The Alon-Saks-Seymour and Rank-Coloring Conjectures

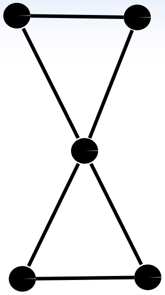
Michael Tait

Department of Mathematical Sciences
University of Delaware
Newark, DE 19716
tait@math.udel.edu

April 20, 2011

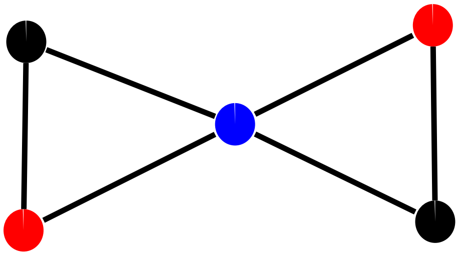
Preliminaries

- A **graph** is a set of vertices $V(G)$ and a set of edges $E(G)$, where each edge is an unordered pair of vertices.
- The adjacency matrix of a graph is a $|V(G)| \times |V(G)|$ matrix with rows and columns indexed after the vertices. The xy'th entry is 1 is xy is an edge in G and 0 otherwise. This matrix is denoted by $A(G)$
- We denote the rank of $A(G)$ by $\mathrm{rank}(A(G))$.



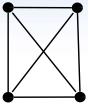
Preliminaries

A proper *k*-coloring of a graph *G* assigns *k* colors to the vertices of *G* in such a way that if two vertices are adjacent they do not have the same color. The **chromatic number** of a graph is the minimum number *k* such that a proper *k* coloring of *G* exists and is denoted $\chi(G)$.

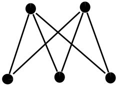


Preliminaries

- The complete graph on n vertices is the graph on n vertices with all $\binom{n}{2}$ possible edges and is denoted K_n.

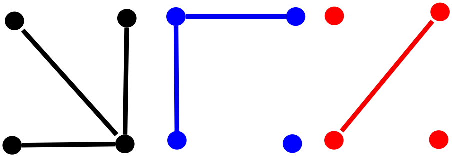


- An independent set is a set of vertices that are pairwise nonadjacent.
- A complete bipartite graph (also called biclique) is an independent set of size a and an independent set of size b with all $a \cdot b$ edges between them and is denoted $K_{a,b}$.



Preliminaries

- The **biclique partition number** of a graph G is the minimum number of bicliques necessary to partition the edge set of G, and is denoted $\mathrm{bp}(G)$.



- So for example, $\mathrm{bp}(K_4) \leq 3$.

The Graham-Pollak Theorem

- In fact, $\mathrm{bp}(K_n) \leq n-1$ for any n.
- We can prove by induction. To see this, we can take a $K_{1,n-1}$ out of the edge set of K_n, and what we are left with is the edge set of K_{n-1}.
- This problem begins with the Graham-Pollak Theorem. In 1971, Graham and Pollak proved that the inequality also goes the other direction, i.e. that $\mathrm{bp}(K_n) \geq n-1$.

Theorem (Graham-Pollak Theorem)

The edge set of the complete graph on n vertices cannot be partitioned into fewer than $n-1$ complete bipartite subgraphs.

- Several proofs of this fact have since been discovered (e.g. Witsenhausen, Peck, Tverberg, Vishwanathan).

The Alon-Saks-Seymour Conjecture

- Since $\chi(K_n) = n$, the Graham-Pollak Theorem can be rephrased as $\chi(K_n) = \mathrm{bp}(K_n) + 1$.
- This prompted Alon, Saks, and Seymour to make the following conjecture in 1991.

Alon-Saks-Seymour Conjecture - 1991

If the edge set of a graph G can be partitioned into the edge disjoint union of k bicliques, then $k + 1 \geq \chi(G)$.

- Rephrasing, the conjecture says for any graph G, the inequality $\chi(G) \leq \mathrm{bp}(G) + 1$ holds.

The Rank-Coloring Conjecture

- We also notice that $\mathrm{rank}(A(K_n)) = n$.
- In 1976, van Nuffelen stated what became known as the Rank-Coloring Conjecture.

Rank-Coloring Conjecture

For any simple graph G, $\chi(G) \leq \mathrm{rank}(A(G))$.

