There are a number of important consequences of the Kővári–Sós–Turán theorem. The first is that it immediately gives us a bound on $\mathrm{ex}(n, H)$ for all bipartite H. Indeed, note that if H_1 is a subgraph of H_2, then

$$\mathrm{ex}(n, H_1) \leqslant \mathrm{ex}(n, H_2)$$

for all n, as any H_1-free graph is also H_2-free. Now, if H is a bipartite graph, then H is a subgraph of $K_{s,t}$ for some $s \leqslant t$. So

$$\mathrm{ex}(n, H) \leqslant \mathrm{ex}(n, K_{s,t}) \leqslant O(n^{2-1/s}).$$

In particular, this proves that $\mathrm{ex}(n, H) = o(n^2) = o(1) \cdot \binom{n}{2}$ for bipartite H. Indeed,

$$\lim_{n\to\infty} \frac{\mathrm{ex}(n, H)}{n^2} \leqslant \lim_{n\to\infty} \frac{O(n^{2-1/s})}{n^2} = \lim_{n\to\infty} O(n^{-1/s}) = 0.$$

Recall that this was a consequence of the Erdős–Stone–Simonovits theorem.

5.2 Lower bounds

How good are the upper bounds we proved? Let's begin with the one we started with, $\mathrm{ex}(n, C_4) \leqslant O(n^{3/2})$. Can we construct an n-vertex C_4-free graph with roughly that many edges?

As it turns out, we can! The following construction was originally due to Eszter Klein in 1938 (as reported in a paper of Erdős). Note that this is before Turán's theorem, so before the birth of extremal graph theory! As such, no one really appreciated what this construction was or meant, and it was later rediscovered by Erdős, Rényi, and Sós (and independently Brown). These days, it is often called the "Erdős–Rényi" construction, which I find a little odd, both because they weren't the first to discover it, and because there are many other things named after Erdős and Rényi.

Theorem 5.4 (Klein 1938)**.** *For every $n \geqslant 1$, there is an n-vertex C_4-free graph with at least $n^{3/2}/64$ edges.*

Proof. First, suppose that $n = 2p^2$ for some prime p; we will later get rid of this assumption. Consider the integers mod p, which form a field that we denote $\mathbb{F}_p$. (If you don't know what the word "field" means, just believe me that among the integers mod p, we can use addition, multiplication, and division and have them work basically the same way they do in $\mathbb{R}$.)

Let $\mathbb{F}_p^2$ denote the two-dimensional plane over $\mathbb{F}_p$, i.e. the set of points (x, y) with $x, y \in \mathbb{F}_p$. For $m, b \in \mathbb{F}_p$, let $\ell_{m,b}$ denote the line $y = mx + b$ in $\mathbb{F}_p^2$. In other words, $\ell_{m,b}$ is the set of points $(x, y) \in \mathbb{F}_p^2$ satisfying $y = mx + b$.

We define a bipartite graph G with parts P, L, where $P = \mathbb{F}_p^2$ and $L = \{\ell_{m,b} : m, b \in \mathbb{F}_p\}$. The edges of G are given by *incidence*: we connect $(x, y) \in P$ to $\ell_{m,b} \in L$ if and only if (x, y) lies on the line $\ell_{m,b}$, i.e. if and only if $y = mx + b$.

Note that $|P| = |L| = p^2$, so G has $n = 2p^2$ vertices. Moreover, every line $\ell_{m,b} \in L$ has exactly p points on it, so every vertex in L has degree p in G. Therefore, $e(G) = p|L| = p^3 = (n/2)^{3/2}$.

Finally, we claim that G is C_4-free. To see this, note that G is bipartite, so the only way we could have a copy of C_4 in G is to have distinct $p_1, p_2 \in P$ and distinct $\ell_1 \ell_2 \in L$ so that $p_1 \ell_1 p_2 \ell_2$ forms a 4-cycle. But this means that p_1 lies on the lines ℓ_1, ℓ_2, and that p_2 also lies on both these lines. So we have two lines which intersect at two distinct points!

