Chapter 2

Complete Bipartite Graphs

Complete bipartite graphs are structurally very simple, but we will still see how the study of their Turán numbers leads us into unchartered geometric and algebraic territory and leaves us with many open questions.

2.1 Forbidding $K_{2,s}$

The following upper bound on the Turán number of $K_{2,2}$ appeared even earlier than Turán's Theorem, as an auxiliary lemma in a number theoretic paper of Paul Erdős. Then it went into relative oblivion as a combinatorial statement. Paul Erdős later refered to it as the point when he "missed discovering Extremal Graph Theory".

Theorem 2.1 (Erdős; 1938) *We have*

$$ex(n, K_{2,2}) \leq \frac{1}{2}n^{3/2} + \frac{1}{4}n.$$

Proof. Take a graph $G \not\supseteq K_{2,2}$. Let us doublecount its "cherries", i.e., its subgraphs isomorphic to $K_{2,1}$.

On the one hand, for a fixed vertex v there are $\binom{d(v)}{2}$ cherries centered at it. The total number of cherries is then equal to $\sum_v \binom{d(v)}{2}$.

On the other hand, since there is no $K_{2,2}$ in G, for every pair of vertices u, v, there is at most one cherry whose endpoints are u and v. We have

$$\sum_v \binom{d(v)}{2} = \# \text{ of cherries } \leq 1 \cdot \binom{n}{2}.$$

By Jensen's inequality ($\binom{x}{2} = \frac{x(x-1)}{2} : \mathbb{R} \to \mathbb{R}$ is convex!) we obtain

$$\sum_v \binom{d(v)}{2} \geq n\binom{\bar{d}(G)}{2},$$

where $\bar{d}(G) = \sum_v d(v)/n = 2e(G)/n$ is the average degree of G. These imply $\bar{d}(G) \leq \frac{\sqrt{4n-3}+1}{2} \leq \sqrt{n} + \frac{1}{2}$, and hence

$$e(G) = \frac{n\bar{d}(G)}{2} \leq \frac{1}{2}n^{3/2} + \frac{1}{4}n.$$

□

The natural question arises: Is this upper bound on the Turán number of $K_{2,2}$ a good one? Is there a $K_{2,2}$-free graph with (order of) $n^{3/2}$ edges?

2.1.1 How about a randomized construction?

Still elated by the success of the random graph for the symmetric Ramsey problem, let us try some natural randomized approaches.

We look at the random graph $G(n,p)$ (i.e. the probability space of graphs, where each of the $\binom{n}{2}$ edges appears with probability p, independently from all other edges).

The first thought is to determine the largest function $p = p(n)$ for which the random graph does not contain a $K_{2,2}$ with positive probability, or rather with probability tending to a strictly positive number. For this, an always good indication is to calculate the expected number of (unlabeled) $K_{2,2}$s in $G(n,p)$. Let $\#K_{2,2}$ denote the random variable representing the number of unlabeled copies of $K_{2,2}$ in $G(n,p)$.

$$\mathsf{E}[\#K_{2,2}] = \frac{n(n-1)(n-2)(n-3)}{8}p^4.$$

For $p = \frac{1}{n}$ this expectation is less than 1, hence there is an instance of a graph in $G(n,p)$ which is $K_{2,2}$-free. "Of course there is" – one can say – "the empty graph for example!" The empty graph though would not help greatly our quest for a dense $K_{2,2}$-free graph. But random graph theory tells us more: once this expectation is $o(1)$, then $G(n,p)$ is $K_{2,2}$-free with probability tending to 1. Hence for large n with probability at least 0.99 $G(n,p)$ is $K_{2,2}$-free, provided $p = o(1/n)$. Random graph theory also tells us that for large n with probability at least 0.99 $G(n,p)$ has at least $\frac{1}{2}\binom{n}{2}p$ edges. Thus with probability at least 0.98 the random graph $G(n,p)$ is $K_{2,2}$-free *and* has an almost linear number of edges! In particular there is an instance of such a graph. Wow! Short meditation tarnishes the enthusiasm a bit, when one comes up with an explicit construction of a $K_{2,2}$-free graph with $n-1$ edges: the path P_n. Unfortunately we cannot really hope for more from the random graph in this simple manner: when $\mathsf{E}[\#K_{2,2}] \to \infty$, i.e. for $p \gg \frac{1}{n}$, the probability that $G(n,p)$ is $K_{2,2}$-free tends to 0. These are standard facts from random graph theory, which we don't prove here, because they show that *most* graphs with $\gg \binom{n}{2} \cdot \frac{1}{n} = \Theta(n)$ edges are *not* $K_{2,2}$-free. That is, should we want hit upon a denser $K_{2,2}$-free graph, we indeed need to find a "needle in the haystack".

