The Regularity Lemma and Applications

Ryan Martin

rymartin@andrew.cmu.edu

Zeev Nehari Visiting Assistant Professor

Mathematics Department

Carnegie Mellon University

Arithmetic progressions

An arithmetic sequence is a set of integers of the form

$$\{a_0 + di : i = 0, 1, 2, 3 \ldots\}$$

where a_0 and d are positive integers.

Arithmetic progressions

An arithmetic progression is a set of integers of the form

Arithmetic progressions

An arithmetic progression is a set of integers of the form

$$\{a_0 + di : i = 0, 1, \ldots, k-1\}$$

where a_0 and d are positive integers.

Arithmetic progressions

An arithmetic progression of length k is a set of integers of the form

$$\{a_0 + di : i = 0, 1, \ldots, k-1\}$$

where a_0 and d are positive integers.

Van der Waerden's theorem

The van der Waerden theorem has a number of equivalent forms. One is:

Van der Waerden's theorem

The van der Waerden theorem has a number of equivalent forms. One is:

THEOREM [van der Waerden, 1927]
Let k and r be positive integers.

Van der Waerden's theorem

The van der Waerden theorem has a number of equivalent forms. One is:

THEOREM [van der Waerden, 1927]
Let k and r be positive integers. There exists a constant $n(k, r)$ (the van der Waerden number)

Van der Waerden's theorem

The van der Waerden theorem has a number of equivalent forms. One is:

THEOREM [van der Waerden, 1927]
Let k and r be positive integers. There exists a constant $n(k,r)$ such that if $n_0 \geq n(k,r)$ and $\{1,2,\ldots,n_0\} \subset C_1 \cup C_2 \cup \cdots \cup C_r$,

Van der Waerden's theorem

The van der Waerden theorem has a number of equivalent forms. One is:

THEOREM [van der Waerden, 1927]
Let k and r be positive integers. There exists a constant $n(k,r)$ such that if $n_0 \geq n(k,r)$ and $\{1,2,\ldots,n_0\} \subset C_1 \cup C_2 \cup \cdots \cup C_r$,
then some set C_i contains an arithmetic progression of length k.

The van der Waerden number

The quantity $n(k, r)$ (the van der Waerden number) is well-studied.

The van der Waerden number

The quantity $n(k, r)$ (the van der Waerden number) is well-studied.

Shelah (1988) proved that the van der Waerden numbers are primitive recursive.

The van der Waerden number

The quantity $n(k,r)$ (the van der Waerden number) is well-studied.

Gowers (1998) announced that

$$n(r,k) \le e^{e^{(1/r)^{e^{e^{k+110}}}}}$$

The van der Waerden number

The quantity $n(k, r)$ (the van der Waerden number) is well-studied.

Gowers (1998) announced that

$$n(r, k) \leq e^{e^{(1/r)e^{e^{k+110}}}}$$

But it has yet to be published.

The van der Waerden number

The quantity $n(k, r)$ (the van der Waerden number) is well-studied.

Gowers (1998) announced that

$$n(r, k) \leq e^{e^{(1/r)e^{e^{k+110}}}}$$

An argument using a probabilistic technique (the Lovász Local Lemma) gives that

$$n(k, r) > \left(\frac{r^k}{erk}\right)(1 + o(1)).$$

Plain language

Our version of van der Waerden says:

> If we color the first n_0 positive integers with r colors, we get a monochromatic arithmetic progression of length k, as long as n_0 is large enough.

Plain language

Our version of van der Waerden says:

> If we color the first n_0 positive integers with r colors, we get a monochromatic arithmetic progression of length k, as long as n_0 is large enough.

Which color?

The average

If we color a set with r colors, then one of those colors will contain $1/r$ of the set.

The average

If we color a set with r colors, then one of those colors will contain $1/r$ of the set.

That set is a very likely candidate. Is that the one?

