Math 530 Spring 2022, Lecture 28: More about paths

website: `https://www.cip.ifi.lmu.de/~grinberg/t/22s`

1. More about paths (cont'd)

Today, we shall discuss some more obscure properties of paths in digraphs and graphs.

1.1. The Gallai–Milgram theorem

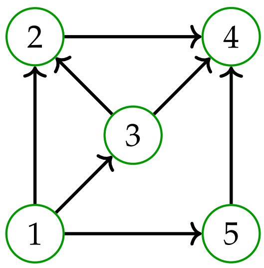
In order to state the first of these properties, we need the following three definitions:

Definition 1.1.1. Two vertices u and v of a multidigraph D are said to be **adjacent** if they are adjacent in the undirected graph D^{und}. (In other words, they are adjacent if and only if D has an arc with source u and target v or an arc with source v and target u.)

Definition 1.1.2. An **independent set** of a multidigraph D means a subset S of $\operatorname{V}(D)$ such that no two elements of S are adjacent. In other words, it means an independent set of the undirected graph D^{und}.

Definition 1.1.3. A **path cover** of a multidigraph D means a set of paths of D such that each vertex of D is contained in exactly one of these paths.

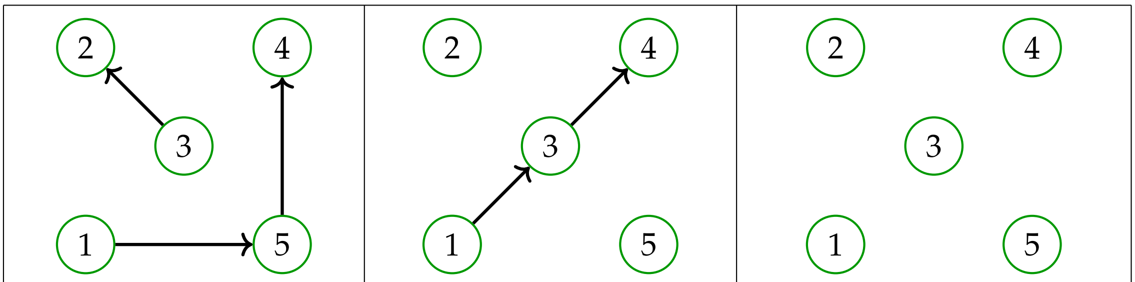
Example 1.1.4. Let D be the following digraph:



Then, $\{(1,*,5,*,4), (3,*,2)\}$ is a path cover of D (we are again writing asterisks for the arcs, since the arcs of D are uniquely determined by their sources and their targets). Another path cover of D is $\{(1,*,3,*,4), (2), (5)\}$. Yet another path cover of D is $\{(1), (2), (3), (4), (5)\}$. There are many more.

Note that the set $\{(1,*,5,*,4), (3,*,2,*,4)\}$ is not a path cover of D, since the vertex 4 is contained in two (not one) of its paths.

Let us draw the three path covers we have mentioned (by simply drawing the arcs of the paths they contain, while omitting all other arcs of D):



$\{(1,*,5,*,4)\,,\ (3,*,2)\}$	$\{(1,*,3,*,4)\,,\ (2)\,,\ (5)\}$	$\{(1)\,,\ (2)\,,\ (3)\,,\ (4)\,,\ (5)\}$

(Note that we have already seen path covers of a "complete" simple digraph $(V,\ V \times V)$ in Lecture 11; we called them "path covers of V".)

Remark 1.1.5. Let D be a digraph. A path cover of D consisting of only 1 path is the same as a Hamiltonian path of D. (More precisely: A single path $\mathbf{p}$ forms a path cover $\{\mathbf{p}\}$ of D if and only if $\mathbf{p}$ is a Hamiltonian path.)

Now, the **Gallai–Milgram theorem** states the following:

Theorem 1.1.6 (Gallai–Milgram theorem)**.** Let D be a loopless digraph. Then, there exist a path cover $\mathcal{P}$ of D and an independent set S of D such that S has exactly one vertex from each path in $\mathcal{P}$ (in other words, for each path $\mathbf{p} \in \mathcal{P}$, exactly one vertex of $\mathbf{p}$ belongs to S).

Example 1.1.7. Let D be the digraph from Example 1.1.4. Then, Theorem 1.1.6 tells us that there exist a path cover $\mathcal{P}$ of D and an independent set S of D such that S has exactly one vertex from each path in $\mathcal{P}$. For example, we can take $\mathcal{P} = \{(1,*,5,*,4)\,,\ (3,*,2)\}$ and $S = \{5,3\}$.

We will now prove Theorem 1.1.6, following Diestel's book [Dieste17, Theorem 2.5.1]:

Proof of Theorem 1.1.6. Write the multidigraph D as $D = (V, A, \varphi)$. We introduce a notation:

- If $\mathcal{P}$ is a path cover of D, then a **cross-cut** of $\mathcal{P}$ means a subset S of V that contains exactly one vertex from each path in $\mathcal{P}$.

Thus, the claim we must prove is saying that there exist a path cover $\mathcal{P}$ of D and an independent cross-cut of $\mathcal{P}$.

We will show something stronger:

> *Claim 1:* Any minimum-size path cover of D has an independent cross-cut.

Note that the size of a path cover means the number of paths in it. Thus, a minimum-size path cover means a path cover with the smallest possible number of paths.

We will show something even stronger than Claim 1. To state this stronger claim, we need more notations:

- If $\mathcal{P}$ is a path cover, then $\operatorname{Ends}\mathcal{P}$ means the set of the ending points of all paths in $\mathcal{P}$. Note that $|\operatorname{Ends}\mathcal{P}| = |\mathcal{P}|$.
- A path cover $\mathcal{P}$ is said to be **end-minimal** if no proper subset of $\operatorname{Ends}\mathcal{P}$ can be written as $\operatorname{Ends}\mathcal{Q}$ for a path cover $\mathcal{Q}$.