Counterexamples

- Neither conjecture is true!
- In 1989, Alon and Seymour constructed the first counterexample to the Rank-Coloring Conjecture with a graph that has rank 29 and chromatic number 32.
- In 1992, Razborov found the first counterexample with a superlinear gap between rank and chromatic number by constructing an infinite family of graphs G_n such that $\chi(G_n) \geq c(\mathrm{rank}(A(G_n)))^{4/3}$ for some fixed $c > 0$.
- At the current time, a construction of Nisan and Wigderson yields the largest gap between rank and chromatic number.
- The Alon-Saks-Seymour Conjecture remained open for 20 years until Huang and Sudakov constructed graphs H_n such that $\chi(H_n) \geq c(\mathrm{bp}(H_n))^{6/5}$ for some fixed $c > 0$.

Thesis Outline

- We construct new infinite families of counterexamples to both conjectures.
- These families generalize the constructions of Razborov and of Huang and Sudakov.
- We explain the relationship between these conjectures and questions in theoretical computer science.
- We consider a generalization of the Graham-Pollak Theorem to hypergraphs.

University of Delaware

Construction

- We construct graphs $G(n,k,r)$ with $n^{2k+2r+1}$ vertices for all integers $n \geq 2$, $k \geq 1$, $r \geq 1$.
- $$\chi(G(n,k,r)) \geq \frac{n^{2k+2r}}{2r+1}. \qquad (1)$$
- For $k \geq 2$,

$$2k(2r+1)(n-1)^{2k+2r-1} \leq \mathrm{bp}(G(n,k,r)) < 2^{2k+2r-1} n^{2k+2r-1} \qquad (2)$$

and

- $$2k(2r+1)(n-1)^{2k+2r-1} \leq \mathrm{rank}(A(G(n,k,r))) < 2k(2r+1)n^{2k+2r-1}. \qquad (3)$$
- So for fixed k, r, and n large enough, $G(n,k,r)$ is a counterexample to both conjectures.

Construction

- Let Q_n be the n-dimensional cube with vertex set $\{0,1\}^n$. Let the all ones and all zeros vectors be denoted by 1^n and 0^n.
- Let Q_n^- be defined as $Q_n \setminus \{1^n, 0^n\}$.
- Given integers n, k, r, we define $G(n,k,r)$ as follows.
- $V(G(n,k,r)) = [n]^{2k+2r+1} = \{(x_1, ..., x_{2k+2r+1}) | x_i \in [n], 1 \leq i \leq 2k+2r+1\}$.
- For any two vertices $x = (x_1, ..., x_{2k+2r+1})$ and $y = (y_1, ..., y_{2k+2r+1})$, let

$$\rho(x,y) = (\rho_1(x,y), ..., \rho_{2k+2r+1}(x,y))$$

  where $\rho_i(x,y) = 1$ if $x_i \neq y_i$ and $\rho_i(x,y) = 0$ if $x_i = y_i$.
- We define adjacency as $x \sim y$ if and only is $\rho(x,y) \in S$ where

$$S = Q_{2k+2r+1} \setminus [(1^{2k} \times Q_{2r+1}^-) \cup \{0^{2k} \times 0^{2r+1}\} \cup \{0^{2k} \times 1^{2r+1}\}].$$

Chromatic Number

Proposition:

For $n \geq 2$ and $k, r \geq 1$, $\chi(G(n,k,r)) \geq \frac{n^{2k+2r}}{2r+1}$.

Proof (Very brief sketch): Using the definition of the set S, we show that an independent set in G can have size at most $(2r+1)n$. Using the fact that (for any graph) $\chi(G) \geq \frac{|V(G)|}{\alpha(G)}$, the bound follows.

Biclique Partition Number

Proposition:
For $n, k \geq 2$ and $r \geq 1$, $\mathrm{bp}(G(n,k,r)) < 2^{2k+2r-1} n^{2k+2r-1}$.

Proof (Very brief sketch):

- First we prove that S can be partitioned into 2-dimensional subcubes.
- This allows us to write G as the edge disjoint union of subgraphs $G_1, \ldots, G_t$, where $t < 2^{2k+2r-1}$ and each G_i is an n^2 blowup of some graph G_i' which has $n^{2k+2r-1}$ vertices.
- Since any blowup of a biclique is still a biclique, we see that $\mathrm{bp}(G_i) \leq \mathrm{bp}(G_i')$.
- Then because the edge set of G is partitioned by the edges of the G_i's, we have

$$\mathrm{bp}(G) \leq \sum_{i=1}^{t} \mathrm{bp}(G_i) \leq \sum_{i=1}^{t} \mathrm{bp}(G_i') \leq \sum_{i=1}^{t} |V(G_i')| - 1 < 2^{2k+2r-1} n^{2k+2r-1}$$

Rank

Proposition:

For $n \geq 2$ and $k, r \geq 1$,
$2k(2r+1)(n-1)^{2k+2r-1} \leq \text{rank}(A(G(n,k,r))) < 2k(2r+1)n^{2k+2r-1}$.