The Alteration Method

Despite these discouraging news on the random front one can still try to put a twist on the random approach. The idea is not to shoot for a $K_{2,2}$-free graph right away, but obtain first with random methods a graph G', possibly with much much more than $\Theta(n)$ edges, which contains some $K_{2,2}$s though, but not so many, in fact less than, say, half of the number of its edges. Deleting one edge from each $K_{2,2}$ makes G' $K_{2,2}$-free, and the resulting graph G still has half of the edges of the original graph G'.

This "semi-random" technique is called the *deletion* or *alteration method.*

To this end we would like to find p *as large as possible*, such that

$$\mathsf{E}[e(G(n,p)) - \#K_{2,2}] \geq \frac{1}{2}\binom{n}{2}p.$$

For such a probability p there surely exists an instance of a graph G' for which we have $e(G') - \#\{K_{2,2} \text{ in } G'\} \geq \frac{1}{2}\binom{n}{2}p$. If we remove an edge from each copy of $K_{2,2}$ appearing in G', then we obtain a $K_{2,2}$-free graph in which the number of edges is at least $\frac{1}{2}\binom{n}{2}p$.

To evaluate the hopelessly complicated-looking random variable $e(G(n,p)) - \#K_{2,2}$ we of course use the linearity of expectation.

$$\begin{aligned}\mathsf{E}[e(G(n,p)) - \#K_{2,2}] &= \mathsf{E}[e(G(n,p))] - \mathsf{E}[\#K_{2,2}]\\ &= \binom{n}{2}p - \frac{n(n-1)(n-2)(n-3)}{8}\cdot p^4\\ &\geq \frac{1}{2}\binom{n}{2}p.\end{aligned}$$

Resolving we get

$$1 \geq \frac{(n-2)(n-3)}{2}p^3.$$

Hence for $p = cn^{-2/3}$ (with small enough c), there exists a graph G' for which we have $e(G') - \#\{K_{2,2} \text{ in } G'\} \geq \frac{1}{2}\binom{n}{2}p = c'n^{4/3}$. So after removing an edge from each $K_{2,2}$ of G we have a $K_{2,2}$-free graph with $c'n^{4/3}$ edges.

This is now at least a superlinear number of edges, but still far away from the order $n^{3/2}$ of the upper bound. We don't claim that this simple construction is the only possible random idea one can have, but we certainly don't have any other. We don't know how to use more sophisticated techniques to give a better randomized lower bound.

2.1.2 A heuristic geometric construction of a $K_{2,2}$-free graph with $n^{3/2}$ edges

The graph constructed in this subsection is infinite. You might wonder what $n^{3/2}$ means when n is not finite... Please loosen up for the moment: The whole subsection is somewhat of a nonsense, but nevertheless, it is a good introduction to what comes next.

We construct a bipartite graph G with one partite set A being the set of all points in the (Euclidean) plane, and the other partite set B being the set of all lines in the plane. We join vertices $p \in A$ and $l \in B$ with an edge if and only if the point p is on the line l.

If this graph contained a $K_{2,2}$, that would mean that there are two different lines that have two different points in common. This is of course not possible, and that is why G is $K_{2,2}$-free.