Example 1.1.8. For instance, if D is as in Example 1.1.4, and if

$$\begin{aligned}\mathcal{P} &= \{(1,*,5,*,4)\,,\ (3,*,2)\}\,,\\ \mathcal{Q} &= \{(1,*,3,*,4)\,,\ (2)\,,\ (5)\}\,,\\ \mathcal{R} &= \{(1)\,,\ (2)\,,\ (3)\,,\ (4)\,,\ (5)\}\end{aligned}$$

are the three path covers from Example 1.1.4, then

$$\operatorname{Ends}\mathcal{P} = \{4,2\}\,, \qquad \operatorname{Ends}\mathcal{Q} = \{4,2,5\}\,, \qquad \operatorname{Ends}\mathcal{R} = \{1,2,3,4,5\}\,,$$

which shows immediately that neither $\mathcal{Q}$ nor $\mathcal{R}$ is end-minimal (since $\operatorname{Ends}\mathcal{P}$ is a proper subset of each of $\operatorname{Ends}\mathcal{Q}$ and $\operatorname{Ends}\mathcal{R}$). It is easy to see that $\mathcal{P}$ is end-minimal (and also minimum-size).

Back to the general case. Clearly, any minimum-size path cover of D is also end-minimal[1]. Thus, the following claim is stronger than Claim 1:

> *Claim 2:* Any end-minimal path cover of D has an independent cross-cut.

It is Claim 2 that we will be proving.[2]

[*Proof of Claim 2:* We proceed by induction on $|V|$.

Base case: Claim 2 is obvious when $|V| = 0$ (since $\varnothing$ is an independent cross-cut in this case).

Induction step: Consider a multidigraph $D = (V, A, \psi)$ with $|V| = N$. Assume (as the induction hypothesis) that Claim 2 is already proved for any multidigraph with $N-1$ vertices.

---

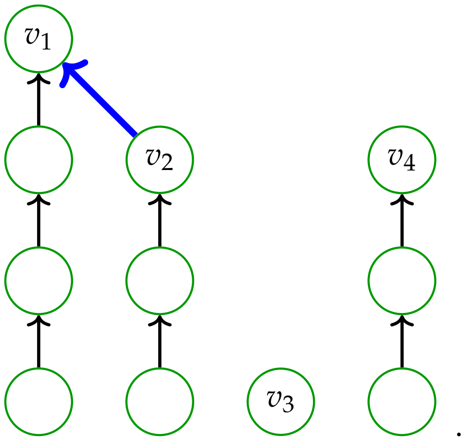
[1] *Proof.* Let $\mathcal{P}$ be a minimum-size path cover of D. If $\mathcal{P}$ was not end-minimal, then there would be a path cover $\mathcal{Q}$ with $|\operatorname{Ends}\mathcal{Q}| < |\operatorname{Ends}\mathcal{P}|$ and therefore $|\mathcal{Q}| = |\operatorname{Ends}\mathcal{Q}| < |\operatorname{Ends}\mathcal{P}| = |\mathcal{P}|$; but this would contradict the fact that $\mathcal{P}$ is minimum-size. Hence, $\mathcal{P}$ is end-minimal.

[2] On a sidenote: Is Claim 2 really stronger than Claim 1? Yes, because it can happen that some end-minimal path cover fails to be minimum-size. For example, the path cover

Let $\mathcal{P}$ be an end-minimal path cover of D. We must show that $\mathcal{P}$ has an independent cross-cut.

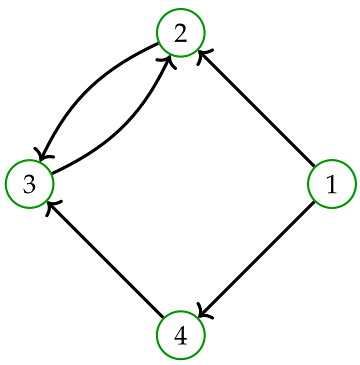
Let $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k$ be the paths in $\mathcal{P}$ (listed without repetitions), and let $v_1, v_2, \ldots, v_k$ be their respective ending points. Thus, $\{v_1, v_2, \ldots, v_k\} = \operatorname{Ends} \mathcal{P}$ and $k = |\operatorname{Ends} \mathcal{P}|$.

Recall that we must find an independent cross-cut of $\mathcal{P}$. If the set $\{v_1, v_2, \ldots, v_k\}$ is independent, then we are done (since this set $\{v_1, v_2, \ldots, v_k\}$ is clearly a cross-cut of $\mathcal{P}$). Thus, we WLOG assume that this is not the case. Hence, there is an arc from some vertex v_i to some vertex v_j. These two vertices v_i and v_j are distinct (because D is loopless). Since we can swap our paths $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k$ (and thus their ending points $v_1, v_2, \ldots, v_k$) at will, we can thus WLOG assume that $i = 2$ and $j = 1$. Assume this. Thus, there is an arc from v_2 to v_1. We shall refer to this arc as the *blue arc*, and we will draw it accordingly:[3]



We can extend the path $\mathbf{p}_2$ beyond its ending point v_2 by inserting the blue arc and the vertex v_1 at its end. This results in a new path, which we denote by $\mathbf{p}_2 + v_1$; this path has ending point v_1.

$\{(1, *, 2, *, 3), (4)\}$ in the digraph

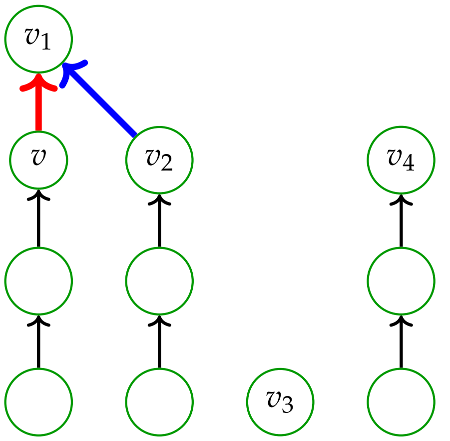


has this property.

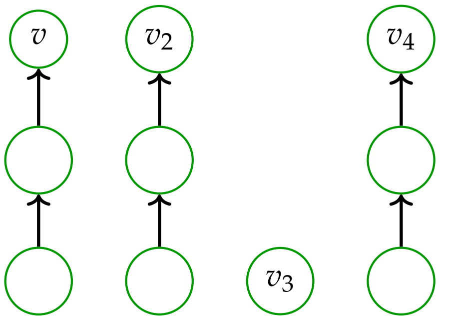
[3]This picture illustrates just one representative case, with $k = 4$. The four columns (from left to right) are the four paths $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{p}_4$. Of course, the digraph D can have many more arcs than we have drawn on this picture, but we are not interested in them right now.