Dense subsets

More generally, we can ask the following:

> If we have a subset of $\{1, 2, \ldots, n\}$ with positive density, does it have an arithmetic progression of length k?

Dense subsets

Let $A_n \subset \{1, 2, \ldots, n\}$. The family A_n is said to have positive upper density if there exists an ϵ such that $|A_n| > \epsilon n$.

Arithmetic progressions and density

Endre Szemerédi proved that positive upper density is sufficient for the existence of a k-term arithmetic progression.

Arithmetic progressions and density

Endre Szemerédi proved that positive upper density is sufficient for the existence of a k-term arithmetic progression.

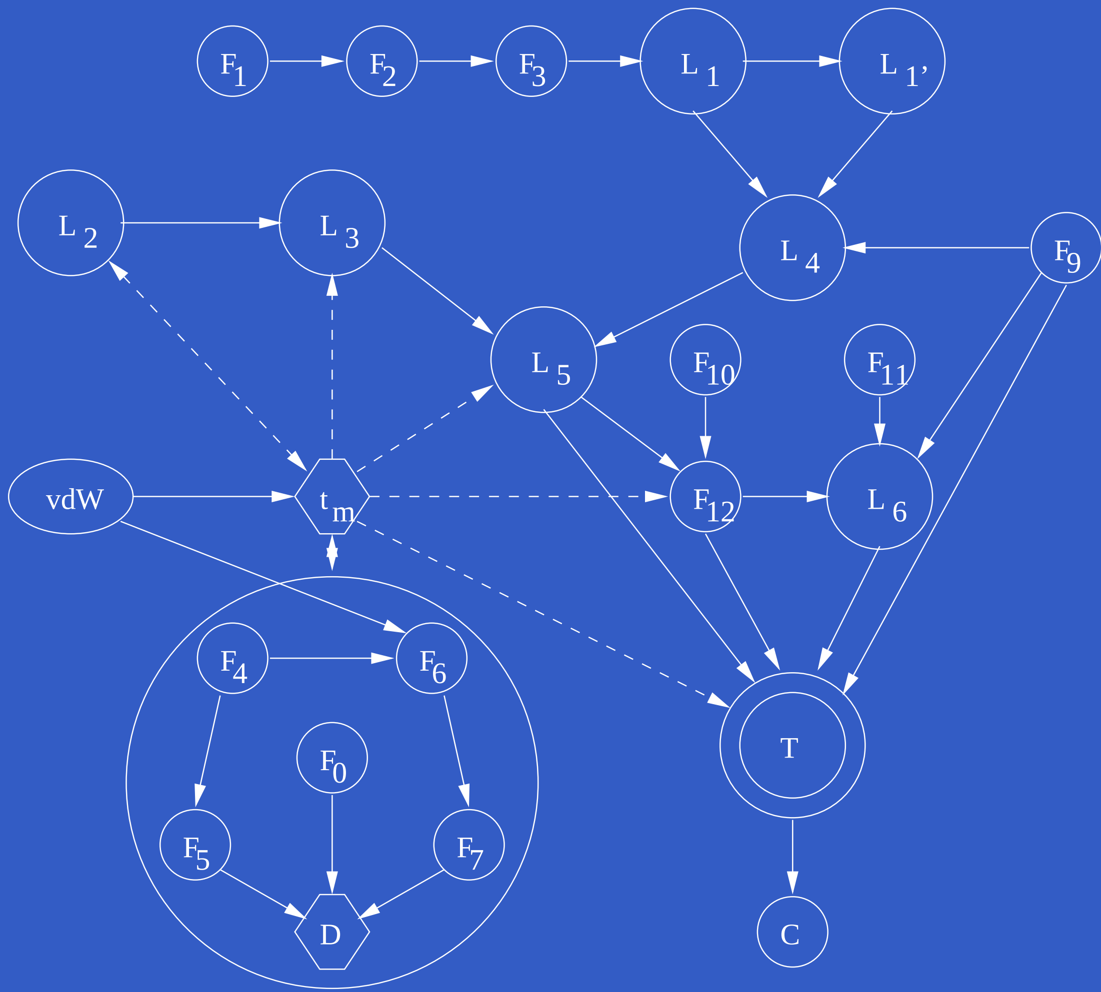
THEOREM [Szemerédi, 1975]
For every integer $k > 2$ and every $\epsilon > 0$, there exists a threshold $n_0 = n_0(k, \epsilon)$ such that if, for some $n \geq n_0$, $A \subset \{1, 2, \ldots, n\}$ and $|A| > \epsilon n$, then A must contain an arithmetic progression of k terms.