Let us now look at the number of edges. The set of points in the plane is just the set of ordered pairs with both coordinates from $\mathbb{R}$. So the cardinality of A is $|\mathbb{R}|^2$. Every line in the plane can be uniquely represented by its slope and its signed distance from the origin. Therefore $|B|$ is of cardinality $|\mathbb{R} \cup \{\infty\}| \cdot |\mathbb{R}|$. Now this is again roughly $|\mathbb{R}|^2$ (we ignored a lower order term of $|\mathbb{R}|$:-)). So the order of the graph is

$$|V(G)| \approx 2 \cdot |\mathbb{R}|^2.$$

On the other hand, on each line there are $|\mathbb{R}|$ points meaning that every vertex from B has degree $|\mathbb{R}|$. Since G is bipartite the total number of edges in G is

$$|E(G)| = \underbrace{|\mathbb{R} \cup \{\infty\}| \cdot |\mathbb{R}|}_{|B|} \cdot |\mathbb{R}| \approx |\mathbb{R}|^3 \approx \left(\frac{|V(G)|}{2}\right)^{3/2}.$$

2.1.3 Explicit constructions of dense $K_{2,2}$-free graphs

How to make sense of the previous geometric idea, i.e. how to "finitize" it? First one has to realize that the required property, the "$K_{2,2}$-freeness", depends only on the fact that a system of two linear equations in two variables (which are not a constant multiple of each other) has at most one solution. Then the answer is quite obvious: instead of the field of real numbers $\mathbb{R}$, take a finite field $\mathbb{F}_q$, in which the same rules apply regarding solving an equation system. Recall that the q-element field $\mathbb{F}_q$ exists if and only if q is a prime power. For primes p, the set $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ of residue classes with the usual ($\pmod p$) addition and multiplication is the p-element field.

Construction 0. We construct a bipartite graph G with one partite set A being all points in the *affine plane* $\mathbb{F}_q^2$, and the other partite set B being all lines in $\mathbb{F}_q^2$. Lines are defined similarly to lines in $\mathbb{R}^2$, as the set of solutions of a linear equation $a_1x_1 + a_2x_2 = a_3$, where not both of a_1 and a_2 are 0. This graph is $K_{2,2}$-free. Like in the previous section, two different lines do not have two common points, since the corresponding linear equation system (of two equations) has *at most* one solution.

We have q^2 points in A and q^2+q lines in B which make up to the total of $n = 2q^2+q$ vertices of G. The degree of each vertex in B (i.e. the cardinality of each line in $\mathbb{F}^2$) is q. Then,

$$e(G) = (q^2+q) \cdot p = \frac{1}{2\sqrt{2}}n^{3/2} + \frac{1}{8}n - O(\sqrt{n}) \approx \left(\frac{n}{2}\right)^{3/2}.$$

Now we should be very happy, since the number of edges in this construction not only beats the random construction, but matches the order of magnitude of the upper bound. Ever discontent, however, we see it still not satisfactory since the constant multiplier of the leading term is smaller than the $\frac{1}{2}$ from Theorem 2.1.

To fix that, one can have a couple of ideas.

Idea 1. We feel that there is some loss in terms of the number of edges when two lines do *not* intersect. In an affine plane this does happen when the two lines are parallel. Taking a projective plane $PG(q, 2)$ instead, where *every two* lines intersect, might be a good idea.

It turns out that this idea improves only the second order term.

Idea 2. The property of $K_{2,2}$-freeness comes from the structure of lines and points, which in turn depends on solving a system of linear equations. Therefore, we don't have to distinguish between points and lines, we can talk about elements of $\mathbb{F}_q^2$ or $PG(q,2)$ as the vertex set. This will abandon the bipartiteness of the construction and basically keep the degrees intact.
Since suddenly the number of vertices is cut in half and the degrees are preserved, the constant multiplyer of the leading term will be improved.

Before we utilize these two observations to get a better construction, let us recall some basic notions about projective planes, in particular the projective plane $PG(q,2)$ over the p-element field consisting of a set $\mathcal{L}$ of lines and a set $\mathcal{P}$ of points. Recall that $\mathbb{F}_q^*$ denotes the nonzero elements of $\mathbb{F}_q$.