If v_1 is the only vertex on the path $\mathbf{p}_1$ (that is, if the path $\mathbf{p}_1$ has length 0), then we can therefore replace the path $\mathbf{p}_2$ by $\mathbf{p}_2 + v_1$ and remove the length-0 path $\mathbf{p}_1$ from our path cover $\mathcal{P}$, and we thus obtain a new path cover $\mathcal{Q}$ such that Ends $\mathcal{Q}$ is a proper subset of Ends $\mathcal{P}$. But this is impossible, since we assumed that $\mathcal{P}$ is end-minimal. Therefore, v_1 is not the only vertex on $\mathbf{p}_1$.

Thus, let v be the second-to-last vertex on $\mathbf{p}_1$ (that is, the vertex that is immediately followed by v_1). Then, the path $\mathbf{p}_1$ contains an arc from v to v_1. We shall refer to this arc as the *red arc*, and we will draw it accordingly:



Let D' be the digraph $D \setminus v_1$ (that is, the digraph obtained from D by removing the vertex v_1 and all arcs that have v_1 as source or target). Let $\mathbf{p}'_1$ be the result of removing the vertex v_1 and the red arc from the path $\mathbf{p}_1$. Then, $\mathcal{P}' := \{\mathbf{p}'_1, \mathbf{p}_2, \mathbf{p}_3, \ldots, \mathbf{p}_k\}$ is a path cover of D'. Note that the path $\mathbf{p}'_1$ has ending point v (since it is obtained from $\mathbf{p}_1$ by removing the last vertex and the last arc, but we know that the second-to-last vertex on $\mathbf{p}_1$ is v), whereas the paths $\mathbf{p}_2, \mathbf{p}_3, \ldots, \mathbf{p}_k$ have ending points $v_2, v_3, \ldots, v_k$. Thus, $\operatorname{Ends}(\mathcal{P}') = \{v, v_2, v_3, \ldots, v_k\}$. Here is an illustration of the digraph $D' = D \setminus v_1$ and its path cover $\mathcal{P}'$:



Consider the path cover $\mathcal{P}'$ of D'. If we can find an independent cross-cut of $\mathcal{P}'$, then we will be done, because any such cross-cut will also be an independent cross-cut of our original path cover $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_k\} = \mathcal{P}$. Since the digraph $D \setminus v_1$ has $N - 1$ vertices[4], we can find such an independent cross-

[4]because the digraph D has $|V| = N$ vertices

cut by our induction hypothesis if we can prove that the path cover $\mathcal{P}'$ is end-minimal (as a path cover of D').

So let us prove this now. Indeed, assume the contrary. Thus, D' has a path cover $\mathcal{Q}'$ such that $\operatorname{Ends}(\mathcal{Q}')$ is a proper subset of $\operatorname{Ends}(\mathcal{P}')$. Consider this $\mathcal{Q}'$. Note that[5]

$$\operatorname{Ends}\left(\mathcal{Q}'\right) \subsetneq \operatorname{Ends}\left(\mathcal{P}'\right) = \{v, v_2, v_3, \ldots, v_k\}.$$

As a consequence, $|\operatorname{Ends}(\mathcal{Q}')| < |\{v, v_2, v_3, \ldots, v_k\}| = k$.

Now, we are in one of the following three cases:

Case 1: We have $v \in \operatorname{Ends}(\mathcal{Q}')$.

Case 2: We have $v \notin \operatorname{Ends}(\mathcal{Q}')$ but $v_2 \in \operatorname{Ends}(\mathcal{Q}')$.

Case 3: We have $v \notin \operatorname{Ends}(\mathcal{Q}')$ and $v_2 \notin \operatorname{Ends}(\mathcal{Q}')$.

Let us consider these cases one by one:

- We first consider Case 1. In this case, we have $v \in \operatorname{Ends}(\mathcal{Q}')$. In other words, some path $\mathbf{p} \in \mathcal{Q}'$ ends at v. Let us extend this path $\mathbf{p}$ beyond v by inserting the red arc and the vertex v_1 at its end. Thus, we obtain a path of D, which we call $\mathbf{p} + v_1$. Replacing $\mathbf{p}$ by $\mathbf{p} + v_1$ in $\mathcal{Q}'$, we obtain a path cover $\mathcal{Q}$ of D such that $\operatorname{Ends}\mathcal{Q}$ is a proper subset of $\operatorname{Ends}\mathcal{P}$ [6]. But this contradicts the fact that $\mathcal{P}$ is end-minimal. Thus, we have obtained a contradiction in Case 1.

- Next, we consider Case 2. In this case, we have $v \notin \operatorname{Ends}(\mathcal{Q}')$ but $v_2 \in \operatorname{Ends}(\mathcal{Q}')$. Combining $\operatorname{Ends}(\mathcal{Q}') \subseteq \{v, v_2, v_3, \ldots, v_k\}$ with $v \notin \operatorname{Ends}(\mathcal{Q}')$, we obtain

$$\operatorname{Ends}\left(\mathcal{Q}'\right) \subseteq \{v, v_2, v_3, \ldots, v_k\} \setminus \{v\} = \{v_2, v_3, \ldots, v_k\}.$$

  From $v_2 \in \operatorname{Ends}(\mathcal{Q}')$, we see that some path $\mathbf{p} \in \mathcal{Q}'$ ends at v_2. Let us extend this path $\mathbf{p}$ beyond v_2 by inserting the blue arc and the vertex v_1 at its end. Thus, we obtain a path of D, which we call $\mathbf{p} + v_1$. Replacing $\mathbf{p}$ by $\mathbf{p} + v_1$ in $\mathcal{Q}'$, we obtain a path cover $\mathcal{Q}$ of D such that $\operatorname{Ends}\mathcal{Q}$ is a proper

[5] The symbol "$\subsetneq$" (note that the stroke only crosses the straight line, not the curved one) means "proper subset of".

