1 Bipartite forbidden subgraphs

We have seen the Erdős-Stone theorem which says that given a forbidden subgraph $H$, the extremal number of edges is $ex(n, H) = \frac{1}{2}(1 - 1/(\chi(H) - 1) + o(1))n^2$. Here, $o(1)$ means a term tending to zero as $n \to \infty$. This basically resolves the question for forbidden subgraphs $H$ of chromatic number at least 3, since then the answer is roughly $cn^2$ for some constant $c > 0$. However, for bipartite forbidden subgraphs, $\chi(H) = 2$, this answer is not satisfactory, because we get $ex(n, H) = o(n^2)$, which does not determine the order of $ex(n, H)$. Hence, bipartite graphs form the most interesting class of forbidden subgraphs.

2 Graphs without any 4-cycle

Let us start with the first non-trivial case where $H$ is bipartite, $H = C_4$. I.e., the question is how many edges $G$ can have before a 4-cycle appears. The answer is roughly $n^{3/2}$.

**Theorem 1.** For any graph $G$ on $n$ vertices, not containing a 4-cycle,

$$E(G) \leq \frac{1}{4}(1 + \sqrt{4n - 3})n.$$  

**Proof.** Let $d_v$ denote the degree of $v \in V$. Let $F$ denote the set of “labeled forks”:

$$F = \{(u, v, w) : (u, v) \in E, (u, w) \in E, v \neq w\}.$$  

Note that we do not care whether $(v, w)$ is an edge or not. We count the size of $F$ in two possible ways: First, each vertex $u$ contributes $d_u(d_u - 1)$ forks, since this is the number of choices for $v$ and $w$ among the neighbors of $u$. Hence,

$$|F| = \sum_{u \in V} d_u(d_u - 1).$$

On the other hand, every ordered pair of vertices $(v, w)$ can contribute at most one fork, because otherwise we get a 4-cycle in $G$. Hence,

$$|F| \leq n(n - 1).$$

By combining these two inequalities,

$$n(n - 1) \geq \sum_{u \in V} d_u^2 - \sum_{u \in V} d_u$$
and by applying Cauchy-Schwartz, we get
\[ n(n - 1) \geq \frac{1}{n} \left( \sum_{u \in V} d_u \right)^2 - \sum_{u \in V} d_u = \frac{(2m)^2}{n} - 2m. \]
This yields a quadratic equation \( 4m^2 - 2mn - n^2(n - 1) \leq 0 \). A solution yields the theorem. \( \square \)

This bound can be indeed achieved, i.e. there exist graphs with \( \Omega(n^{3/2}) \) edges, not containing any 4-cycle. One example is the incidence graphs between lines and points of a finite projective plane. We give a similar example here, which is algebraically defined and easier to analyze.

**Example.** Let \( V = \mathbb{Z}_p \times \mathbb{Z}_p \), i.e. vertices are pairs of elements of a finite field \((x, y)\). The number of vertices is \( n = p^2 \). We define a graph \( G \), where \((x, y)\) and \((x', y')\) are joined by an edge, if \( x + x' = yy' \). For each vertex \((x, y)\), there are \( p \) solutions of this equation (pick any \( y' \in \mathbb{Z}_p \) and \( x' \) is uniquely determined). One of these solutions could be \((x, y)\) itself, but in any case \((x, y)\) has at least \( p - 1 \) neighbors. Hence, the number of edges in the graph is \( m \geq \frac{1}{2}p^2(p - 1) = \Omega(n^{3/2}) \).

Finally, observe that there is no 4-cycle in \( G \). Suppose that \((x, y)\) has neighbors \((x_1, y_1)\) and \((x_2, y_2)\). This means \( x + x_1 = yy_1 \) and \( x + x_2 = yy_2 \), therefore \( x_1 - x_2 = y(y_1 - y_2) \). Hence, given \((x_1, y_1) \neq (x_2, y_2)\), \( y \) is determined uniquely, and then \( x \) can be also computed from one of the equations above. So \((x_1, y_1)\) and \((x_2, y_2)\) can have only one shared neighbor, which means there is no \( C_4 \) in the graph.

