My favorite application using eigenvalues: Eigenvalues and the Graham-Pollak Theorem

Michael Tait

Winter 2013

Abstract

The famous Graham-Pollak Theorem states that one needs at least $n-1$ complete bipartite subgraphs to partition the edge set of the complete graph on n vertices. Originally proved in conjunction with addressing for networking problems, this theorem is also related to perfect hashing and various questions about communication complexity. Since it's original proof using Sylvester's Law of Intertia, many other proofs have been discovered. Though the statement is purely combinatorial in nature, it is a surprising fact that most proofs have been algebraic. In this essay, we give a beautiful result about how the eigenvalues of an adjacency matrix of a graph relate to the minimum number of bipartite subgraphs necessary to partition its edge set.

1 Introduction

Originally motivated by a problem of loop switching in networks, Graham and Pollak [6] became interested in partitioning the edge set of a multigraph by complete bipartite subgraphs (henceforth *bicliques*). If G is a finite, loopless multigraph, the *biclique partition number*, denoted $\mathrm{bp}(G)$, is the minimum number of bicliques whose edge sets partition $E(G)$. Since every edge is a biclique, this parameter is well-defined and finite. Graham and Pollak showed that a problem on loop switching is equivalent to partitioning a multigraph, and in the process proved their celebrated theorem [8].

Theorem 1 (Graham-Pollak Theorem)**.** *The edge set of the complete graph on n vertices cannot be partitioned into fewer than $n-1$ complete bipartite subgraphs.*

As the edges of K_n can be partitioned into $n-1$ bicliques using edge-disjoint stars (there are also many other ways, cf. Babai and Frankl [2] Exercise 1.4.5), the Graham-Pollak Theorem gives the result $\mathrm{bp}(K_n) = n-1$. Since Graham and Pollak's result, many other proofs of this fact have been discovered. Though the result is purely combinatorial, most of the proofs are algebraic, including proofs by G.W. Peck [9], Tverberg [10], and Vishwanathan [11]. Vishwanathan [12] also discovered a proof that replaces linear algebraic techniques by using the pigeon-hole principle in a way that does not necessitate the use of an underlying field. However, his proof uses intermediate structures of large size (on the order of n^n). He asked whether there was a "better" combinatorial proof. Recently, Chung [3] answered this question in the affirmative, giving an elementary proof.

In this note, we describe a beautiful result, attributed Witsenhausen [6], which gives the Graham-Pollak Theorem as a corollary.

Theorem 2 (Witsenhausen, 1980s)**.** *Let G be a finite, loopless graph, and A its adjacency matrix. Then, if $n_+(A)$ and $n_-(A)$ denote the number of positive eigenvalues and negative eigenvalues of A respectively,*

$$\mathrm{bp}(G) \geq \max(n_+(A), n_-(A)).$$

Since K_n has eigenvalue -1 with multiplicity $n-1$, the Graham-Pollak Theorem is a corollary.

2 Sketch of the proof

In this section, we sketch the proof of Witsenhausen's result.

Proof. Assume the edge set of a graph G is partitioned into $\mathrm{bp}(G)$ bicliques. If S is a subset of the vertices of G, then the characteristic vector of S is the n-dimensional $(0,1)$ column vector whose i-th position equals 1 if vertex i is in S and equals 0 otherwise. Denote by u_i and v_i the characteristic vectors of the partite sets of the i-th biclique of our decomposition. Define $D_i = u_i v_i^T + v_i u_i^T$. Then D_i is the adjacency matrix of the i-th biclique as a subgraph of G, and $A = \sum_{i=1}^{\mathrm{bp}(G)} D_i$. Let

$$W = Span\{w \in \mathbb{R}^n | w^T u_i = 0, \forall 1 \leq i \leq \mathrm{bp}(G)\}$$
$$P = Span\{\text{Eigenvectors of the positive eigenvalues of } A\}.$$

Since W is made up of n-dimensional vectors that are all orthogonal to $\mathrm{bp}(G)$ vectors, we have that $dim(W) \geq n - \mathrm{bp}(G)$. On the other hand, since $p^T A p > 0$ for all nonzero $p \in P$, we have that $W \cap P = \{0\}$. Therefore

$$dim(W) \leq n - dim(P) = n - n_+(A).$$

It follows that $n - \mathrm{bp}(G) \leq dim(W) \leq n - n_+(A)$ which implies that $\mathrm{bp}(G) \geq n_+(A)$. The argument for $n_-(A)$ follows similarly. Thus $\mathrm{bp}(G) \geq \max\{n_+(A), n_-(A)\}$. □

3 Motivation and reflections

In this section, we discuss the eigenvalue approach to proving the theorem, and then we discuss motivations and applications. First we note again that although the statement of Graham and Pollak's theorem is combinatorial, Witsenhausen's proof relies on linear algebra. In the same vein, we reflect that one benefit of this proof is that we can obtain bounds on one graph invariant (biclique partition number) from another (eigenvalues). Since biclique partition number is a hard parameter to compute, this is valuable.

Secondly, the eigenvalue approach is valuable because we have proved a statement more general than the Graham-Pollak Theorem. This theorem gives a lower bound for the biclique partition number of a general graph. Indeed, one can pair Witsenhausen's theorem with easy constructions to obtain exact or asymptocally exact values for the biclique partition number of many classes of graphs.