The proof of Szemerédi's theorem

The proof itself is quite long, complicated and ingenious.

For example, consider a diagram of the structure of the proof.

The proof of Szemerédi’s theorem



The Regularity Lemma

Buried in the proof of Szemerédi's theorem is a primitive version of the Regularity Lemma.

But the Regularity Lemma is not a statement about number theory, it's a statement about graphs.

Graph definitions

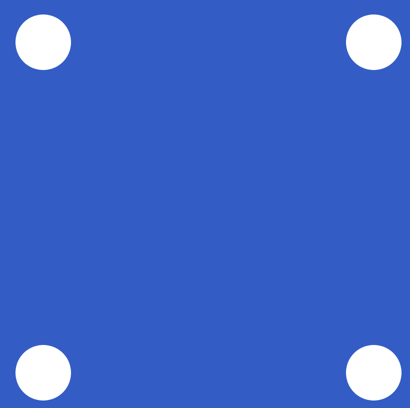
Let's define some basic graph terms.

We have a set of

Graph definitions

Let's define some basic graph terms.

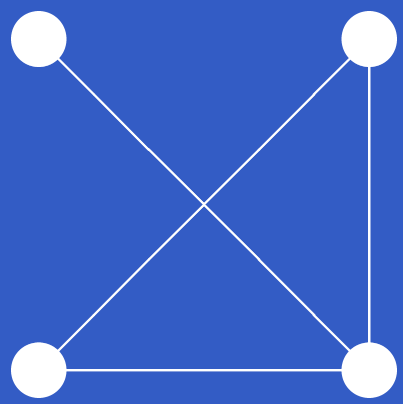
We have a set of vertices,



Graph definitions

Let’s define some basic graph terms.

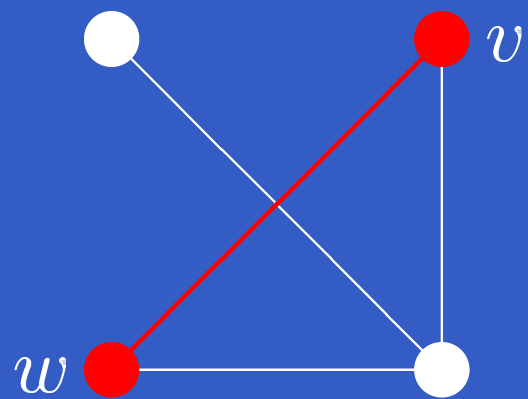
We have a set of vertices, connected by edges



Graph definitions

Let's define some basic graph terms.

We have a set of vertices, connected by edges

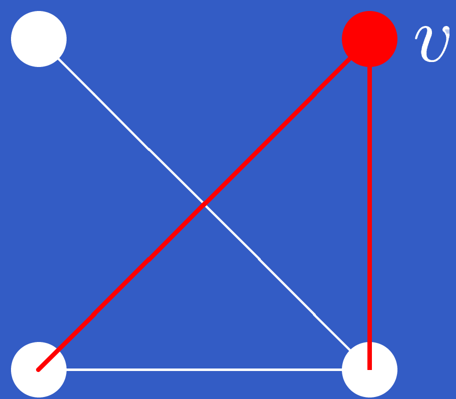


v is said to be adjacent to w if there is an edge between them.