# 3 Graphs without a complete bipartite subgraph

Observe that another way to view \( C_4 \) is as a complete bipartite subgraph \( K_{2,2} \). More generally, we can ask how many edges force a graph to contain a complete bipartite graph \( K_{t,t} \).

**Theorem 2.** Let \( t \geq 2 \). Then there is a constant \( c > 0 \) such that any graph on \( n \) vertices without \( K_{t,t} \) has at most \( cn^{2-1/t} \) edges.

**Proof.** Let \( G \) be a graph without \( K_{t,t} \), \( V(G) = \{1, 2, \ldots, n\} \) and let \( d_i \) denote the degree of vertex \( i \). The neighborhood of vertex \( i \) contains \( \binom{d_i}{t} \) \( t \)-tuples of vertices. Let’s count such \( t \)-tuples over the neighborhoods of all vertices \( i \). Note that any particular \( t \)-tuple can be counted at most \( t - 1 \) times in this way, otherwise we would get a copy of \( K_{t,t} \). Therefore,
\[ \sum_{i=1}^{n} \binom{d_i}{t} \leq (t - 1) \binom{n}{t}. \]

Observe that the average degree in the graph is much more than \( t \), otherwise we have nothing to prove. Due to the convexity of \( \binom{d_i}{t} \) as a function of \( d_i \), the left-hand side is minimized if all the degrees are equal, \( d_i = 2m/n \). Therefore,
\[ \sum_{i=1}^{n} \frac{d_i}{t} \geq n \frac{2m/n}{t} \geq n \frac{(2m/n - t)t}{t!} \]

\( \Omega(f(n)) \) denotes any function which is lower-bounded by \( cf(n) \) for some constant \( c > 0 \) for sufficiently large \( n \).
and 

\[(t - 1)\binom{n}{t} \leq (t - 1)^{n-1}n^t/t!.
\]

We conclude that

\[n(2m/n - t)^t \leq (t - 1)n^t\]

which means that 

\[m \leq \frac{1}{2}(t - 1)^{1/t}n^{2-1/t} + \frac{1}{2}tn \leq n^{2-1/t} + \frac{1}{2}tn.\]

As an exercise, the reader can generalize the bound above to the following.

**Theorem 3.** Let \(s \geq t \geq 2\). Then for sufficiently large \(n\), any graph on \(n\) vertices without \(K_{s+t}\) has \(O(s^{1/t}n^{2-1/t})\) edges.

Another extremal bound of this type is for forbidden even cycles. (Recall that for a forbidden odd cycle, each vertex has at least \(\rho_n\) edges, hence the number of edges can be as large as \(\frac{1}{2}n^2\).)

**Theorem 4.** If \(G\) has \(n\) vertices and no cycle \(C_{2k}\), then the number of edges is \(m \leq cn^{1+1/k}\) for some constant \(c > 0\).

We prove a weaker version of this bound, for graphs that do not contain any cycle of length at most \(2k\).

**Theorem 5.** If \(G\) has \(n\) vertices and no cycles of length shorter than \(2k + 1\), then the number of edges is \(m < n(n^{1/k} + 1)\).

**Proof.** Let \(\rho(G) = |E(G)|/|V(G)|\) denote the density of a graph \(G\). First, we show that there is a subgraph \(G'\) where every vertex has degree at least \(\rho(G)\): Let \(G'\) be a graph of maximum density among all subgraphs of \(G\) (certainly \(\rho(G') \geq \rho(G)\)). We claim that all degrees in \(G'\) are at least \(\rho(G')\). If not, suppose \(G'\) has \(n'\) vertices and \(m' = \rho(G')n'\) edges; then by removing a vertex of degree \(d' < \rho(G')\), we obtain a subgraph \(G''\) with \(n'' = n' - 1\) vertices and \(m'' = m' - d' > \rho(G')(n' - 1)\) edges, hence \(\rho(G'') = m''/n'' > \rho(G')\) which is a contradiction.

