Warren Shull Research Statement

My primary interests lie in graph theory, the study of networks as abstract mathematical objects. My work includes a close look at the relationship between spanning trees, forbidden subgraphs, and independent sets. I also have a collaboration on partially ordered sets currently being reviewed, have begun looking at the saturation spectrum of a family of graphs called books, and have returned to a previously dormant project on Sylvester colorings of snarks. The remainder of this statement is divided into four sections, one describing each topic. Aspects of all four topics are likely to be accessible for qualified undergraduates.

Spanning Trees and Branch Vertices

A spanning tree of a graph is a connected acyclic subgraph with the same vertex set. It should be noted that a path is a kind of tree, and a spanning path is often called a Hamiltonian path. Hamiltonian paths are widely applicable, thoroughly studied, and notoriously hard to compute. Checking for such paths algorithmically is NP-complete, which is why sufficient conditions are commonly sought.

Paths are precisely the category of trees with no *branch vertices*, which are vertices whose degree is at least 3. A spanning tree with few branch vertices may thus be considered "close" to a Hamiltonian path, and is likely to share some similar applications. In my thesis work, my advisor and I proved a sufficient condition for a spanning tree with few branch vertices, as outlined below. In the process, I devised a useful concept and coined three new terms to describe it: *oblique neighbors*, *pseudoadjacency*, and *pseudoindependence*. In the coming pages, I highlight these ideas and propose a couple directions to build on my work.

When choosing a spanning tree, we naturally will have more options in a graph with more edges. Independent sets (collections of non-adjacent vertices), meanwhile, are larger in a graph with fewer edges. Since adding edges leads to "better" spanning trees (ones with fewer branch vertices), and removing edges leads to larger independent sets, we can establish parameters that require one object or the other to always exist. If we restrict our attention to claw-free graphs, we get the following result:

Theorem 1. [9] *Let k be a non-negative integer and let G be a connected claw-free graph. Then G contains either a spanning tree with at most k branch vertices, or an independent set of $2k+3$ vertices.*

The authors of Theorem 1 demonstrate it to be sharp using the example shown in Figure 1, while suggesting an additional condition for a possibly stronger result. If a graph with many edges has a large independent set, the degrees of vertices in the independent set often add up to some large number. As edges are removed, however, even the largest independent sets will see these sum totals diminish. In Figure 1, the degrees of a maximum independent set must add up to at least $|V(G)|-3$, hence Conjecture 1.

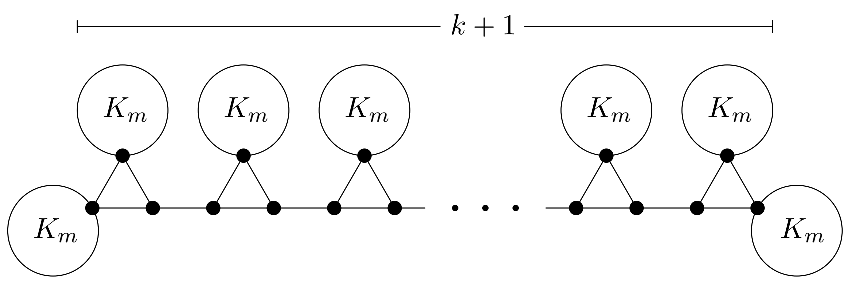


Figure 1: Any spanning tree of this graph G must contain more than k branch vertices, while a maximum independent set contains $2k+3$ vertices with degrees adding up to at least $|V(G)|-3$.

Conjecture 1. [9] *Let k be a non-negative integer and let G be a connected claw-free n-vertex graph. Then G contains either a spanning tree with at most k branch vertices, or an independent set of $2k+3$ vertices whose degrees add up to at most $n-3$.*

As described above, Theorem 1 and Conjecture 1 are both made sharp by the example in Figure 1. The truth of Conjecture 1 was essentially automatic for $k=0$ and shown in the same paper for $k=1$. I then showed it for $k=2$ in [4], and later proved it outright.

Theorem 2. [4] *Let G be a connected claw-free n-vertex graph. Then G contains either a spanning tree with at most 2 branch vertices, or an independent set of 7 vertices whose degrees add up to at most $n-3$.*

Theorem 3. [5] *Let k be a non-negative integer and let G be a connected claw-free n-vertex graph. Then G contains either a spanning tree with at most k branch vertices, or an independent set of $2k+3$ vertices whose degrees add up to at most $n-3$.*

In order to prove Theorem 3, I developed the concept of pseudoadjacency and pseudoindependence. These terms are similar to adjacency and independence, but are defined with respect to a fixed spanning tree.