Points: The set of points $\mathcal{P}$ are the equivalence classes of $\mathbb{F}_q^3 \setminus \{(0,0,0)\}$, where two triples are in relation if they are a non-zero constant multiples of each other,

$$[x_0, x_1, x_2] := \left\{(cx_0, cx_1, cx_2) \in \mathbb{F}_q^3 \setminus \{(0,0,0)\} \ : \ c \in \mathbb{F}_q^*\right\} .$$

Hence each equivalence class contains $q-1$ triples and the number of points in the projective plane is $\frac{q^3-1}{q-1} = q^2+q+1$.

Lines: Given a triple $(a_0, a_1, a_2) \in \mathbb{F}_q^3 \setminus \{(0,0,0)\}$ we define the line $L(a_0,a_1,a_2)$ as follows:

$$L(a_0, a_1, a_2) := \left\{[x_0, x_1, x_2] \in \mathcal{P} \ : \ a_0x_0 + a_1x_1 + a_2x_2 = 0\right\} .$$

This definition makes sense, since containment in $L(a_0,a_1,a_2)$ obviously does not depend on which representative of $[x_0,x_1,x_2]$ we use in the equation. The set of lines $\mathcal{L}$ in $PG(q,2)$ consists of all the lines $L(a_0,a_1,a_2)$ defined above. Of course $L(a_0,a_1,a_2) = L(ca_0,ca_1,ca_2)$. So the number of lines is also q^2+q+1.

The projective plane $PG(q,2)$ has the following properties:

- There are $q+1$ lines through a point.
- Every line has $q+1$ points on it.
- Every two (different) lines intersect in exactly one point.
- For any two points there is exactly one line containing them.

Let us now return to improving on Construction 0.

Construction 1. (Eszter Klein; 1938) Utilizing Idea 1., let us first consider the point/line incidence graph G of the projective plane $PG(q,2)$. Then G is a $K_{2,2}$-free, $(q+1)$-regular bipartite graph with partite sets of order q^2+q+1. That is $n(G) = 2(q^2+q+1)$ and

$$e(G) = \underbrace{(q^2+q+1)}_{n/2}(q+1) = \frac{n}{2}\left(\sqrt{\frac{n}{2}-\frac{3}{4}}+\frac{1}{2}\right) = \left(\frac{n}{2}\right)^{3/2} + \frac{n}{4} - O(\sqrt{n}).$$

One can see that there is a slight improvement in the second order term compared to our original construction.

Exercise 2.1 *Define $ex(n, m, H)$ as the largest number e, such that there is an H-free bipartite graph with partite sets of size n and m, respectively, containing e edges. Show that $ex(q^2+q+1, q^2+q+1, K_{2,2}) = (q^2+q+1)(q+1)$ for every prime power q, i.e. the above graph is* an *optimal construction.*

Construction 2. Utilizing Idea 2., we can make our original graph non-bipartite the following way. Define a graph G whose vertex set is $\mathbb{F}_q^2 \setminus \{(0,0)\}$ and the distinct vectors $(a_1, a_2), (b_1, b_2) \in V(G)$ are adjacent if

$$a_1b_1 + a_2b_2 = 1.$$

This graph is easily seen to be $K_{2,2}$-free. Its number of vertices is $n(G) = q^2 - 1$. For the degree one can see that for any (a_1, a_2), $a_1x_1 + a_2x_2 = 1$ has exactly q solutions (x_1, x_2) (each of them non-zero). This would mean that the graph is q-regular, but this is not quite so. One of the solutions could be equal to (a_1, a_2), which would correspond to a *loop* in G. Since we are interested in simple graphs, these edges are not allowed. In any case, every vertex has degree at least $q-1$, so the number of edges in our graph is

$$e(G) \geq \frac{1}{2}(q^2-1)(q-1) = \frac{n^{3/2}}{2} - \frac{n}{2} + O(\sqrt{n}).$$

Exercise 2.2 *Prove that there are exactly $q-1$, or $q+1$ loops in G depending on whether -1 is a quadratic residue in $\mathbb{F}_q$, or not. Conclude that the number of edges in our graph is in fact $n^{3/2}/2 - O(\sqrt{n})$.*

We are extremely happy at this point, since this construction matches our upper bound even in the multiplicative constant of the leading term! We have an asymptotically correct answer!!

