Paths, trails, walks

We will define basic notions of graph theory and use them to define connectedness (intuitively, that the graph consists of one piece). We always assume that the graph is finite.

Let us start with the definition of a subgraph.

Definition. Let $G = (V, E)$ be a graph. A subgraph $G' = (V', E')$ is a graph with $V' \subseteq V$ and $E' \subseteq E$. Of course, we have to assume that for every edge $e' \in E'$, the endpoints are in V'. The subgraph is called a *spanning* subgraph if $V = V'$. If E' consists of all edges in E, whose endpoints are in V', then G' is called an *induced* subgraph.

This means that in case of a spanning subgraph deleting some edges of E yields E', in case of an induced subgraph we delete some nodes (points, vertices). In case of an arbitrary subgraphs, we can delete both vertices and edges.

Definition. Let G be a simple graph. The *complement* of G is the graph $\bar{G}$, where two points are joined by an edge if and only if there is no edge between them in G.

This means that the number of edges of G plus the number of edges in $\bar{G}$ is $\binom{n}{2}$.

Definition. Let $G = (V, E)$ be a graph. A *walk* is a set of adjacent edges. More precisely, it is a sequence $v_1, e_1, v_2, e_2 \dots v_{s-1}, e_{s-1}, v_s$, where the endpoints of e_i are v_i and v_{i+1}, $i = 1, \dots, s-1$. The walk is *closed* if $v_1 = v_s$ (*open* otherwise). The number of edges in the sequence is called the *length* of the walk (so it is $s - 1$). If the graph is simple, then we do not need to list the edges. (Actually, even in a non-simple graph we need to list an edge, if there are parallel edges joining the same endpoints.) If we have a closed walk, then one can start the sequence at any of its vertices.

If the edges in a walk are pairwise distinct, then we call it a *trail.* A closed trail is also called a *circuit.* If there is no repetition among the vertices of the sequence (all the v_i's are different), then we call it a *path.* If $v_1 = v_s$, but there is no other repetition, then it is called a *cycle.* (So, in case of a cycle, $v_1, \dots v_{s-1}$ are pairwise different and $v_1 = v_s$.)

We say that that walk (trail, path) joins v_1 and v_s. We also call v_1 the starting points of the walk (trail, path), and v_s the endpoint of it.

Note that a walk (trail, path) can be reversed (by listing its vertices, edges backwards) and two walks can be joined (or concatenated) if the endpoint of the first one is the starting point of the second one (by just concatenating the two sequences, and listing the endpoint of the first walk once). Note that the concatenation of two paths (or trails) is not necessarily a path (or trail) in general, just a walk. Note also that deleting an edge from a cycle (circuit) gives path (open trail), whose starting point and endpoint are the endpoints of the deleted edge. The is a „degenerate" closed walk, the concatenation of a trail and its reverse trail. In many cases we have to exclude this possibility.

Proposition. *Assume that there is walk between two points $u \neq v \in V$. Then there is also a path between u and v. If there is a closed walk C in G, which contains an edge which appears once, then there is a cycle in G.*

The first assertion follows from the fact that in case of having $v_i = v_j$, we can shorten the walk by deleting the sequence between v_i and v_j (keeping one of them). If e is an edge appearing once, then deleting that edge from the closed walk yields an open walk between its endpoints. This open walk can be shorted to a path by the first assertion. Together with the edge e it gives a cycle.

Proposition. *If the degree of each node is at least two, then the graph contains a cycle.*

Let us start from any node. We try to build a path greedily. So, at any node we try to find an edge whose other endpoint is a new node (not on the path we constructed so far). Since the graph is finite, this is not possible at a certain vertex. At this point, all the edges emanating from this vertex go to a vertex of the path built so far. Hence we get a cycle.

Definition. The graph G is connected, if there is a walk (or path) between any two of its vertices.