Definition 1. *Let T be a spanning tree of a graph G and let $v \in V(G)$ and $e \in E(T)$. Denote $g(e,v)$ as the vertex incident to e farthest away from v in T. We say v is an* ***oblique neighbor*** *of e* ***with respect to T*** *if $vg(e,v) \in E(G)$.*

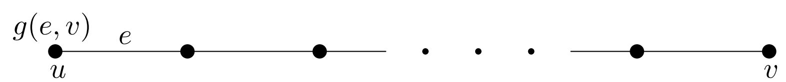


Figure 2: A path between vertices u and v within some tree T, showing $u = g(e,v)$ as described in Definition 1. If T is a spanning tree of some graph G, note that v is an oblique neighbor of e with respect to T if and only if $uv \in E(G)$.

Note that both vertices incident to a given edge of T are among its oblique neighbors.

Definition 2. *Let T be a spanning tree of a graph. Two vertices are* ***pseudoadjacent with respect to T*** *if there is some $e \in E(T)$ which has them both as oblique neighbors. A vertex set is* ***pseudoindependent with respect to T*** *if no two vertices in the set are pseudoadjacent with respect to T.*

Note that psuedoadjacency (with respect to any tree) is a weaker condition than adjacency, while pseudoindependence is a stronger condition than independence. This is a useful concept for the sums of degrees, since the number of edges with v as an oblique neighbor is exactly the degree of v in G. The next paragraph sums up the basic structure of our proof:

Take a carefully chosen spanning tree T and a carefully chosen vertex set X. Show that X is pseudoindependent with respect to T by ruling out all possible edges of T that could have any two vertices of X as oblique neighbors. The pseudoindependence of X limits the sum of its degrees to $n-1$, the number of edges in our spanning tree, and we need only find two edges with no oblique neighbors in X to lower it to $n-3$.

Figures 3 and 4 illustrate how pseudoindependence is shown in the simplest of cases, a spanning tree with just two branch vertices and four leaves. In Figure 3, assuming we have chosen the tree T with the fewest possible branch vertices, it can be shown that the edge e cannot have more than one leaf as an oblique neighbor (which would need to be shown for all edges of T to establish the leaves as a pseudoindependent set). Suppose e has two oblique neighbor leaves in the same direction, say w and x. Then we have edges vw and vx. To avoid the claw $G[v, u, w, x]$, at least one of the edges wx, uw, or ux must exist. If ux is an edge, this allows for the spanning tree in Figure 4 with only one branch vertex, contradicting our initial choice of T. The same contradiction can be reached with uw or wx, or if the leaves are in opposite directions.

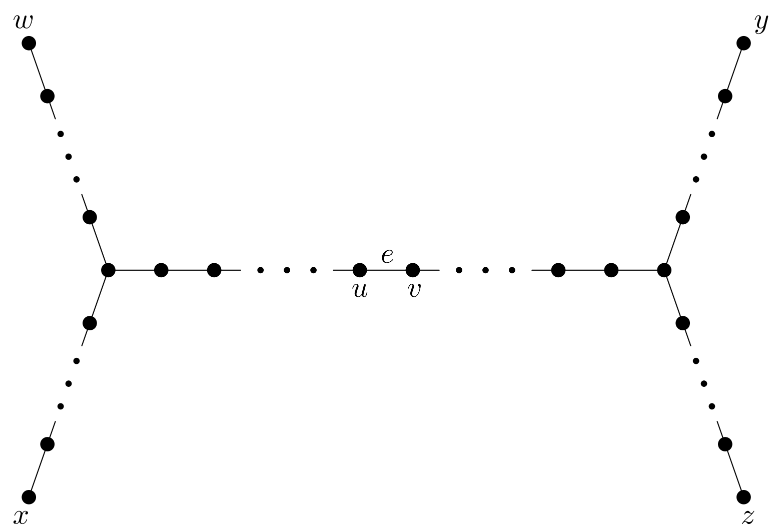


Figure 3: An important step in our proof is to choose a spanning tree with the fewest possible branch vertices. Imagine this tree T was the best we could do.