[6] *Proof.* We obtained $\mathcal{Q}$ from $\mathcal{Q}'$ by replacing $\mathbf{p}$ by $\mathbf{p} + v_1$. As a consequence of this replacement, the ending point v of $\mathbf{p}$ has been replaced by the ending point v_1 of $\mathbf{p} + v_1$. Thus,

$$\begin{aligned} \operatorname{Ends}\mathcal{Q} &= \underbrace{\left(\operatorname{Ends}\left(\mathcal{Q}'\right) \setminus \{v\}\right)}_{\substack{\subseteq \{v_2, v_3, \ldots, v_k\} \\ (\text{since } \operatorname{Ends}(\mathcal{Q}') \subseteq \{v, v_2, v_3, \ldots, v_k\})}} \cup \{v_1\} \\ &\subseteq \{v_2, v_3, \ldots, v_k\} \cup \{v_1\} = \{v_1, v_2, \ldots, v_k\} = \operatorname{Ends}\mathcal{P}. \end{aligned}$$

For the same reason, we have $|\operatorname{Ends}\mathcal{Q}| = |\operatorname{Ends}(\mathcal{Q}')| < k = |\operatorname{Ends}\mathcal{P}|$, so that $\operatorname{Ends}\mathcal{Q} \neq \operatorname{Ends}\mathcal{P}$. Combining this with $\operatorname{Ends}\mathcal{Q} \subseteq \operatorname{Ends}\mathcal{P}$, we conclude that $\operatorname{Ends}\mathcal{Q}$ is a proper subset of $\operatorname{Ends}\mathcal{P}$.

subset of Ends $\mathcal{P}$ [7]. But this contradicts the fact that $\mathcal{P}$ is end-minimal. Thus, we have obtained a contradiction in Case 2.

- Finally, we consider Case 3. In this case, we have $v \notin \operatorname{Ends}(\mathcal{Q}')$ and $v_2 \notin \operatorname{Ends}(\mathcal{Q}')$. Combining this with $\operatorname{Ends}(\mathcal{Q}') \subseteq \{v, v_2, v_3, \ldots, v_k\}$, we obtain

$$\operatorname{Ends}(\mathcal{Q}') \subseteq \{v, v_2, v_3, \ldots, v_k\} \setminus \{v, v_2\} = \{v_3, v_4, \ldots, v_k\},$$

so that $|\operatorname{Ends}(\mathcal{Q}')| \leq |\{v_3, v_4, \ldots, v_k\}| = k - 2$. Now, adding the trivial path (v_1) to $\mathcal{Q}'$ yields a path cover $\mathcal{Q}$ of D such that Ends $\mathcal{Q}$ is a proper subset of Ends $\mathcal{P}$ [8]. But this contradicts the fact that $\mathcal{P}$ is end-minimal. Thus, we have found a contradiction in Case 3.

So we have obtained a contradiction in each case. Thus, our assumption was false. This shows that the path cover $\mathcal{P}'$ is end-minimal. As we already said above, this allows us to apply the induction hypothesis to D' instead of D, and conclude that the end-minimal path cover $\mathcal{P}'$ of D' has an independent cross-cut. This independent cross-cut is clearly an independent cross-cut of $\mathcal{P}$ as well, and thus we have shown that $\mathcal{P}$ has an independent cross-cut. This proves Claim 2.]

As explained above, this completes the proof of Theorem 1.1.6. $\square$

Here are two simple applications of the Gallai–Milgram theorem:

- Remember the Easy Rédei theorem, which we proved in Lecture 12. It says that each tournament has a Hamiltonian path.

  We can now prove it again using the Gallai–Milgram theorem:

[7] *Proof.* We obtained $\mathcal{Q}$ from $\mathcal{Q}'$ by replacing $\mathbf{p}$ by $\mathbf{p} + v_1$. As a consequence of this replacement, the ending point v_2 of $\mathbf{p}$ has been replaced by the ending point v_1 of $\mathbf{p} + v_1$. Thus,

$$\begin{aligned}\operatorname{Ends}\mathcal{Q} &= \underbrace{\left(\operatorname{Ends}(\mathcal{Q}') \setminus \{v_2\}\right)}_{\substack{\subseteq \{v_3, v_4, \ldots, v_k\} \\ (\text{since } \operatorname{Ends}(\mathcal{Q}') \subseteq \{v_2, v_3, \ldots, v_k\})}} \cup \{v_1\} \\ &\subseteq \{v_3, v_4, \ldots, v_k\} \cup \{v_1\} = \{v_1, v_3, v_4, \ldots, v_k\} \\ &\subsetneq \{v_1, v_2, \ldots, v_k\} = \operatorname{Ends}\mathcal{P}.\end{aligned}$$

In other words, Ends $\mathcal{Q}$ is a proper subset of Ends $\mathcal{P}$.

[8] *Proof.* We obtained $\mathcal{Q}$ from $\mathcal{Q}'$ by adding the trivial path (v_1), whose ending point is v_1. Thus,

$$\begin{aligned}\operatorname{Ends}\mathcal{Q} &= \underbrace{\operatorname{Ends}(\mathcal{Q}')}_{\subseteq \{v_3, v_4, \ldots, v_k\}} \cup \{v_1\} \subseteq \{v_3, v_4, \ldots, v_k\} \cup \{v_1\} = \{v_1, v_3, v_4, \ldots, v_k\} \\ &\subsetneq \{v_1, v_2, \ldots, v_k\} = \operatorname{Ends}\mathcal{P}.\end{aligned}$$

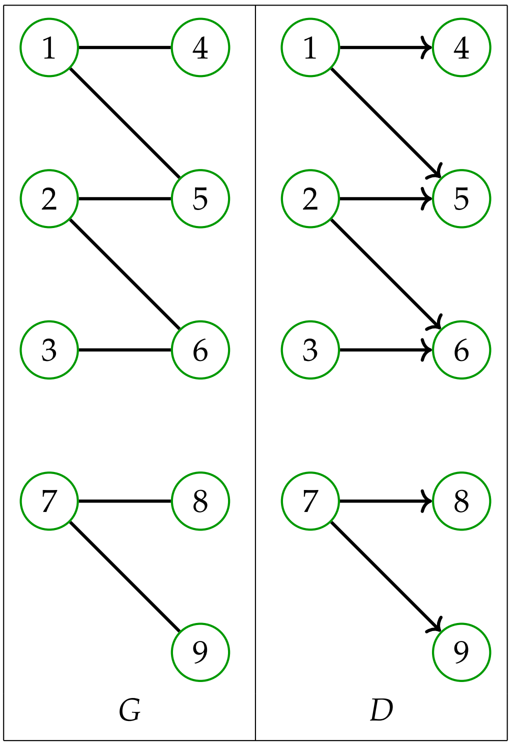
In other words, Ends $\mathcal{Q}$ is a proper subset of Ends $\mathcal{P}$.