Proof (Very brief sketch):

- We notice that G can be defined by something called the Non-complete Extended P-Sum (NEPS). Because of this, we can determine the spectrum of G by

$$f(x_1, \ldots, x_{2k+2r+1}) = \sum_{(s_1,\ldots,s_{2k+2r+1}) \in S} \prod_{i=1}^{2k+2r+1} x_i^{s_i}$$

  where f is evaluated at all possible combinations where the x_i's are eigenvalues of the complete graph K_n.

- This looks complicated but actually simplifies nicely! By plugging in values carefully, we obtain lower bounds on the number of both zero and non zero eigenvalues of G and show

$$2k(2r+1)(n-1)^{2k+2r-1} \leq \text{rank}(A(G(n,k,r))) < 2k(2r+1)n^{2k+2r-1}$$

Taking a Step Back

- That was technical, but most importantly, remember that we've constructed graphs $G(n, k, r)$ on $n^{2k+2r+1}$ vertices.
- $$\chi(G(n, k, r) \geq \frac{n^{2k+2r}}{2r+1}.$$
- For $k \geq 2$,
  $$2k(2r+1)(n-1)^{2k+2r-1} \leq \mathrm{bp}(G(n, k, r)) < 2^{2k+2r-1} n^{2k+2r-1}$$
  and
- $$2k(2r+1)(n-1)^{2k+2r-1} \leq \mathrm{rank}(A(G(n, k, r))) < 2k(2r+1)n^{2k+2r-1}.$$
- So for fixed k, r, and n large enough, $G(n, k, r)$ is a counterexample to both conjectures.

Applications

- Next we talk about the applications of the Alon-Saks-Seymour and Rank-Coloring Conjectures to theoretical computer science.
- We talk about a deterministic model of communication complexity that was first introduced by Yao in 1979.
- The basic model is that there are two parties (traditionally named Alice and Bob), and two finite sets X and Y. The task is to evaluate a boolean function

$$f : X \times Y \to \{0, 1\}$$

- The function is publicly known, the difficulty is that Alice is the only one who can see the input $x \in X$ and Bob is the only one that can see the input $y \in Y$.

Applications

$a_1 = f_1(x)$

$b_1 = g_1(y, a_1)$

$a_2 = f_2(x, a_1, b_1)$

$\cdots$

$b_t = g_t(y, a_1, b_1, \ldots, a_t)$

$x \in X$

$b_t = f(x, y)$!

$y \in Y$

Applications

- Given a protocol p, we define the cost of evaluating the function $\alpha_p(x, y)$ to be the number of bits that Alice and Bob need to exchange before $f(x, y)$ can be computed.
- Then the *deterministic communication complexity* of f is defined the be the cost of the "best" protocol given the "worst" inputs x and y and we will denote it by $C(f)$. More precisely

  $$C(f) = \min_{p \in P} \max_{x \in X, y \in Y} \alpha_p(x, y)$$

  where P is the set of all protocols.
- For any boolean function f we can define a matrix M_f where the rows are indexed after X and the columns after Y where $(M_f)_{x,y} = f(x, y)$.

Theorem (Yao/Mehlhorn and Schmidt)

$C(f) \geq \log_2 \text{rank}(M_f)$.

Log-Rank Conjecture

- Lovaśz and Saks have conjectured that this bound is “almost” tight.

Conjecture (Still open!)

(Log-Rank Conjecture) There exists a constant $k > 0$ such that for any function f

$$C(f) \leq (\log_2 \mathrm{rank}(M_f))^k.$$

- Next we explain the connection between the Log-Rank Conjecture and the Rank-Coloring Conjecture.

Log-Rank/Rank-Coloring

Proposition

The Log-Rank Conjecture is true if and only if there exists a constant $l > 0$ such that for any graph G

$$\log_2 \chi(G) \leq (\log_2 \mathrm{rank}(A(G)))^l$$

- Further, for any graph G such that $\mathrm{rank}(A(G)) < \chi(G)$ there is a corresponding boolean function $f : V(G) \times V(G) \to \{0, 1\}$ such that $\log_2(\mathrm{rank}(M_f) - 1) < C(f)$.
- We constructed graphs $G(n, k, r)$ such that $\chi(G(n, k, r)) \geq \frac{n^{2k+2r}}{2r+1}$ and $\mathrm{rank}(A(G(n, k, r))) < 2k(2r + 1)n^{2k+2r-1}$.
- These graphs correspond to functions f defined by $M_f = J - A(G(n, k, r))$ such that

$$C(f) \geq \frac{2k + 2r}{2k + 2r - 1} \log_2(\mathrm{rank}(M_f)) - c$$

for a fixed constant c.