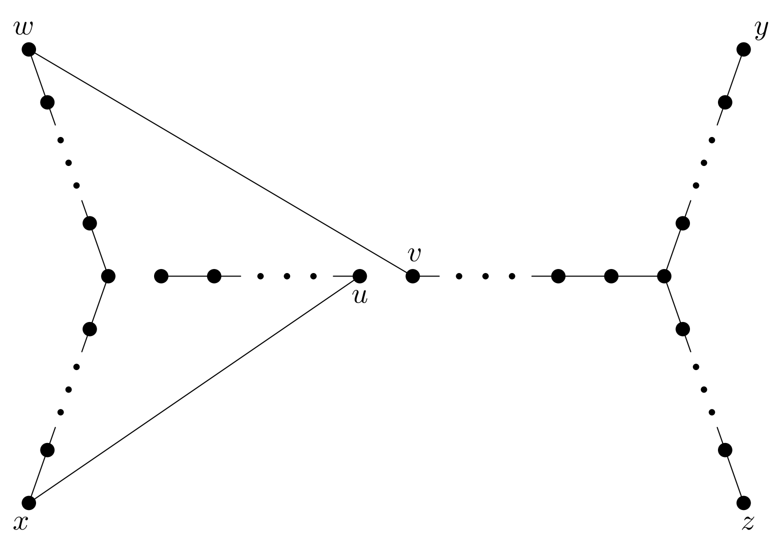


Figure 4: A spanning tree with one less branch vertex than T. This contradicts the possibility that the edges vw and vx both exist in G, since we chose the spanning tree with the fewest branch vertices.

To accomplish this proof, we combined pseudoadjacency and pseudoindependence with a technique borrowed from [9]. We chose a spanning tree and a root vertex, while considering the degrees of other branch vertices and their distances from the root. The biggest challenges were the exact choices of what spanning tree to consider and which vertices to select for our independent set. Selecting options that establish pseudoindependence took several attempts before we found one that worked, each attempt withstanding a large majority of cases in a complex, multi-layered case argument.

Future Work on Spanning Trees

This section discusses two possible directions for future work. The second of these begins with the paragraph "As a second direction for future work..."

Looking at the sharpness example in Figure 1, one might notice that many of its vertices have degree 3 (in the whole graph). A natural question is: how much stronger is our result if we require the graph to have a minimum degree of 4 or larger? If our graph must have minimum degree at least t, then Figure 5 shows that we cannot guarantee an independent set any larger than before, though we might be able to make their degrees add up to a smaller number.

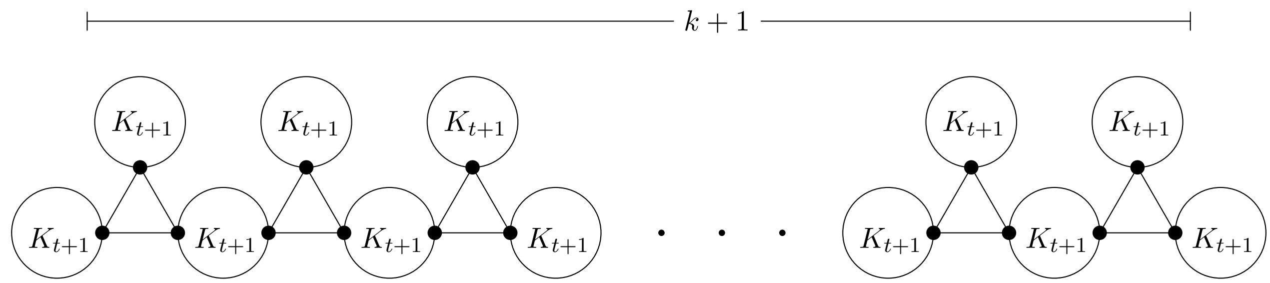


Figure 5: This graph has minimum degree t and contains no spanning trees with at most k branch vertices. A maximum independent set contains $2k+3$ vertices as before, and their degrees must add up to at least $|V(G)| - 2k - 3$.