New proof of the Easy Rédei theorem: Indeed, let D be a tournament. The Gallai–Milgram theorem shows that D has a path cover with an independent cross-cut[9]. Consider this path cover and this cross-cut. But since D is a tournament, any independent set of D has size ≤ 1. Thus, our independent cross-cut must have size ≤ 1. Hence, our path cover must consist of 1 path only (because the size of the path cover equals the size of its cross-cut). But this means that it is a Hamiltonian path (or, more precisely, it consists of a single path, which is necessarily a Hamiltonian path). Hence, D has a Hamiltonian path. So we have proved the Easy Rédei theorem again.

- Less obviously, Hall's Marriage Theorem (Theorem 1.3.4 in Lecture 24) and the Hall-König matching theorem (Theorem 1.4.7 in Lecture 24) can be proved again using Gallai–Milgram. Here is how:

  New proof of the Hall-König matching theorem: Let (G, X, Y) be a bipartite graph.

  Let D be the digraph obtained from G by directing each edge so that it goes from X to Y (in other words, each edge with endpoints $x \in X$ and $y \in Y$ becomes an arc with source x and target y). Thus, in the digraph D, no vertex can simultaneously be the source of some arc and the target of some arc. Thus, any path of D has length ≤ 1. Here is an illustration of a bipartite graph (G, X, Y) (drawn as agreed in Lecture 24) and the corresponding digraph D:

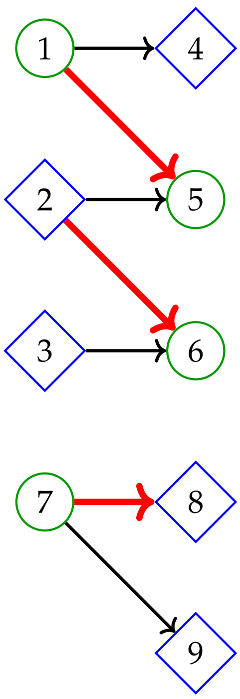


.

[9] See the above proof of Theorem 1.1.6 for the definition of a "cross-cut".

As we said, any path of D has length ≤ 1. Thus, any path of D corresponds either to a vertex of G or to an edge of G (depending on whether its length is 0 or 1). Hence, any path cover $\mathcal{P}$ of D necessarily consists of length-0 paths (corresponding to vertices of G) and length-1 paths (corresponding to edges of G); moreover, the edges of $\mathcal{P}$ (that is, the edges corresponding to the length-1 paths in $\mathcal{P}$) form a matching of G, and the vertices of $\mathcal{P}$ (that is, the vertices corresponding to length-0 paths in $\mathcal{P}$) are precisely the vertices that are not matched in this matching.

Now, Theorem 1.1.6 shows that there exist a path cover $\mathcal{P}$ of D and an independent cross-cut S of $\mathcal{P}$. Consider these $\mathcal{P}$ and S. For the purpose of illustration, let us draw a path cover $\mathcal{P}$ (by marking the arcs in red) and an independent cross-cut S of $\mathcal{P}$ (by drawing each vertex $s \in S$ as a blue diamond instead of a green circle):



.

We have $|S| = |X \cap S| + |Y \cap S|$ (since the set S is the union of its two disjoint subsets $X \cap S$ and $Y \cap S$).

The set S is an independent set of the digraph D, thus also an independent set of the graph $D^{\text{und}} = G$. From this, we easily obtain $N(X \cap S) \subseteq Y \setminus S$ (since (G, X, Y) is a bipartite graph)[10]. Therefore, $|N(X \cap S)| \leq |Y \setminus S|$, so

[10] *Proof.* Let $v \in N(X \cap S)$. Thus, v is a vertex with a neighbor in $X \cap S$. Let x be this neighbor. Then, $x \in X \cap S \subseteq X$, so that the vertex v has a neighbor in X (namely, x). Since (G, X, Y) is a bipartite graph, this entails that $v \in Y$. Furthermore, we have $x \in X \cap S \subseteq S$. If we had $v \in S$, then the set S would contain two adjacent vertices (namely, v and x), which would contradict the fact that S is an independent set of G. Thus, we have $v \notin S$. Combining $v \in Y$

that $|Y \setminus S| \geq |N(X \cap S)|$. Hence,

$$|Y| = \underbrace{|Y \setminus S|}_{\geq |N(X \cap S)|} + |Y \cap S| \geq |N(X \cap S)| + \underbrace{|Y \cap S|}_{\substack{=|S|-|X \cap S| \\ (\text{since } |S|=|X\cap S|+|Y\cap S|)}}$$
$$= |N(X \cap S)| + |S| - |X \cap S|. \tag{1}$$

Now, let M be the set of edges of G corresponding to the length-1 paths in our path cover $\mathcal{P}$. As we already mentioned, this set M is a matching of G (since two paths in $\mathcal{P}$ cannot have a vertex in common). The vertices that are not matched in M are precisely the vertices that don't belong to any of the length-1 paths in $\mathcal{P}$; in other words, they are the vertices that belong to length-0 paths in $\mathcal{P}$ (since $\mathcal{P}$ is a path cover, and any path has length ≤ 1). We let p be the number of such vertices that lie in X, and we let q be the number of such vertices that lie in Y.

