by $\frac{1}{n}$ on every backward edge of P. Since this new flow would contradict the choice of ϕ, it follows that $t \notin X$.

It follows from the definition of X that every edge $e \in \delta^+(X)$ satisfies $\phi(e) = c(e)$ and every edge $f \in \delta^-(X)$ satisfies $\phi(f) = 0$. Thus, our flow ϕ has value equal to the capacity of the cut $\delta^+(X)$ and the proof is complete. $\square$

Note: The above proof for rational valued flows combined with a simple convergence argument yields the proof in general. However, the algorithm inherent in the above proof does not yield a finite algorithm for finding a flow of maximum value for arbitrary capacity functions.

Corollary 5.19 (edge-digraph version of Menger) *Let D be a digraph and let $s,t \in V(D)$. Then exactly one of the following holds:*

(i) *There exist k pairwise edge disjoint directed paths $P_1, \ldots, P_k$ from s to t.*

(ii) *There exists $X \subseteq V(D)$ with $s \in X$ and $t \notin X$ so that $|\delta^+(X)| < k$*

Proof: It is immediate that (i) and (ii) are mutually exclusive, so it suffices to show that at least one holds. Define a capacity function $c : E(D) \to \mathbb{R}$ by the rule that $c(e) = 1$ for every edge e. Apply the Ford-Fulkerson Theorem to choose an admissible integer valued (s,t)-flow $\phi : E(D) \to \mathbb{Z}$ and a cut $\delta^+(X)$ with $s \in X$ and $t \notin X$ so that the value of ϕ and the capacity of $\delta^+(X)$ are both equal to the integer q. Now, let $H = D - \{e \in E(D) : \phi(e) = 0\}$. Then H is a digraph with the property that $\delta^+_H(s) - \delta^-_H(s) = q = \delta^-_H(t) - \delta^+_H(t)$ and $\delta^+_H(v) = \delta^-_H(v)$ for every $v \in V(H) \setminus \{s,t\}$. By Problem 3 of Homework 10, we find that H contains q edge disjoint directed paths from s to t. So, if $q \leq k$, then (i) holds, and if $q > k$ (ii) holds. $\square$