Given the above example, the following corollary to Theorem 1 and new conjecture are sharp, no matter how high a minimum degree we require:

Corollary 1. *Let G be a connected claw-free graph with minimum degree at least 4. Then G contains either a spanning tree with at most k branch vertices or an independent set of $2k+3$ vertices.*

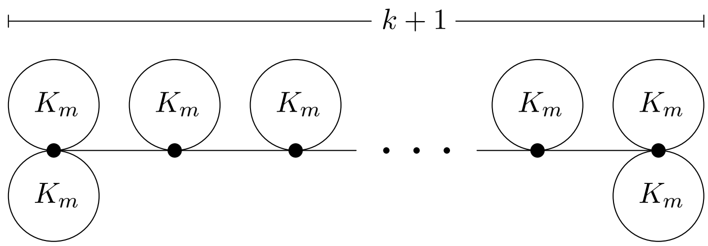
Conjecture 2. *Let G be a connected claw-free n-vertex graph with minimum degree at least 4. Then G contains either a spanning tree with at most k branch vertices or an independent set of $2k+3$ vertices whose degrees add up to at most $n-2k-3$.*

My future research will explore ways to modify the argument to reduce this sum of degrees from its current level at $n-3$, or to prove Conjecture 2 for small values of k. An instrumental tool for the small cases of Conjecture 1 was the following theorem about spanning trees with a small number of *leaves*:

Theorem 4. [7] *Let k be a non-negative integer and let G be a connected claw-free n-vertex graph. Then G contains either a spanning tree with at most $k+2$ leaves, or an independent set of $k+3$ vertices whose degrees add up to at most $n-k-3$.*

This result, either in its current form or improved for graphs of larger minimum degree, is likely to shed light on Conjecture 2 at least for small values of k. I believe these cases warrant some exploration before a time-intensive commitment to the whole of the conjecture. Furthermore, small cases of a conjecture like this one seem a fine starting point for new collaborators, especially students.

As a second direction for future work, I've begun to consider what would happen if we didn't require our graph G to be claw-free. How large an independent set (in terms of k and n) must exist assuming the desired spanning tree does not, and how small a total degree must one of them have? If we take the sharpness example at the beginning and simply collapse each triangle to a single point, we get the following graph:



The graph above has a total of $(m-1)(k+3)+(k+1) = mk-3m-2$ vertices. Each spanning tree of this graph has at least $k+1$ vertices, and the largest possible independent set has $k+3$ vertices. Any such independent set will have total degree $(m-1)(k+3) = n-k-1$, so the new conjecture is:

Conjecture 3. *Let k be a positive integer and let G be a connected n-vertex graph. Then G contains* either *a spanning tree with at most k branch vertices,* or *an independent set of $k+3$ vertices with total degree at most $n-k-1$.*

The observant reader may notice that, unlike the conjectures about claw-free graphs, Conjecture 3 requires that k be positive. This is because the bipartite graph, with $\lceil \frac{n+2}{2} \rceil$ vertices in one partite set and $\lfloor \frac{n-2}{2} \rfloor$ in the other, violates the conjecture for $k=0$. Our original goal, however, involved *first* finding the largest independent set we could, and *then* choosing from equally large independent sets for the smallest total degree. Since the aforementioned bipartite graph has independent sets larger than 3, perhaps we need not worry so much about the lack of an independent 3-set with a small enough total degree. To reflect this goal, I have rephrased the conjecture:

Conjecture 4. *Let k be a non-negative integer and let G be a connected graph. Then G contains* either *a spanning tree with at most k branch vertices,* or *an independent set of $k+3$ vertices with total degree at most $|V(G)|-k-1$,* or *an independent set larger than $k+3$ vertices.*

And if, instead of minimizing branch vertices, we wish to minimize leaves (another way of being *close* to a path), the related conjecture is:

Conjecture 5. *Let k be a non-negative integer and let G be a connected graph. Then G contains* either *a spanning tree with at most $k+2$ leaves,* or *an independent set of $k+3$ vertices with total degree at most $|V(G)|-1$,* or *an independent set larger than $k+3$ vertices.*

To the best of my knowledge, both these conjectures are open even for $k=0$, which I also anticipate to be a good starting point for students.

Partially Ordered Sets and Local Dimension

The *dimension* of a partially ordered set, or poset, is a well-studied property, and *local dimension* is a much newer variant on it. A few definitions are needed to understand these ideas.