Thus, our path cover $\mathcal{P}$ contains exactly $p+q$ length-0 paths: namely, p length-0 paths consisting of a vertex in X and q length-0 paths consisting of a vertex in Y. Hence, the path cover $\mathcal{P}$ consists of $|M| + p + q$ paths in total (since it contains $|M|$ many length-1 paths). The set S contains exactly one vertex from each of these $|M| + p + q$ paths (since S is a cross-cut of $\mathcal{P}$); therefore,

$$|S| = |M| + p + q. \tag{2}$$

Each vertex $y \in Y$ that is matched in M belongs to exactly one M-edge (namely, to its M-edge), and conversely, each M-edge contains exactly one vertex in Y (which, of course, is matched in M). Thus, the map

$$\{\text{vertices in } Y \text{ that are matched in } M\} \to M,$$
$$y \mapsto (\text{the } M\text{-edge of } y)$$

is a bijection. Hence, the bijection principle yields

$$(\# \text{ of vertices in } Y \text{ that are matched in } M) = |M|. \tag{3}$$

On the other hand, the set Y contains exactly q vertices that are not matched in M (by the definition of q). Therefore, Y contains exactly $|Y| - q$ vertices that are matched in M. In other words,

$$(\# \text{ of vertices in } Y \text{ that are matched in } M) = |Y| - q.$$

Comparing this with (3), we obtain $|M| = |Y| - q$. In other words,

$$|M| + q = |Y|. \tag{4}$$

with $v \notin S$, we obtain $v \in Y \setminus S$.

Forget that we fixed v. We thus have shown that $v \in Y \setminus S$ for each $v \in N(X \cap S)$. In other words, $N(X \cap S) \subseteq Y \setminus S$.

The same argument (but applied to X and p instead of Y and q) yields

$$|M| + p = |X|. \tag{5}$$

Now, from (4), we obtain

$$\begin{aligned} |M| + q &= |Y| \\ &\geq |N(X \cap S)| + \underbrace{|S|}_{\substack{=|M|+p+q \\ \text{(by (2))}}} - |X \cap S| \qquad \text{(by (1))} \\ &= |N(X \cap S)| + |M| + p + q - |X \cap S| \\ &= \underbrace{|M| + p}_{\substack{=|X| \\ \text{(by (5))}}} + |N(X \cap S)| - |X \cap S| + q \\ &= |X| + |N(X \cap S)| - |X \cap S| + q. \end{aligned}$$

Cancelling q, we obtain

$$\begin{aligned} |M| &\geq |X| + |N(X \cap S)| - |X \cap S| \\ &= |N(X \cap S)| + |X| - |X \cap S|. \end{aligned} \tag{6}$$

Thus, we have found a matching M of G and a subset U of X (namely, $U = X \cap S$) such that $|M| \geq |N(U)| + |X| - |U|$. This proves the Hall-König matching theorem (once again).

New proof of Hall's Marriage Theorem: Proceed as in the proof of the Hall-König matching theorem that we just gave. But now assume that our bipartite graph (G, X, Y) satisfies the Hall condition (i.e., we have $|N(A)| \geq |A|$ for each subset A of X). Hence, in particular, $|N(X \cap S)| \geq |X \cap S|$. Therefore, (6) becomes

$$|M| \geq \underbrace{|N(X \cap S)|}_{\geq |X \cap S|} + |X| - |X \cap S| \geq |X|.$$

Hence, Proposition 1.3.1 **(e)** in Lecture 24 shows that the matching M is X-complete. Thus, G has an X-complete matching (namely, M). This proves Hall's Marriage Theorem (once again).

1.2. Path-missing sets

We move on to less well-trodden ground.

Menger's theorem (one of the many) is from 1927; the Gallai–Milgram theorem is from 1960. One might think that everything that can be said about paths in graphs has been said long ago.

Apparently, this is not the case. In 2017, when trying to come up with a homework exercise for a previous iteration of this course, I was experimenting with paths in Python. Specifically, I was looking at digraphs $D = (V, A, \psi)$ with two distinct vertices s and t selected. Inspired by the arc-Menger theorems, I was looking at the subsets B of A that could be removed without disconnecting s from t (more precisely, without destroying all paths from s to t). I noticed that the number of such subsets B seemed to be even whenever D has a cycle or a "useless arc" (i.e., an arc that is used by no path from s to t) [11], and odd otherwise.

I could not prove this observation. Soon after, Joel Brewster Lewis and Lukas Katthän came up with a proof and multiple stronger results. The proofs can now be found in a joint preprint [GrKaLe21], although I believe that they are far from optimal (this is one reason we have not submitted the preprint to a journal yet).

The first way to strengthen the observation is to replace the parity claim (i.e., the claim that the number is even or odd depending on cycles and useless arcs) by a stronger claim about an alternating sum. This is an instance of a general phenomenon, in which a statement of the form "the number of some class of things is even" can often be replaced by a stronger statement of the form "we can assign a plus or minus sign to each of these things, and then the total number of plus signs equals the total number of minus signs". The stronger statement is as follows:

Theorem 1.2.1 (Grinberg–Lewis–Katthän)**.** Let $D = (V, A, \psi)$ be a multidigraph. Let s and t be two distinct vertices of D. A subset B of A will be called **path-missing** if D has a path from s to t that does not use any of the arcs in B (that is, a path from s to t that would not be destroyed if we remove all arcs in B from D). (In the terminology of Lecture 27, this is the same as saying that B is **not** an s-t-arc-separator.)

Let $\mathbf{M}$ be the set of all path-missing subsets of A.

(a) If D has an arc that is not used by any path from s to t (this is what we call a "useless arc"), then

$$\sum_{B \in \mathbf{M}} (-1)^{|B|} = 0$$

(and thus $|\mathbf{M}|$ is even).

(b) If D has a cycle, then

$$\sum_{B \in \mathbf{M}} (-1)^{|B|} = 0$$

(and thus $|\mathbf{M}|$ is even).

[11] With one exception: If $A = \varnothing$, then it is odd.