Clique vs. Independent Set Problem

- We apply the question of deterministic communication complexity to the Clique vs. Independent Set Problem (CL-IS).
- In this problem there is a publicly known graph G. Alice gets a complete subgraph C of G and Bob gets an independent set I of G.
- Letting X be the set of all cliques and Y the set of all independent sets, the objective function is given by $f : X \times Y \to \{0,1\}$ where $f(C, I) = |C \cap I|$.
- We denote the deterministic communication complexity of the function by $C(CL - IS_G)$.
- To find a lower bound, notice that we can consider each vertex as both a clique and an independent set of size 1. Then there are $|V(G)|$ vertices that may be given to Alice and Bob. This means that $I_{|V(G)|}$ is a submatrix of M_f, which means that $\mathrm{rank}(M_f) \geq \mathrm{rank}(I_{|V(G)|}) = |V(G)|$.This implies that $C(CL - IS_G) \geq \log_2 |V(G)|$.
- Surprisingly, this is the best lower bound known.

Clique vs. Independent Set Problem

- We discuss the connection between the CL-IS problem and the Alon-Saks-Seymour Conjecture.

Proposition

(Alon and Haviv) For and graph G with $\chi(G) > \mathrm{bp}(G) + 1$ there is a corresponding graph H with $C(CL - IS_H) > \log_2 |V(H)|$.

- We constructed graphs $G(n, k, r)$ with $\chi(G(n, k, r)) \geq \frac{n^{2k+2r}}{2r+1}$ and $\mathrm{bp}(G(n, k, r)) < 2^{2k+2r-1} n^{2k+2r-1}$.
- These correspond to graphs $H = H(n, k, r)$ such that

$$C(CL - IS_H) \geq \frac{2k + 2r}{2k + 2r - 1} \log_2 |V(H)| - c$$

for a fixed constant c.

Hypergraphs

- Next we talk about a generalization of the Graham-Pollak Theorem.
- The complete r-uniform hypergraph on n vertices has vertex set $[n]$ and edge set $\binom{[n]}{r}$ and is denoted $K_n^{(r)}$.
- If $X_1, ..., X_r$ are disjoint subsets of $[n]$, then the complete r-partite r-uniform subgraph with partite sets $X_1, ..., X_r$ has edge set $\{(x_1, ..., x_r) | x_i \in X_i\}$.
- In 1986, Alon asked the question, how many complete r-partite r-uniform subgraphs are necessary to partition the edge set of $K_n^{(r)}$ and we denote this value by $f_r(n)$.
- Indeed this is a generalization of the Graham-Pollak Theorem, because for $r = 2$ the question asks how many bicliques are necessary to partition the edge set of K_n.

Hypergraphs

- The value of $f_r(n)$ is not known for $r \geq 4$.

The best published bounds are given by Cioabă, Küngden, and Verstraëte, who improved a result of Alon and proved the following theorem.

Theorem

If $f_r(n)$ denotes the minimum number of complete r-partite r-uniform subgraphs necessary to partition the edge set of the complete r-uniform graph on n vertices, then

$$\frac{2\binom{n-1}{k}}{\binom{2k}{k}} \leq f_{2k}(n) \leq \binom{n-k}{k} \tag{4}$$

and

$$f_{2k}(n-1) \leq f_{2k+1}(n) \leq \binom{n-k-1}{k}. \tag{5}$$

Hypergraphs

- We find the value of $f_r(n)$ exactly in the case when $n = r + 2$.

Theorem
$f_{2k}(2k+2) = f_{2k+1}(2k+3) = \lceil \frac{2k^2+5k+3}{4} \rceil.$

- We make a slight improvement on the upper bound of $f_{2k}(n)$ by showing

$$f_{2k}(n) < \binom{n-k}{k} - \frac{n}{20}\binom{\lfloor \frac{n}{2} \rfloor - k + 4}{k-4}.$$

Open Questions

In the final chapter of the thesis, we list open problems:

- Is the Log-Rank Conjecture true? Equivalently, does there exist a constant $l > 0$ such that for all graphs G

$$\log_2 \chi(G) \leq (\log_2 \mathrm{rank}(A(G)))^l.$$

- Do there exist graphs G_n with arbitrarily large biclique partition number k_n and chromatic number at least $2^{c \log^2 k_n}$ for some fixed constant $c > 0$?
- What is the correct value for $f_{2k}(n)$ and $f_{2k+1}(n)$?