Of course there are those people who still want more... Why not shoot for the best and find *the* best construction. For that we need to combine Ideas 1. and 2.

Construction 3. (Brown, 1966; Erdős, Rényi, T. Sós, 1966) We define the so-called *polarity graph* G on the points of the projective plane $PG(q,2) = (\mathcal{P}, \mathcal{L})$ in the following way. Let $V(G) = \{[x_0, x_1, x_2] \in \mathcal{P}\}$ and $E(G) = \{\{x, y\} : x \neq y, x_0y_0 + x_1y_1 + x_2y_2 = 0\}$. By the properties of the projective plane, discussed above, the polarity graph is $K_{2,2}$-free, $n = q^2 + q + 1$ and the degree of any vertex in G is either $q+1$ or q (because of possible loops). In fact we will see later that there are exactly $q+1$ vertices of degree q. Thus the number of edges of G is

$$e(G) = \frac{1}{2}(q^2+q)(q+1) = \frac{1}{2}(n-1)\left(\sqrt{n-\frac{3}{4}} + \frac{1}{2}\right) = \frac{1}{2}n^{3/2} + \frac{1}{4}n + O(\sqrt{n}).$$

Exercise 2.3 *Show that the polarity graph indeed contains exactly $q+1$ vertices of degree q.*

Füredi showed that

$$ex(q^2+q+1, K_{2,2}) = \frac{1}{2}q(q+1)^2$$

for every prime power $q > 13$. Hence the polarity graph is *an* optimal construction. The understanding of the exact value of $ex(n, K_{2,2})$, when $n \neq q^2+q+1$ is quite sparse. Extensive computer searches were performed for small values of n, up to 31.

2.1.4 The upper bound on $ex(n, K_{t,s})$ and an application

A famous question of Erdős is the following: How many unit distances can n points in the plane determine? (Plane now means the usual Euclidean plane.) Erdős offered 500 dollars for a solution. (`http://maven.smith.edu/~orourke/TOPP/node40.html`)

First attempts at a construction would result in a linear number of pairs of points at unit distance. So at first sight it is surprising even that there is a construction of a pointset that has $nf(n)$ unit distances, where $f(n) \to \infty$. The best known constructions have $f(n) = n^{\frac{c}{\log \log n}}$. However, it is conjectured by Erdős that there cannot be more than $n^{1+\varepsilon}$ pairs at unit distance, for any constant $\varepsilon > 0$.

For a point set P in the plane let us define its *unit-distance graph*, which has P as its vertex set and two points are adjacent if their distance is 1. The problem of Erdős is then asking for the maximum number of edges a unit-distance graph can have.

Remark: Another notorious open problem is inquiring about the chromatic number of the unit distance graph of the plane (an infinite graph). Solving this problem is even worth 1000 dollars. `http://maven.smith.edu/~orourke/TOPP/P57.html`

Exercise 2.4 *Prove that the chromatic number of the unit distance graph lies between 4 and 7. (It's worthwhile to note that there are no better bounds known at the moment.)*

Two unit circles intersect in at most two points, thus for every two points in the plane, there are at most two points that are at unit distance from both of them. Hence the unit distance graph on n points does not contain a $K_{2,3}$ as a subgraph.

We are going to give an upper bound on the number of edges of a graph that does not contain a $K_{t,s}$.

Theorem 2.2 (Kővári, T. Sós, Turán; 1954) *Let $s \geq t \geq 2$ be arbitrary integers. Then*

$$ex(n, K_{t,s}) \leq \frac{1}{2}(s-1)^{1/t}n^{2-1/t} + \frac{1}{2}(t-1)n \approx c_{s,t}n^{2-\frac{1}{t}}.$$

Proof. Let G be a graph containing no $K_{t,s}$. We are going to doublecount the number of t-stars (i.e., $K_{1,t}$). For every vertex v there are $\binom{d(v)}{t}$ t-stars centered at v. On the other hand, for every t-tuple T there are at most $(s-1)$ t-stars that have T as endpoints, since otherwise G would contain a $K_{t,s}$. Thus, we get

$$\binom{n}{t}(s-1) \geq \ \# \text{ of } K_{1,t} \text{ in } G = \sum_v \binom{d(v)}{t}.$$