(c) If $A = \varnothing$, then

$$\sum_{B \in \mathbf{M}} (-1)^{|B|} = 0$$

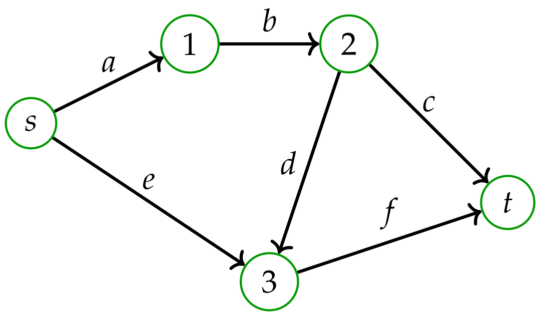
(and thus $|\mathbf{M}|$ is even).

(d) In all other cases, we have

$$\sum_{B \in \mathbf{M}} (-1)^{|B|} = (-1)^{|A| - |V'|},$$

where V' is the set of all vertices of D that have outdegree > 0 (and thus $|\mathbf{M}|$ is odd).

Example 1.2.2. Let $D = (V, A, \varphi)$ be the following digraph:



Let s and t be the vertices labelled s and t here. Then, D has neither a cycle nor a "useless arc", and its arc set A is nonempty; thus, Theorem 1.2.1 **(d)** applies. The path-missing subsets of A are the three sets $\{a, b, c, d\}$, $\{c, e\}$ and $\{d, e, f\}$ as well as all their subsets (such as $\{b, c, d\}$). In other words,

$$\begin{aligned}
\mathbf{M} &= \{\text{all subsets of } \{a, b, c, d\}\} \cup \{\text{all subsets of } \{c, e\}\} \\
&\qquad \cup \{\text{all subsets of } \{d, e, f\}\} \\
&= \{\varnothing, \{a\}, \{b\}, \{c\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \\
&\qquad \{c, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}, \{a, b, c, d\}, \\
&\qquad \{e\}, \{c, e\}, \{f\}, \{d, e\}, \{e, f\}, \{d, f\}, \{d, e, f\}\}.
\end{aligned}$$

Hence, the sum $\sum_{B \in \mathbf{M}} (-1)^{|B|}$ has 11 addends equal to -1 and 12 addends equal to 1; thus, this sum equals to 1. This is precisely the value $(-1)^{|A| - |V'|} = (-1)^{6-4} = 1$ predicted by Theorem 1.2.1 **(d)**.

Proof of Theorem 1.2.1. See [GrKaLe21, Theorem 1.3] (where $\mathbf{M}$ is denoted by $\operatorname{PM}(D)$, and where arcs are called "edges"). Of course, part **(c)** is obvious, and part **(a)** is easy (since inserting a useless arc into a set $B \in \mathbf{M}$ or removing it from a set $B \in \mathbf{M}$ always results in a set in $\mathbf{M}$). Parts **(b)** and **(d)** are the interesting ones. The proof in [GrKaLe21, Theorem 1.3] relies on a recursive argument ("deletion-contraction") in which we pick an arc with source s and consider the two smaller digraphs $D \setminus a$ and $D \diagup a$ obtained (respectively) by deleting the arc a from D and by "contracting" a "to a point". $\square$

Further levels of strength can be reached by treating $\mathbf{M}$ as a topological space. Indeed, $\mathbf{M}$ is not just a random collection of sets of arcs, but actually a **simplicial complex** (since any subset of a path-missing subset of A is again path-missing). Simplicial complexes are known to be a combinatorial model for topological spaces, and in particular they have homology groups, homotopy types, etc.. Thus, in particular, we can ask ourselves how the topological space corresponding to the simplicial complex $\mathbf{M}$ looks like. This, too, has been answered in [GrKaLe21, Theorem 1.3]: It is homotopic to a sphere or a ball (depending on the existence of cycles or "useless arcs"); its dimension can also be determined explicitly. (The sum $\sum\limits_{B \in \mathbf{M}} (-1)^{|B|}$ discussed above is, of course, its reduced Euler characteristic.)

1.3. Elser's sums

We now return to undirected (multi)graphs. Here is a result found by Veit Elser in 1984 ([Elser84, Lemma 1]), as a lemma for his work in statistical mechanics:[12]

Theorem 1.3.1 (Elser's theorem, in my version)**.** Let $G = (V, E, \varphi)$ be a multigraph with at least one edge. Fix a vertex $v \in V$.

If $F \subseteq E$, then an F**-path** shall mean a path of G such that all edges of this path belong to F. In other words, it means a path of the spanning subgraph $(V, F, \varphi\mid_F)$.

If $e \in E$ is an edge and $F \subseteq E$ is a subset, then we say that F **infects** e if there exists an F-path from v to some endpoint of e. (The terminology is inspired by the idea that some infectious disease starts at v and spreads along the F-edges.)

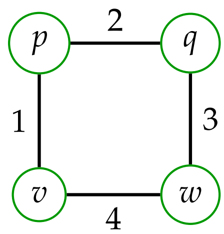
(Note that if an edge e contains the vertex v, then **any** subset F of E (even the empty set) infects e, because (v) is a trivial F-path from v to v.)

Then,

$$\sum_{\substack{F \subseteq E \text{ infects} \\ \text{every edge } e \in E}} (-1)^{|F|} = 0.$$

[12] I have restated the result beyond recognition; see [Grinbe21, Remark 1.4] for why Theorem 1.3.1 actually implies [Elser84, Lemma 1].

Example 1.3.2. Let $G = (V, E, \varphi)$ be the following graph:



,

and let v be the vertex labelled v. Then, the subsets of E that infect every edge are

$$\{1,2\}, \quad \{1,4\}, \quad \{3,4\}, \quad \{1,2,3\}, \quad \{1,3,4\}, \quad \{1,2,4\}, \quad \{2,3,4\}, \quad \{1,2,3,4\}.$$

Thus,

$$\begin{aligned} & \sum_{\substack{F \subseteq E \text{ infects} \\ \text{every edge } e \in E}} (-1)^{|F|} \\ & = (-1)^2 + (-1)^2 + (-1)^2 + (-1)^3 + (-1)^3 + (-1)^3 + (-1)^3 + (-1)^4 \\ & = 0, \end{aligned}$$

exactly as predicted by Theorem 1.3.1.

Remark 1.3.3. It might appear more natural to study subsets $F \subseteq E$ infecting vertices rather than edges. However, Theorem 1.3.1 would be false if we replaced "every edge $e \in E$" by "every vertex $v \in V$". The graph in Example 1.3.2 provides a counterexample.