Definition 3. *A* **linear extension** $\leq_i$ *of a poset P is a total ordering on its elements such that for any $x \leq y$ in P, we always have $x \leq_i y$. In other words, all orderings within P are respected by $\leq_i$.*

Definition 4. *A* **realizer** *of a poset $P = (X, \leq)$ is a set $\{\leq_1, \ldots, \leq_d\}$ of linear extensions such that $x \leq y$ if and only if $x \leq_i y$ for every i. The* **dimension** *of P is the minimum size of a realizer.*

Local dimension is a parameter bounded above by dimension, and has so far appeared much harder to compute.

Definition 5. *A* **partial linear extension** $\leq_i$ *of a poset P is a linear extension of a sub-poset of P.*

Definition 6. *A* **local realizer** *of a poset $P = (X, \leq)$ is a set $\{\leq_1, \ldots, \leq_l\}$ of partial linear extensions such that*

$$x < y \Rightarrow x <_i y \text{ for some } i \in \{1, \ldots, d\},$$

and

$$x \text{ is incomparable to } y \Rightarrow x <_i y \text{ for some } i \in \{1, \ldots, d\}$$

for any $x, y \in X$. The **local dimension** *of P is the minimum number of times the most frequent element appears in a local realizer.*

There is a standard example to distinguish dimension from local dimension. Consider the poset $\{a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n\}$, where the only relations are $a_i < b_j$ whenever $i \neq j$. To create a realizer, we must have $a_i > b_i$ in at least one linear extension for every i, and no linear extension can contain more than one of these relations. This forces the dimension of the standard example to be at least n, and indeed it is exactly n. We can, however, take the partial linear extensions $a_1 < a_2 < \cdots < a_n < b_1 < b_2 < \cdots < b_n$, and $a_n < a_{n-1} < \cdots < a_2 < a_1 < b_n < b_{n-1} < \cdots < b_2 < b_1$, and $a_i > b_i$ for each i, and each element only appears three times, proving that the local dimension of the standard example is at most 3, no matter how large n is. For $n \geq 3$, the standard example indeed has local dimension 3; it is known that any poset with dimension at least 3 has local dimension at least 3.

In a resubmitted paper, we have a variety of new results, two of which I will highlight here. Relevant to them both is a classical result of Chung, Erdos, and Spencer:

Theorem 5. [1] *For asymptotically large n, there exists an n-vertex graph such that any edge covering with complete bipartite graphs $\{B_1, B_2, \ldots, B_k\}$ must have $\displaystyle\sum_{i=1}^{k} |V(B_i)| \geq O\left(\frac{n^2}{\ln n}\right)$.*

One of our results improves on this once we define a category of bipartite graphs called *semigraphs.*

Definition 7. *A **semigraph** $H(a, b, f)$ is a bipartite graph on $a+b$ vertices $U \cup W$ where $U = \{u_1, \cdots, u_a\}$ and $W = \{w_1, \cdots, w_b\}$ together with a non-increasing function $f : [a] \to [b]$ such that $N(v_i) = \{w_1, \cdots, w_{f(i)}\}$ for all $i \in [a]$ and $f(1) = b$.*

It should be noted that semigraphs are a generalization of complete bipartite graphs. We have thus improved on Theorem 5 as follows:

Theorem 6. [8] *For asymptotically large n, there exists an n-vertex bipartite graph such that any edge covering with semigraphs $\{S_1, S_2, \ldots, S_k\}$ must have $\displaystyle\sum_{i=1}^{k} |V(S_i)| \geq O\left(\frac{n^2}{\ln n}\right)$.*

Corollary 2. *For asymptotically large n, there exists a bipartite graph G such that for any edge covering with semigraphs $\{S_1, S_2, \ldots, S_k\}$, some vertex appears in at least $O\left(\frac{n}{\ln n}\right)$ of the semigraphs.*

This corollary implies the sharpness of our other main result:

Theorem 7. [8] *For any poset P with n points, with n asymptotically large, the local dimension of P is $O\left(\frac{n}{\ln n}\right)$.*

We reach the sharpness conclusion by sorting partial linear extensions (PLE's) of a height 2 poset into equivalence classes; two PLE's belong in the same class if they contain the same elements and the necessary swaps to move between them are between elements on the same level (of the two levels of the height 2 poset). These equivalence classes are in bijection with semigraphs, implying the existence of a height 2 poset (for asymptotically large n) which can only be realized with at least one element appearing at least $O\left(\frac{n}{\ln n}\right)$ times. Since bipartite subgraphs are a special case of semigraphs, the following result gives the necessary upper bound:

Theorem 8. [2] *Let $G = (V, E)$ be a graph on n vertices. The edge set E can be partitioned into complete bipartite subgraphs such that each vertex $v \in V$ is contained by $O\left(\frac{n}{\ln n}\right)$ of the bipartite subgraphs.*

The reader may have noticed that our lower bound was restricted to height 2 posets; this extends to all posets using an operation called the "split" which consists of placing copies of the set on two distinct levels, turning every previously existing relation into one between the levels, with one level always above and the other always below.