We would like to apply Jensen's inequality, but there is a slight technical problem, since the function $x \to \binom{x}{t} = \frac{x(x-1)\cdots(x-t+1)}{t!}$ is not convex in the interval $[0,\infty)$, only for $x \geq t-1$. For that we assume that $e(G) = ex(n, K_{t,s})$, which implies $\delta(G) \geq t-1$. Then we have

$$(s-1)\frac{n^t}{t!} \geq \sum_v \binom{d(v)}{t} \geq n\binom{\bar{d}(G)}{t} \geq n\frac{(\bar{d}(G)-(t-1))^t}{t!},$$

From this we obtain $\bar{d}(G) \leq (s-1)^{1/t}n^{1-1/t} + t - 1$, which implies the theorem. □

In particular, we get that $ex(n, K_{2,3}) \leq cn^{3/2}$, which means that we cannot expect to have a unit-distance graph with more than $cn^{3/2}$ edges. The best known upper bound in this problem of Erdős stands at $cn^{4/3}$. Any improvement would be extremely interesting.

Exercise 2.5 *(Unit distance problem in the euclidean 3-space.) Show that for any set P of n points in $\mathbb{R}^3$ there are at most $O(n^{5/3})$ pairs of points in $\binom{P}{2}$ that have Euclidean distance exactly one.*

Remark. *The best known upper bound on the number of unit distances in the eucledian 3-space is $O(n^{3/2})$ (due to Zahl, and Kaplan, Matoušek, Safernová, Sharir). A point set with $n^{4/3}\log\log n$ pairs of unit distances was constracted by Erdős in 1960.*

Obviously, it is "easier" to avoid $K_{2,3}$s than $K_{2,2}$s: any $K_{2,2}$-free graph is $K_{2,3}$-free as well. This implies that $ex(n, K_{2,3})$ is strictly larger than $ex(n, K_{2,2})$, but according to Theorem 2.2 not by much: they are of the same order of magnitude.

The following conjecture claims that a similar phenomenon is true for arbitrary pairs of integers $s \geq t \geq 2$: the upper bound of Kővári, T. Sós and Turán is essentially tight and thus the order of magnitude of $ex(n, K_{t,s})$ essentially depends only on the smaller of the parameters.

Conjecture 4 *For any two fixed integers $s \geq t \geq 2$*

$$ex(n, K_{t,s}) = \Theta\left(n^{2-\frac{1}{t}}\right).$$

In Sections 2.2 and 2.3 we investigate further the status of this conjecture for $\min\{t,s\} \geq 3$, but first let's finish off the case when $\min\{t,s\} = 2$.

2.1.5 A lower bound for $ex(n, K_{2,s})$

In the last subsection we proved an upper bound on $ex(n, K_{t,s})$ in which the exponent of n depended only on the smaller of the two parameters t and s, namely for $t \le s$ we had

$$ex(n, K_{t,s}) \le \frac{1}{2}(s-1)^{1/t}n^{2-1/t} + \frac{1}{2}(t-1)n.$$

In particular if a $K_{t,s}$-free graph is found with $\Theta(n^{2-1/t})$ edges then $ex(n, K_{t,s'})$ is determined up to a constant factor for *every* $s' \ge s$. This is the case for $t = 2$, where we have that

$$\frac{1}{2}n^{3/2} \lessapprox ex(n, K_{2,2}) \le ex(n, K_{2,3}) \le \cdots \le ex(n, K_{2,s}) \lessapprox \frac{1}{2}\sqrt{s-1}\,n^{3/2}.$$

Füredi generalized the affine construction (Construction 2) we presented in Subsection 2.1.3 to establish the asymptotics of $ex(n, K_{2,s})$ for every s. He introduced an equivalence relation on the vertex set, modified the definition of adjacency accordingly and thus reduced the number of vertices by a constant factor while keeping the degrees intact.