Using our intuition from $\mathbb{R}$, we expect this to be impossible, and it's impossible over $\mathbb{F}_p$ as well. Formally, let $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$, and $\ell_1 = \ell_{m_1,b_1}, \ell_2 = \ell_{m_2,b_2}$. Then we have the equations

$$\begin{aligned} y_1 &= m_1 x_1 + b_1 & \qquad y_2 &= m_1 x_2 + b_1 \\ y_1 &= m_2 x_1 + b_2 & \qquad y_2 &= m_2 x_2 + b_2 \end{aligned}$$

Rearranging the first column, we see that $m_1 x_1 + b_1 = m_2 x_1 + b_2$, or equivalently that $(m_2 - m_1)x_1 = b_1 - b_2$. If $m_1 = m_2$ then this implies that $b_1 = b_2$, contradicting that ℓ_1, ℓ_2 are distinct. So we have that $m_1 \neq m_2$, so $x_1 = (b_1 - b_2)/(m_2 - m_1)$. But from the second column of equations, we conclude that $x_2 = (b_1 - b_2)/(m_2 - m_1)$ as well, so $x_1 = x_2$. But if we plug this into any of the equations, we conclude that $y_1 = y_2$, and thus that $p_1 = p_2$, a contradiction. So G is C_4-free.

The only remaining thing is to deal with the fact that n need not equal twice the square of a prime. So let n be arbitrary. There is an important result in number theory, called Bertrand's postulate, which says that there is always a prime between m and $2m$ for all positive integers m. Let $m = \lfloor \sqrt{n}/4 \rfloor$, and let p be a prime between m and $2m$, so that $n/8 \leqslant 2p^2 \leqslant n$. Using the construction above, we obtain a C_4-free graph G on $2p^2$ vertices with p^3 edges. We add to this graph $n - 2p^2$ isolated vertices, and we obtain a new C_4-free n-vertex graph with $p^3 \geqslant (n/16)^{3/2} = n^{3/2}/64$ edges. □

Using these finite fields and finite geometries might seem like a neat trick, but it turns out that it's essentially the only thing one can do. Indeed, all constructions we know of for C_4-free graphs with many edges use such techniques. Moreover, there is a powerful result of Füredi, which says that for those n for which such a construction (appropriately defined) exists, the *unique* C_4-free n-vertex graph with the most edges comes from such a construction.

So we conclude that $\mathrm{ex}(n, C_4) = \Theta(n^{3/2})$, where the big-$\Theta$ means that we have upper and lower bounds that agree up to a constant factor. Since $\mathrm{ex}(n, K_{2,t}) \geqslant \mathrm{ex}(n, C_4)$ for all $t \geqslant 2$, we conclude that $\mathrm{ex}(n, K_{2,t}) = \Theta(n^{3/2})$ for all $t \geqslant 2$.

What about $\mathrm{ex}(n, K_{3,3})$? We proved in Theorem 5.3 that $\mathrm{ex}(n, K_{3,3}) \leqslant O(n^{5/3})$. As it turns out, this is also tight.

Theorem 5.5 (Brown 1966)**.** *For every n, there exists an n-vertex $K_{3,3}$-free graph G with $n^{5/3}/100$ edges.*

Proof sketch. I won't present the proof in detail, but will explain the big idea. Suppose that $n = p^3$. Construct a graph G with vertex set $\mathbb{F}_p^3$, where we connect two vertices (x, y, z) and

(x', y', z') by an edge if and only if

$$(x-x')^2 + (y-y')^2 + (z-z')^2 = 1.$$

In other words, the neighborhood of any vertex looks like a "unit sphere" centered at that vertex, except that "spheres" don't really exist over $\mathbb{F}_p$.

Nonetheless, if we were working in $\mathbb{R}^3$, then we'd expect that any three unit spheres can intersect in at most two points: two unit spheres can intersect in a circle, and that circle can intersect a thid unit sphere in only two points. So we'd expect that G is $K_{3,3}$-free, since any three vertices have at most two common neighbors.