Several smaller results are analogous to known results for dimension. Future research on local dimension could explore open questions like those below. The *removable pair conjecture* remains open for both dimension and local dimension:

Conjecture 6. [10] *For every poset P there is some $x, y \in P$ such that* $\dim(P) \leq \dim(P - \{x, y\}) + 1$.

Conjecture 7. [8] *For every poset P there is some $x, y \in P$ such that* $\mathrm{ldim}(P) \leq \mathrm{ldim}(P - \{x, y\}) + 1$.

We have also conjectured the following for local dimension, which is already known for dimension:

Conjecture 8. [8] *For all $n \geq 1$, the Boolean lattice on an n-set has local dimension n.*

It remains unknown for $n \geq 4$, and appears harder to compute than first expected. New, creative ideas are needed for these and many other open questions on dimension, local dimension, and other emerging variations.

Saturation

Given a positive integer n and a graph H, the *extremal number* $ex(n, H)$ is the maximum number of edges of an n-vertex graph that does not contain any copy of H as a subgraph, and the n-vertex graph that achieves this maximum is called the *extremal graph*. This idea originated in 1906 with Mantel, who showed that $ex(n, K_3) = \left\lfloor \frac{n^2}{4} \right\rfloor$; that is, an n-vertex graph without no K_3 (triangle) as a subgraph contains at most $\left\lfloor \frac{n^2}{4} \right\rfloor$ edges, and this bound is achieved with a complete bipartite graph whose partite sets differ in size by at most 1.

A notable feature of extremal graphs is that they are *H-saturated*. That is, the addition of a single edge will always create a copy of H. Saturation, however, is not limited to extremal graphs. A complete bipartite graph will be K_3-saturated whenever both partite sets are nonempty, even though its total number of edges may be as small as $n - 1$. The *saturation number*, denoted $sat(n, H)$, is the minimum number of edges in an H-saturated n-vertex graph, and indeed $sat(n, K_3) = n - 1$.

For any forbidden subgraph H, once we know the extremal number and the saturation number, we can ask whether saturated graphs G exist with $|E(G)| = m$ for every number m between $sat(n, H)$ and $ex(n, H)$. This range of numbers is called the *saturation spectrum*. The doctoral thesis of Jessica Fuller, a graduate colleague of mine, focused on the saturation spectrum of $K_4 - e$; that is, the complete graph on 4 vertices, with just one edge removed [3]. This may be thought of as two triangles sharing an edge.

Any number of triangles t all sharing the same edge is called a *book* and denoted B_t. Since the completion of Dr. Fuller's thesis, I have been working on the saturation spectrum of B_3 with my advisor and another graduate colleague. So far, we have constructions covering parts toward the middle of the spectrum, and a proof ruling out a chunk at the lower end. Many numbers remain unknown, especially at the upper end. The unsolved parts of the spectrum provide plenty of accessible entry points for new collaborators.

Sylvester coloring

A long-studied class of problems in graph theory is that of edge coloring. A coloring of the edges of a graph is considered *proper* if no two edges of the same color share a vertex. The minimum number of colors needed to properly color the edges of the graph is called the *chromatic index* of the graph. It is not hard to see that the chromatic index of any graph G (denoted $\chi'(G)$) is at least the maximum degree of G (denoted $\Delta(G)$); a famous result central to edge coloring is Vizing's Theorem, which says that $\Delta(G) \leq \chi'(G) \leq \Delta(G) + 1$ for any graph G. Consequently, a graph G for which $\chi'(G) = \Delta(G)$ is called a "class one" graph, and one for which $\chi'(G) = \Delta(G) + 1$ is called a "class two" graph.

Edge colorings have been studied especially often for graphs that are *3-regular*–that is, those in which every vertex has degree 3. Those of class two are called *snarks*, and have proven especially hard to say anything general about. The simplest snark is the famous Petersen graph.