Proposition. *If G is connected, C is a cycle in G, e is an edge of C, then after deleting the edge e, the resulting graph (usually denoted as $G - e$) is connected.*

The next proposition tells us how to imagine graphs that are not connected.

Proposition. *Every graph is a vertex disjoint union of connected induced subgraphs.*

Let us define a binary relation on the vertices: $u \sim v$ if there is a walk between u and v. This is an equivalence relation, so it partitions V. The graph is the union of the indeced subgraphs on the equivalence classes of $\sim$.

Definition. The induced subgraphs mentioned in the previous proposition are called the *components* (or: connectivity components) of G.

One can also define connected components as maximal connected subgraphs. To imagine a graph that is not connected, we have to consider disjoint sets and the edges of the graph are in these disjoint sets, there is no edge between the sets.

Proposition. *At least one of the graphs G and $\bar{G}$ is connected.*

If G is not connected, then it has more than one components. In the complementer graph, two points are connected by an edge, if they are in different components of G. If they (say u and v) are in the same component of G, take a vertex w in another component. Then u, w, v will be a path between u and v in $\bar{G}$.

In some of the classical problems in graph theory we want to find walks / trails / paths etc. that visit all the vertices. The first problem in graph theory was the problem of walking through the bridges of Königsberg, solved by Euler.

Definition. A trail (open or closed) using all the edges of a graph is called an *(open or closed) Eulerian trail*. Sometimes it is called an Eulerian walk (or circuit if we speak of a closed one).

So, such an Eulerian trail uses each edge of the graph exactly once.

Theorem (Euler)**.** *Assume that G is connected. Then there is a closed Eulerian trail in G if and only if each vertex of G has even degree. There is an open Eulerian trail if and only if all but two degrees are even. (In this case any Eulerian trail starts at one of the two vertices of odd degree and ends at the other one.)*

Note also that the theorem is valid also for non-simple graphs (multigraphs), so we may allow parallel edges and loops. The condition of connectivity can be dropped. A non-connected graph has a (closed or open) Eulerian trail if and only if the same condition on the degrees holds, and the connectivity components of the graph are isolated vertices with one exception.

If we want to visit the vertices of a graph exactly once, then we are interested in so-called Hamiltonian paths or Hamiltonian cycles.

Definition. A cycle (path) containing every vertex is called a *Hamiltonian cycle* (*Hamiltonian path*).

Such a graph can be easily imagined: it contains a cycle (path) and some extra edges between certain vertices of the cycle (path). The next proposition gives an easy necessary condition for the existence of Hamiltonian cycles/paths.

Proposition. *If a graph contains a Hamiltonian cycle, then for every k we have that $G - \{v_1, \ldots, v_k\}$ has at most k components for every $v_1, \ldots, v_k \in V$. So, by deleting k points arbitrarily, the remaining graph has at most k components.*

If a graph contains a Hamiltonian path, then for every k we have that $G - \{v_1, \ldots, v_k\}$ has at most $k+1$ components for every $v_1, \ldots, v_k \in V$. So, by deleting k points arbitrarily, the remaining graph has at most $k+1$ components.

If we just consider the path (or cycle) but not the extra edges, and delete k points, then this path (or cycle) splits in at most $k+1$ (or k) subpaths.

Unfortunately, there is no necessary and sufficient condition for the existence of a Hamiltonian cycle/path. Actually, it is computanionally difficult to find one. The next condition is a sufficient condition.

Theorem (Gábor Dirac). *Let G be a simple graph on n vertices, $n \geq 3$. Assume that every vertex has degree at least $n/2$. Then the graph contains a Hamiltonian cycle.*

Ore generalized this to the following: If two vertices x, y are not connected by an edge, then $\deg(x) + \deg(y) \geq n$, then the graph contains a Hamiltonian cycle.

We can also note that the conditions are not necessary, if the graph is just a cycle (of length at least 5), then the conditions in Dirac's or Ore's thm are not satisfied.