Since we expect a sphere to be "two-dimensional", we should expect every unit sphere to have roughly p^2 points on it, and this turns out to be true. So G has $n = p^3$ vertices, and every vertex has degree around p^2, so we expect $e(G) \approx p^5 = n^{5/3}$.

All of this intuition can be made precise, some of it with some annoyance. For example, it turns out that this only really works if $p \equiv 3 \pmod 4$. But the high-level idea is correct. □

So we see that the Kővári–Sós–Turán theorem is best possible (up to the constant factor) for $s = 2$ and $s = 3$. The case of $s = 1$ is much easier, but it's also best possible there, as you'll see on the homework. So it is natural to conjecture, as many have done, that the Kővári–Sós–Turán theorem is tight in general.

Conjecture 5.6 (Many people)**.** *For all $s \leqslant t$,*

$$\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s}).$$

Moreover, based on what I've told you so far, it is natural to expect that not only is this conjecture proved, but that the constructions look kind of the same as above. You work with the s-dimensional space $\mathbb{F}_p^s$ over the field $\mathbb{F}_p$, and use some kind of cleverly chosen polynomial or set of polynomial equations to define the adjacency condition. However, despite many people having this same idea, Conjecture 5.6 remains unproved. Moreover, many experts in the field now even question whether it is true.

Nonetheless, some other things are known about $\mathrm{ex}(n, K_{s,t})$. Namely, it is known that $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$ if t is sufficiently large compared to s. The first result of this type is due to Kollár, Rónyai, and Szabó in 1996, who proved that

$$\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s}) \qquad \text{if } t > s!.$$

To do this, they constructed a $K_{s,t}$-free graph, again using the space $\mathbb{F}_p^s$, called the *norm graph*. Their construction was later modified by Alon, Rónyai, and Szabó in 1999, who defined a similar graph called the *projective norm graph* (again over $\mathbb{F}_p^s$), which implies that

$$\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s}) \qquad \text{if } t > (s-1)!.$$

So, for example, we know that $\mathrm{ex}(n, K_{4,7}) = \Theta(n^{7/4})$, but have no such lower bound for $\mathrm{ex}(n, K_{4,4})$.

For about 20 years, the Alon–Rónayi–Szabó result was the best known. But very recently, Bukh proved the following theorem.

Theorem 5.7 (Bukh 2021)**.** *Suppose $s \geqslant 2$ and $t \geqslant 9^s \cdot s^{4s^{2/3}}$ are integers. Then*

$$\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s}).$$

The key point is that for large s, the previous bound on t, namely $(s-1)!$, grew super-exponentially in s. But Bukh's bound, for large s, grows merely exponentially in s. The key to Bukh's construction is again to work over $\mathbb{F}_p^s$, but not to pick a *clever* polynomial. Instead, he picks a *random* polynomial, and then uses arguments from probability, combinatorics, and algebraic geometry to prove that the resulting graph is $K_{s,t}$-free with positive probability.

While similar algebraic constructions exist for certain specific bipartite H, there is essentially only one general-purpose lower bound that is known. In general, algebraic techniques like the ones described above are the best techniques we have for constructing lower bounds, but they often rely on specific structures that we can exploit. The following bound holds for any bipartite graph.

Given a graph H, we define its *2-density* to be

$$m_2(H) \coloneqq \max_{F \subseteq H} \frac{e(F) - 1}{v(F) - 2}.$$

Theorem 5.8. *For any bipartite H, we have*

$$\mathrm{ex}(n, H) \geqslant \Omega(n^{2-1/m_2(H)}).$$

The proof of Theorem 5.8 uses the probabilistic method, and I won't cover it in class. But at a high level, the idea is to pick a *random* graph G with n vertices and roughly $n^{2-1/m_2(H)}$ edges. One can then show that with positive probability, the number of copies of H in G is less than half the number of edges of G. By deleting a single edge from each copy of H, we obtain a graph with half as many edges—so still $\Omega(n^{2-1/m_2(H)})$—and no copy of H.