Graph definitions

Let's define some basic graph terms.

We have a set of vertices, connected by edges



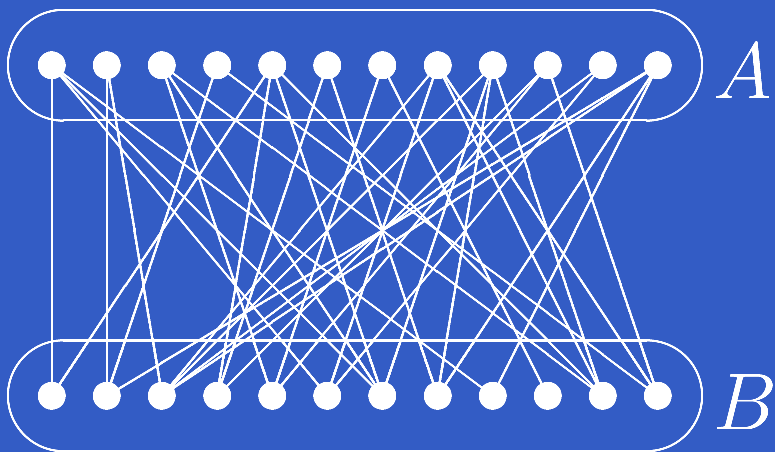
v is said to be adjacent to w if there is an edge between them. The degree of v, $\deg(v)$, is the number of edges touching v.

Bipartite

A graph G is bipartite if we can partition the vertex set into A and B so that all edges are **between** A and B:

Bipartite

A graph G is bipartite if we can partition the vertex set into A and B so that all edges are **between** A and B:



We denote such a graph as $G = (A, B; E)$.

Random bipartite graphs

Let $(A, B; E)$ be a bipartite graph.

$$|A| = |B| = n$$

Random bipartite graphs

Let $(A, B; E)$ be a bipartite graph.

$$|A| = |B| = n$$

The number of edges is: $e(A, B)$.

Random bipartite graphs

Let $(A, B; E)$ be a bipartite graph.

$$|A| = |B| = n$$

The number of edges is: $e(A, B)$.
Define the density to be:

$$d(A, B)$$

Random bipartite graphs

Let $(A, B; E)$ be a bipartite graph.

$$|A| = |B| = n$$

The number of edges is: $e(A, B)$.
Define the density to be:

$$d(A, B) = \frac{e(A, B)}{n^2}$$

Binomial model

For each pair of vertices: $a \in A$, $b \in B$,

Binomial model

For each pair of vertices: $a \in A$, $b \in B$, let ab be an edge in the graph with probability

$$d \in (0, 1)$$

independently for each such pair.

Binomial model

For each pair of vertices: $a \in A$, $b \in B$, let ab be an edge in the graph with probability

$$d \in (0, 1)$$

independently for each such pair.

I.e., take an $a \in A$ and $b \in B$ and flip a biased coin to see whether or not ab is an edge.

Binomial model

For each pair of vertices: $a \in A$, $b \in B$, let ab be an edge in the graph with probability

$$d \in (0, 1)$$

independently for each such pair.

The average density is

$$E[d(A, B)] = \frac{dn^2}{n^2} = d.$$

Subgraphs

Choose $A' \subseteq A$ and $B' \subseteq B$.
What is the probability that

$$d(A', B') \in (d - \epsilon, d + \epsilon)?$$

Subgraphs

Choose $A' \subseteq A$ and $B' \subseteq B$.
What is the probability that

$$d(A', B') \in (d - \epsilon, d + \epsilon)?$$

Easy to compute because $d(A', B')$ is a binomial random variable.

Subgraphs

Choose $A' \subseteq A$ and $B' \subseteq B$.
What is the probability that

$$d(A', B') \in (d - \epsilon, d + \epsilon)?$$

Easy to compute because $d(A', B')$ is a binomial random variable.

I.e., it is the average of coin flips.