Theorem 2.3 (Füredi, 1996)

$$ex(n, K_{2,s}) \gtrapprox \frac{1}{2}\sqrt{s-1}\,n^{3/2}.$$

Proof. We choose a prime p such that $p-1$ is divisible by $s-1$ (see the Remark after the proof). Recall that the multiplicative group $\mathbb{F}_p^* = \{1, 2, \ldots, p-1\}$ of the p-element field is known to be cyclic. Hence there is a unique subgroup $H < \mathbb{F}_p^*$ of order $|H| = s-1$. (If g is a generating element of $\mathbb{F}_p^*$, then $g^{\frac{p-1}{s-1}}$ generates H.)

On the set $\mathbb{F}_p^2 \setminus \{(0,0)\}$ we define the relation $\sim$, by

$$(a, b) \sim (a', b') \quad \text{if} \quad \exists h \in H \ : \ (a', b') = (ha, hb).$$

This is obviously an equivalence relation and it partitions the set $\mathbb{F}_p^2 \setminus \{(0,0)\}$ into equivalence classes, which will be the vertices of our graph G. Note that each class has $s-1$ elements. We denote by $\langle a, b\rangle$ the equivalence class containing (a, b). Then we define

$$V(G) = \{\langle a, b\rangle : (a, b) \in \mathbb{F}_p^2 \setminus \{(0,0)\}\},$$

and we have $|V(G)| = n = \frac{p^2-1}{s-1}$.

Now consider two vertices $\langle a, b\rangle \ne \langle a', b'\rangle$. We define that $\langle a, b\rangle$ and $\langle a', b'\rangle$ are adjacent in G if $aa' + bb' \in H$.

Note that for $s = 2$ the definition of G agrees with Construction 2 of the previous section. Observe the change in the definition of an edge: we not only allow $aa' + bb'$ to be equal to 1 (as in Construction 2), but also to any other h from the subgroup H. This is not solely for the purpose of achieving a larger degree but also *necessary* for having the adjacency definition consistent with the equivalence classes. Indeed, if

$\langle a,b\rangle$ and $\langle x,y\rangle$ are adjacent, and we also have $(a,b)\sim(a',b')$ and $(x,y)\sim(x',y')$, then $ax+by=h\in H, a'a^{-1}=b'b^{-1}=h'\in H, x'x^{-1}=y'y^{-1}=h''\in H$ implying $a'x'+b'y'=hh'h''\in H$, that is, the adjacency relation is well-defined.

How many neighbors does a vertex $\langle a,b\rangle$ have? Exactly the number of pairs (x,y) for which $ax+by\in H$, divided by $s-1$. One of the coordinates of (a,b) is nonzero, say $b\neq 0$. Then for any fixed $x\in\mathbb{F}_p$ and $h\in H$ there is a unique $y=\frac{h-ax}{b}$ satisfying $ax+by=h$. One can select x in p ways, h in $s-1$ ways, so the total number of solutions (x,y) of $ax+by\in H$ is $p(s-1)$. Since solutions come in equivalence classes of size $s-1$, we obtain that for $\langle a,b\rangle$ there are exactly p equivalence classes $\langle x,y\rangle$ with $ax+by\in H$. One of these might be equal to $\langle a,b\rangle$ itself, if a^2+b^2 happens to be in H. This would constitute a loop in our graph; however, in any case every vertex has degree at least $p-1$.

It remains to show that G is $K_{2,s}$-free. To this end consider two distinct vertices $\langle a,b\rangle\neq\langle a',b'\rangle$. All common neighbors $\langle x,y\rangle$ of $\langle a,b\rangle$ and $\langle a',b'\rangle$ satisfy

$$\begin{aligned} ax+by &= h,\\ a'x+b'y &= h', \end{aligned}$$

for some $h,h'\in H$. Let us now fix h and h'. We distinguish two cases.
Case 1. The matrix of the above system of linear equations is nonsingular. Then there exists a unique solution (x,y).
Case 2. The matrix is singular. Then there exists λ such that $a=\lambda a'$ and $b=\lambda b'$. Note that since $\langle a,b\rangle\neq\langle a',b'\rangle$ the pairs (a,b) and (a',b') are not equivalent, hence λ cannot be an element of H. If we multiply the second equation by λ and subtract it from the first, we get $0=h-\lambda h'$, or $\lambda=h(h')^{-1}\in H$, which is a contradiction. Hence if Case 2 applies then there are no solutions.