6 Extremal numbers of hypergraphs

It's time for everything to get more *hyper*.

If we go back to bare basics, a graph is a collection V of vertices, plus a collection E of edges, which are simply unordered pairs of vertices. Why restrict ourselves to pairs?

Definition 6.1. A *k-uniform hypergraph* (sometimes called an k-graph for short) consists of a finite collection V of *vertices*, as well as a collection E of *k-uniform hyperedges*, which are simply subsets of V of size k.

As with graphs, we say that one k-graph $\mathcal{H}$ is a *subhypergraph* (or simply *subgraph*) of another k-graph $\mathcal{G}$ if we can obtain $\mathcal{H}$ from $\mathcal{G}$ by deleting some vertices and edges. We say that $\mathcal{G}$ is *$\mathcal{H}$-free* if $\mathcal{G}$ does not contain $\mathcal{H}$ as a subgraph (and we also say that $\mathcal{G}$ has no copy of $\mathcal{H}$).

As with graphs, we define the *extremal number* of $\mathcal{H}$ as

$$\mathrm{ex}(n, \mathcal{H}) = \max\{e(\mathcal{G}) : \mathcal{G} \text{ is an } n\text{-vertex } \mathcal{H}\text{-free } k\text{-graph}\}.$$

In contrast to graphs (the case $k = 2$), we know extraordinarily little about $\text{ex}(n, \mathcal{H})$ for k-graphs $\mathcal{H}$ with $k \geqslant 3$. For example, even the hypergraph analogue of Mantel's theorem is a famous open problem. To explain this formally, we make the following definition.

Definition 6.2. The *complete k-graph* on r vertices, denoted $K_r^{(k)}$, is the k-graph with r vertices whose edge set consists of all subsets of size k.

Then the amazing fact is that for any $k > r \geqslant 3$, we do not know the value of $\text{ex}(n, K_r^{(k)})$. For literally no pair of (r, k)! This problem was proposed by Turán already in 1941, and he made the following conjecture, which is a natural analogue of Mantel's theorem.

Conjecture 6.3 (Turán 1941)**.**

$$\text{ex}(n, K_4^{(3)}) = \left(\frac{5}{9} + o(1)\right)\binom{n}{3}.$$

The reason for 5/9 is a specific construction of an n-vertex $K_4^{(3)}$-free 3-graph, which Turán came up with, and which was the best he could come up with. You'll see Turán's construction on the homework.

Erdős offered \$500 for the resolution of Conjecture 6.3, and \$1000 for a general formula for $\text{ex}(n, K_r^{(k)})$. So far, very little progress has been made on these questions. The best known bound for $\text{ex}(n, K_4^{(3)})$ is due to Razborov, who proved that

$$\text{ex}(n, K_4^{(3)}) \leqslant (0.561666 + o(1))\binom{n}{3}.$$

Note that $5/9 = 0.555\ldots$, so this is pretty close to Turán's conjecture. Unfortunately, Razborov's technique is unlikely to yield the full resolution of Conjecture 6.3, because his technique uses a computer to do complicated computations to what is essentially a "finite approximation" to the problem.

In general, the best known lower bound is due to de Caen, who proved that

$$\text{ex}(n, K_r^{(k)}) \leqslant \left(1 - \frac{1}{\binom{r-1}{k-1}} + o(1)\right)\binom{n}{k}.$$

The best known general lower bound, due to Sidorenko, is

$$\text{ex}(n, K_r^{(k)}) \geqslant \left(1 - \left(\frac{k-1}{r-1}\right)^{k-1} + o(1)\right)\binom{n}{k}.$$

In the case of $k = 3$, this says that

$$\text{ex}(n, K_r^{(3)}) \geqslant \left(1 - \left(\frac{2}{r-1}\right)^{2} + o(1)\right)\binom{n}{3},$$

and this was conjectured to be optimal by Turán. You'll see the construction in the homework.