However, if we go further and replace $F \subseteq E$ by $W \subseteq V$, then we get something true again – see Theorem 1.3.4 below.

Proof of Theorem 1.3.1. Elser's proof is somewhat complicated. I give a different proof in [Grinbe21, Theorem 1.2], which is elementary and nice if I may say so myself.

My proof should also be not very hard to discover, once you have the following hint: It suffices to prove the equality

$$\sum_{\substack{F \subseteq E \text{ does } \textbf{not} \text{ infect} \\ \text{every edge } e \in E}} (-1)^{|F|} = 0$$

(because the total sum $\sum_{F \subseteq E} (-1)^{|F|}$ is known to be 0). In order to prove this equality, we equip the set E with some total order (it doesn't matter how; we can just rank the edges arbitrarily), and we make the following definition: If

$F \subseteq E$ is a subset that does **not** infect every edge $e \in E$, then we let $\varepsilon(F)$ be the smallest (with respect to our chosen total order) edge that is not infected by F. Now, you can show that if $F \subseteq E$ is a subset that does **not** infect every edge $e \in E$, then the set[13] $F' := F \triangle \{\varepsilon(F)\}$ (that is, the set obtained from F by inserting $\varepsilon(F)$ if $\varepsilon(F) \notin F$ and by removing $\varepsilon(F)$ if $\varepsilon(F) \in F$) has the same property (viz., it does not infect every edge $e \in E$) and satisfies $\varepsilon(F') = \varepsilon(F)$. This entails that the addends in the sum $\sum\limits_{\substack{F \subseteq E \text{ does } \textbf{not} \text{ infect} \\ \text{every edge } e \in E}} (-1)^{|F|}$ cancel each other in pairs (namely, the addend for a given set F cancels the addend for the set $F' = F \triangle \{\varepsilon(F)\}$), and thus the whole sum is 0. $\square$

Elser's theorem, too, can be generalized and strengthened. The strengthening is similar to what we did with Theorem 1.2.1: We treat the set of all "non-pandemic-causing subsets" (i.e., of all subsets $F \subseteq E$ that **don't** infect every edge) as a simplicial complex (since a subset of a non-pandemic-causing subset is again non-pandemic-causing), and analyze this complex as a topological space. The claim of Theorem 1.3.1 then says that the reduced Euler characteristic of this space is 0; but we can actually show that this space is contractible (i.e., homotopy-equivalent to a point). Even better, we can prove that the simplicial complex of all non-pandemic-causing subsets is **collapsible** (a combinatorial property that is stronger than contractibility of the corresponding space). See [Grinbe21, §5] for definitions and proofs.

We can furthermore generalize the theorem. One way to do so is to replace our "patient zero" v by a set of vertices. This leads to a much less trivial situation. The recent paper [DHLetc19] by Dorpalen-Barry, Hettle, Livingston, Martin, Nasr, Vega and Whitlatch proves some results and asks some questions (that are still open as of 2022).

A different direction in which Elser's theorem can be generalized is more fundamental: It turns out that the theorem is not really about graphs and edges. Instead, there is a general structure that I call a "shade map", which always leads to a certain sum being 0. See [Grinbe21, §4] for the details of this generalization. I will not explain it here, but I will state one more particular case of it ([Grinbe21, Theorem 3.2]), which replaces edges by vertices throughout Theorem 1.3.1:

Theorem 1.3.4 (vertex-Elser's theorem)**.** Let $G = (V, E, \varphi)$ be a multigraph with at least two vertices. Fix a vertex $v \in V$.

[13]The symbol $\triangle$ stands for the symmetric difference of two sets. Recall its definition: If X and Y are two sets, then their **symmetric difference** $X \triangle Y$ is defined to be the set

$$\begin{aligned}(X \cup Y) \setminus (X \cap Y) &= (X \setminus Y) \cup (Y \setminus X) \\ &= \{\text{all elements that belong to } \textbf{exactly} \text{ one of } X \text{ and } Y\}.\end{aligned}$$

If $W \subseteq V$, then a W**-vertex-path** shall mean a path $\mathbf{p}$ such that all intermediate vertices of $\mathbf{p}$ belong to W. (Recall that the "intermediate vertices of $\mathbf{p}$" mean all vertices of $\mathbf{p}$ except for the starting and ending points of $\mathbf{p}$.) (Note that any path of length ≤ 1 is automatically a W-vertex-path, since it has no intermediate vertices.)

If $w \in V \setminus \{v\}$ is any vertex, and $W \subseteq V \setminus \{v\}$ is any subset, then we say that W **vertex-infects** w if there exists a W-vertex-path from v to w. (This is always true when w is a neighbor of v.)

Then,

$$\sum_{\substack{W \subseteq V \setminus \{v\} \text{ vertex-infects} \\ \text{every vertex } w \in V \setminus \{v\}}} (-1)^{|W|} = 0.$$

References

[DHLetc19] Galen Dorpalen-Barry, Cyrus Hettle, David C. Livingston, Jeremy L. Martin, George Nasr, Julianne Vega, Hays Whitlatch, *A positivity phenomenon in Elser's Gaussian-cluster percolation model*, arXiv:1905.11330v6, corrected version of a paper published in: Journal of Combinatorial Theory, Series A, **179**:105364, April 2021, doi:10.1016/j.jcta.2020.105364.

[Dieste17] Reinhard Diestel, *Graph Theory*, 5th Edition, Springer 2017. See `https://diestel-graph-theory.com/GrTh5_corrections.pdf` for errata.

[Elser84] Veit Elser, *Gaussian-cluster models of percolation and self-avoiding walks*, J. Phys. A: Math. Gen. **17** (1984), pp. 1515–1523.

[Grinbe21] Darij Grinberg, *The Elser nuclei sum revisited*, arXiv:2009.11527v8. (More detailed version of a paper published in: Discrete Mathematics & Theoretical Computer Science **23** no. 1, Combinatorics (June 3, 2021) dmtcs:7487.)

[GrKaLe21] Darij Grinberg, Lukas Katthän, Joel Brewster Lewis, *The path-missing and path-free complexes of a directed graph*, arXiv:2102.07894v1.