So for each fixed $h,h'\in H$ we have at most one solution (x,y). Solutions again come in equivalence classes of $s-1$ elements, giving that the total number of common neighbors of $\langle a,b\rangle$ and $\langle a',b'\rangle$ is at most $\frac{(s-1)^2}{s-1}=s-1$.

We thus constructed a $K_{2,s}$-free graph with

$$|E(G)|\geq\frac{1}{2}n(p-1)\geq\frac{1}{2}\sqrt{s-1}\,n^{3/2}-\frac{1}{2}n.$$

□

Remark:

1. Note that in the above construction, the number of vertices is of very special form, it is $\frac{p^2-1}{s-1}$ where p is a prime congruent to 1 modulo $s-1$. In order to provide a construction for *every* n, which then gives the correct asymptotics of $ex(n,K_{2,s})$, we need that the primes chosen in the proof of Theorem 2.3 are *frequent enough* for every s. Dirichlet's classic theorem states that in every infinite arithmetic progression, satisfying the obvious necessary condition, there are infinitely many primes. More precisely, for any pair (k,l) of relatively prime integers there are infinitely many primes of the form $ak+l$, $a\in\mathbb{Z}$. The density version of Dirichlet's Theorem also ensures that the number

of these primes up to a large integer N is exactly what one expects according to the Prime Number Theorem, namely $\approx \frac{1}{\phi(k)}\frac{N}{\log N}$, where $\phi(k)$ is Euler's function, the number of positive integers up to k that are relatively prime to k. However, for us this is still not enough: we need that in *every small enough interval there is such a prime.* Such type of statements are the topic of intensive research in number theory. In particular by a result of Huxley and Iwaniec (1975) for every sufficiently large n there exists a prime $p \equiv 1 \pmod{s}$ such that $\sqrt{sn} - n^{1/3} < p < \sqrt{sn}$.

2. Selecting $s-1 = p-1$ would give us that *any* nonzero multiple of (a,b) is equivalent with it, that is $V(G)$ is the 1-dimensional projective space (line). The defined graph is the clique, which is then $K_{2,p}$-free indeed, since the projective line contains only $p+1$ points.

3. From the analysis of the two cases in the proof of Theorem 2.3 it could seem at first sight that, since Case 2 leads to a contradiction, every two vertices have exactly $s-1$ common neighbors. This is false however, and not only because of the presence of the loops, but also because Case 2 leads to a contradiction only if there was a common neighbor in the first place. It could happen, though rarely, that two vertices do not have a common neighbor at all. This occurs only if $(a,b) = (\lambda a, \lambda b)$ with some $\lambda \notin H$.

2.2 Forbidding $K_{3,s}$

The value of $ex(n, K_{t,s})$ for $t = 2$ is fully settled now up to lower order terms. Before going further to $t \geq 3$ we repeat our randomized approach for obtaining a dense H-free graph for arbitrary H. In particular, for $H = K_{t,s}$ we will check how the random attempt relates to the Kővári-Sós-Turán upper bound.

2.2.1 A quick detour to random graphs

The plan is again to find a probability p as large as possible such that in the random graph $G(n,p)$ we have

$$\mathbf{E}[e(G(n,p)) - \#\{\text{copies of } H \text{ in } G(n,p)\}] \geq \frac{1}{2}\binom{n}{2}p.$$

This definitely holds if

$$\frac{1}{2}\binom{n}{2}p \geq n^{n(H)}p^{e(H)} \geq \mathbf{E}[\#\{\text{copies of } H \text{ in } G(n,p)\}],$$

or if $p = O\left(n^{-\frac{n(H)-2}{e(H)-1}}\right)$.

Hence, we know that with some $p = cn^{-\frac{n(H)-2}{e(H)-1}}$ there exists a graph G such that after deleting one edge from every copy of H there are still $\frac{1}{2}\binom{n}{2}p = \Theta\left(n^{2-\frac{n(H)-2}{e(H)-1}}\right)$ edges left. We proved the following.