Related to standard edge colorings, it is possible under certain conditions to color one graph G *with the edges of another graph* H. This is called an H-coloring of G, and is a mapping from the edge set of G to that of H (so if we think of every edge of H as having a different color, we choose colors for edges of G from among the colors of H). An H-coloring of G must satisfy one crucial property: For every vertex v of G, there must be some vertex u of H such that the set of colors of edges incident to v in G exactly matches those of the edges incident to u in H (recall that every edge of H has a *unique* color). Note that this implicitly requires each v to have the same degree as the corresponding u, though this property is automatic if we restrict our attention to 3-regular graphs. It should be noted that a class one 3-regular graph, being 3-colorable, is automatically H-colorable for any 3-regular graph H (or even any H with a single degree 3 vertex).

The Petersen coloring conjecture of Jaeger states that every snark can be Petersen colored (that is, H-colored if H is the Petersen graph). It has remained unsolved since 1988, but a weaker conjecture was proposed in 2012 by Mkrtchyan. It replaces the Petersen graph with a multigraph called the Sylvester graph, shown below. I became familiar with both the Petersen and Sylvester coloring conjectures at the Graduate Research Workshop in Combinatorics, which is also where my local dimension project began. After some initial attempts, the project went dormant, but I have begun thinking about it again in the last month or two and revived communication with one collaborator from the workshop. There are angles from which I hadn't thought to approach the problem, which I am just now beginning to explore.

References

[1] Chung, F.R.K., Erdős, P., and Spencer, J., On the decomposition of graphs into complete bipartite subgraphs. *Studies in pure mathematics*, 95–101, Birkhäuser, 1983.

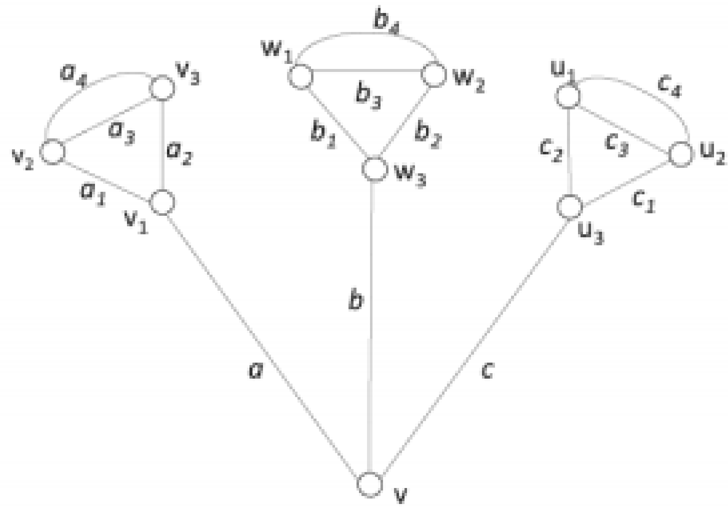


Figure 6: The Sylvester Graph. Image credit [6]

[2] Erdős, P. and Pyber, L., Covering a graph by complete bipartite graphs, *Discrete Math.*, **170**(1-3) (1997), 249-251.

[3] Fuller, J., On Saturation Spectrum, Doctoral Thesis, Emory University, 2017.

[4] Gould, R., Shull, W.: On a Conjecture on Spanning Trees with few Branch Vertices. Journal of Combinatorial Mathematics and Combinatorial Computing (2019).

[5] Gould, R., Shull, W.: On Spanning Trees with few Branch Vertices. Discrete Mathematics (to appear).

[6] Hakobyan, A., Mkrtchyan, V.: On Sylvester Colorings of Cubic Graphs.

[7] Kano, M., Kyaw, A., Matsuda, H., Ozeki, K., Saito, A., Yamashita, T.: Spanning trees with small number of leaves in a claw-free graph. Ars. Combin. **103**, 137-154 (2012)

[8] Kim, Jinha, et. al., Local Dimension. Resubmitted to EJC.

[9] Matsuda, H., Ozeki, K., Yamashita, T.: Spanning Trees with a Bounded Number of Branch Vertices in a Claw-Free Graph. Graphs and Combinatorics **30**, 429-437 (2014)

[10] Trotter, William T. Jr.: Inequalities in Dimension Theory for Posets. Proceedings of the American Mathematical Society **47** (1975), 311-316.