Now consider a graph \(G\) with \(m \geq n(n^{1/k} + 1)\) edges and its subgraph \(G'\) of maximum density, where all degrees are at least \(\rho(G) \geq n^{1/k} + 1\). We start from any vertex \(v_0\) and grow a tree where on each level \(L_j\) we include all the neighbors of vertices in \(L_{j-1}\). As long as we do not encounter any cycle, each vertex has at least \(n^{1/k}\) new children and \(|L_j| \geq n^{1/k}|L_{j-1}|\). Assuming that there is no cycle of length shorter than \(2k + 1\), we can grow this tree up to level \(L_k\) and we have \(|L_k| \geq n\). However, this contradicts the fact that the levels should be disjoint and all contained in a graph on \(n\) vertices. \(\square\)

## 4 Application to additive number theory

The following type of question is studied in additive number theory. Suppose we have a set of integers \(B\) and we want to generate \(B\) by forming sums of numbers from a smaller set \(A\). How small can \(A\) be?

More specifically, suppose we would like to generate a certain sequence of squares, \(B = \{1^2, 2^2, 3^2, \ldots, m^2\}\), by taking sums of pairs of numbers, \(A + A = \{a + b : a, b \in A\}\). How small can \(A\) be so that \(B \subseteq A + A\)? Obviously, we need \(|A| \geq \sqrt{m}\) to generate any set of \(m\) numbers.

**Theorem 6.** For any set \(A\) such that \(B = \{1^2, 2^2, 3^2, \ldots, m^2\} \subseteq A + A\), we need \(|A| \geq m^{2/3-o(1)}\).
Proof. Let $B \subseteq A + A$ and suppose $|A| = n$. We define a graph $G$ whose vertices are $A$ and $(a_1, a_2)$ is an edge if $a_1 + a_2 = x^2$ for some integer $x$. Since we need to generate $m$ different squares, the number of edges is at least $m$.

Consider $a_1, a_2 \in A$ and all numbers $b$ such that $a_1 + b = x^2$ and $a_2 + b = y^2$. Note that we get a different pair $(x, y)$ for each $b$. Then, $a_1 - a_2 = x^2 - y^2 = (x + y)(x - y)$. Now, $(x + y, x - y)$ cannot be the same pair for different numbers $b$. Denoting the number of divisors of $a_1 - a_2$ by $d$, we can have at most $\binom{d}{2}$ such possible pairs, and each of them can be used only for one number $b$. Now we use the following proposition.

**Proposition.** For any $\epsilon > 0$ and $n$ large enough, $n$ has less than $n'$ divisors.

This can be proved by considering the prime decomposition of $n = \prod_{i=1}^{t} p_i^{\alpha_i}$, where the number of divisors is $d = \prod_{i=1}^{t} (1 + \alpha_i)$. We assume $\alpha_i \geq 1$ for all $i$. We claim that for any fixed $\epsilon > 0$ and $n$ large enough,

$$\phi(n) = \frac{\log d}{\log n} = \frac{\sum_{i=1}^{t} \log(1 + \alpha_i)}{\sum_{i=1}^{t} \alpha_i \log p_i} < \epsilon.$$  

Observe that $\frac{\log(1+\alpha_i)}{\alpha_i \log p_i}$ can be larger than $\epsilon/2$ only if $p_i^{\alpha_i} \leq (1+\alpha_i)^{2/\epsilon}$, and this can be true only if $p_i$ and $\alpha_i$ are bounded by some constants $P_\epsilon, A_\epsilon$. All such factors together contribute only a constant $C_\epsilon$ in the decomposition of $n$. For $n$ arbitrarily large, a majority of the terms $\log(1+\alpha_i)$ will be smaller than $\frac{\epsilon}{2} \alpha_i \log p_i$ and hence $\phi(n)$ will drop below $\epsilon$ for sufficiently large $n$.

To summarize, for any pair $a_1, a_2 \in A$, we have less than $n^2\epsilon$ numbers $b$ which are neighbors of both $a_1$ and $a_2$ in the graph $G$. In other words, $G$ does not contain $K_{2,n^2\epsilon}$. By our extremal bound, it has at most $cn^{3/2+\epsilon}$ edges. I.e., $m \leq cn^{3/2+\epsilon}$, for any fixed $\epsilon > 0$.\hfill $\square$