Another advantage to using this approach above others is that it is generalizable past what we have already discussed. One can relax the requirement of partitioning the edge set of G to a more lenient condition of covering every edge of $E(G)$ with no edge being covered more than t times, for some fixed $t \in \mathbb{N}$. Precisely, let $\mathrm{bp}_t(G)$ denote the minimum number of bicliques necessary to cover $E(G)$ such that no edge is covered more than t times. Alon investigated this parameter for $G = K_n$, and showed that such a covering of K_n is equivalent to finding the maximum number of boxes in $\mathbb{R}^n$ that are t-neighborly. In Witsenhausen's proof, we saw that one of the keys to the argument was writing $A(G)$ as a sum of the adjacency matrices of bicliques. In fact, one notices that the argument only requires that $A(G)$ is a linear combination of matrices of bicliques. Thus, using inclusion-exclusion, one can obtain lower bounds on $\mathrm{bp}_t(G)$ as well (see [7] for the best known lower bound). One can also use the same method to obtain a lower bound on the number of bicliques necessary such that the number of times every edge is covered comes from a specified list of integers [5]. Another application of this proof can be seen in the following exercise.

Exercise 1. *Adapt Witsenhausen's proof to show that the minimum number of at most r-partite graphs necessary to partition the edge set of a graph G is at least* $\frac{1}{r-1}\max(n_+(A(G)), n_-(A(G)))$.

Besides the above geometric applications of the Graham-Pollak Theorem, we note two applications from computer science that were alluded to earlier. First we discuss the application to loop-switching that originally motivated the theorem. Consider the problem of communicating among computers. Imagine that the terminals are on one-way communication loops, that are connected at various switching points. If a message needs to be sent from one communication loop to another, it proceeds to a suitable switching point where it may choose to enter a different loop. This process continues until the message reaches its destination. A proposed protocol to decide where a message should switch is by addressing each loop with a string of values from the set $\{0, 1, *\}$ in such a way that the minimum number of loop switches necessary to go from one loop to another is the same as the number of positions where one loop has a 1 and the other has a 0. A message will switch at a switching point if and only if doing so decreases its "Hamming distance" to its destination (we call this a

Hamming distance because of its similarity to the classic Hamming distance, but strictly speaking it is not a metric).

The first question is is such an addressing possible? And if it is possible how long must each address be? One can identify a system of loops with vertices, and make an edge if there is a switching point between the corresponding loops. Graham and Pollak noticed that if G is the graph associated with a loop-system, then the minimum length of the addresses on the loops is the same as the minimum number of bicliques necessary to partition the distance multigraph corresponding to G, where the number of edges between two vertices is their distance in G.

Another application is to *perfect hashing* from computer science. The question can be phrased as follows. Given integers n, k, r with $n \geq r \geq k$, what is the minimum number of functions $f_i : [n] \to [r]$ such that for every $K \subset [n]$ with $|K| = k$, there exists an i such that f_i is injective when restricted to K. If we add the word *unique* before the condition that there exists such a function, we are asking for the minimum number of complete r-partite k-uniform hypergraphs necessary to partition the edge set of the complete k-uniform hypergraph. Thus taking $r = k = 2$ is asking for $\mathrm{bp}(K_n)$.

4 Open Problems

In this note, we discussed an eigenvalue based proof of the Graham-Pollak Theorem as well as some of its applications. The two most interesting open problems, in the opinion of the author, are the following.

Open Problem 1. *What is the minimum number of bicliques necessary to cover every edge of K_n at least once and at most twice? That is, what is $\mathrm{bp}_2(K_n)$?*

The best known bounds for this problem are given by $\sqrt{n-1} \leq \mathrm{bp}_2(K_n) \leq \lceil \sqrt{n} \rceil + \lfloor \sqrt{n} \rfloor - 2$ [1, 7].

Open Problem 2. *What is the minimum number of complete r-partite r-uniform hypergraphs necessary to partition the edge set of the complete r-uniform hypergraph?*

This question seems to be extremely difficult, and the best bounds [4] are far apart. It would be nice to use an eigenvalue approach here as well, but the eigenvalue theory for higher order tensors is still not fully developed.

References

[1] N. Alon, Neighborly families of boxes and bipartite coverings, *The Mathematics of Paul Erdös, R.L. Graham and J. Nesteril*, Springer Verlag, Vol. 2, Berlin, pp. 27 – 31, 1997.

[2] L. Babai and P. Frankl, *Linear Algebraic Methods in Combinatorics*, Lecture Notes, University of Chicago.

[3] F.R.K. Chung, private communication.

[4] S.M. Cioabă, A. Kündgen, and J. Verstraetë, On decompositions of complete hypergraphs, *J. Combin. Thoery Ser. A*, **116**, pp. 1232 – 1234, 2009.

[5] S.M. Cioabă and M. Tait, Variations on a theme of Graham and Pollak, *Discrete Math*, **313**, pp. 665 – 676, 2013.

[6] R.L. Graham and H.O. Pollak, On the addressing problem for loop switching, *Bell Syst. Tech. J.* **50** (8), pp. 2495 – 2519, 1971.

[7] H. Huang and B. Sudakov, *Combinatorica*, **32**, pp. 205 – 219, 2012.

[8] R.L. Graham and H.O. Pollak, On embedding graphs in squashed cubes, *Graph theory and applications*, Lecture Notes in Math., Vol. 303, Springer, Berlin, pp. 99 – 110, 1972.

[9] G.W. Peck, A new proof of a theorem of Graham and Pollak, *Discrete Math*, **49**, pp. 327 – 328, 1984.

[10] H. Tverberg, On the decomposition of K_n into complete bipartite graphs, *J. Graph Theory*, **6**, pp. 493 – 494, 1982.

[11] S. Vishwanathan, A polynomial space proof of the Graham-Pollak Theorem, *J. Combin. Theory Ser. A*, **115**, pp. 674 – 676, 2008.

[12] S. Vishwanathan, A counting proof of the Graham-Pollak Theorem, *Discrete Math*, **313**, pp. 765 –766, 2013.