Chernoff bound

The Chernoff bound gives the probability that the density between these subsets differs from d by more than ϵ.

Chernoff bound

The Chernoff bound gives the probability that the density between these subsets differs from d by more than ϵ.

$$\Pr\{|d(A', B') - d| \geq \epsilon\}$$

Chernoff bound

The Chernoff bound gives the probability that the density between these subsets differs from d by more than ϵ.

$$\Pr\left\{|d(A', B') - d| \geq \epsilon\right\} \leq 2\exp\left(-2\epsilon^2|A'||B'|\right)$$

Chernoff bound

In fact, if $|A'| > \epsilon n$ and $|B'| > \epsilon n$, then we have:

$$\Pr\left\{|d(A', B') - d| \geq \epsilon\right\} \leq 2\exp\left(-2\epsilon^4 n^2\right)$$

Chernoff bound

In fact, if $|A'| > \epsilon n$ and $|B'| > \epsilon n$, then we have:

$$\Pr\{|d(A', B') - d| \geq \epsilon\} \leq 2\exp\left(-2\epsilon^4 n^2\right)$$

This is ... tiny ...

Applies simultaneously to all such sets

In fact, the probability that ALL pairs of sets, each with size $> \epsilon n$, have density in

$$(d-\epsilon, d+\epsilon)$$

Applies simultaneously to all such sets

In fact, the probability that ALL pairs of sets, each with size $> \epsilon n$, have density in

$$(d - \epsilon, d + \epsilon)$$

approaches 1
as $n \to \infty$.

Defining regularity

A pair (A, B) is ϵ-regular with density d if

- $d = d(A, B)$

Defining regularity

A pair (A, B) is ϵ-regular with density d if

- $d = d(A, B)$
- For all $A' \subseteq A$ and $B' \subseteq B$ with

Defining regularity

A pair (A, B) is ϵ-regular with density d if

- $d = d(A, B)$
- For all $A' \subseteq A$ and $B' \subseteq B$ with
  - $|A'| > \epsilon|A|$ and

Defining regularity

A pair (A, B) is ϵ-regular with density d if

- $d = d(A, B)$
- For all $A' \subseteq A$ and $B' \subseteq B$ with
  - $|A'| > \epsilon |A|$ and
  - $|B'| > \epsilon |B|$,

Defining regularity

A pair (A, B) is ϵ-regular with density d if

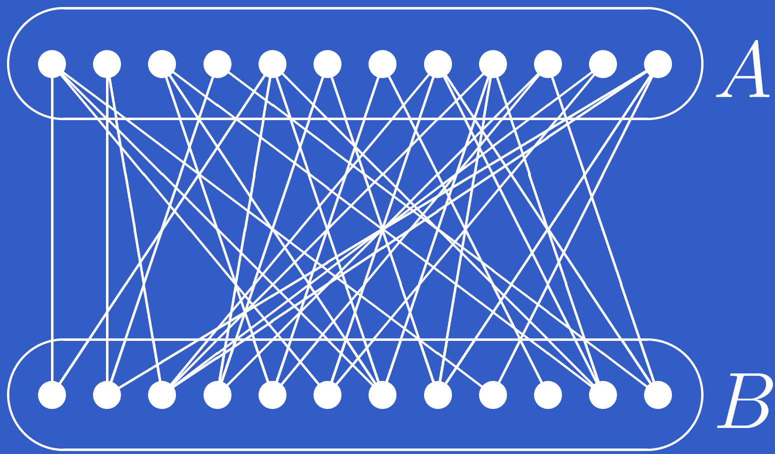
- $d = d(A, B)$
- For all $A' \subseteq A$ and $B' \subseteq B$ with
  - $|A'| > \epsilon|A|$ and
  - $|B'| > \epsilon|B|$,

  we have $|d(A', B') - d| < \epsilon$.

So, ϵ-regular pairs mimic random pairs.

Random graphs

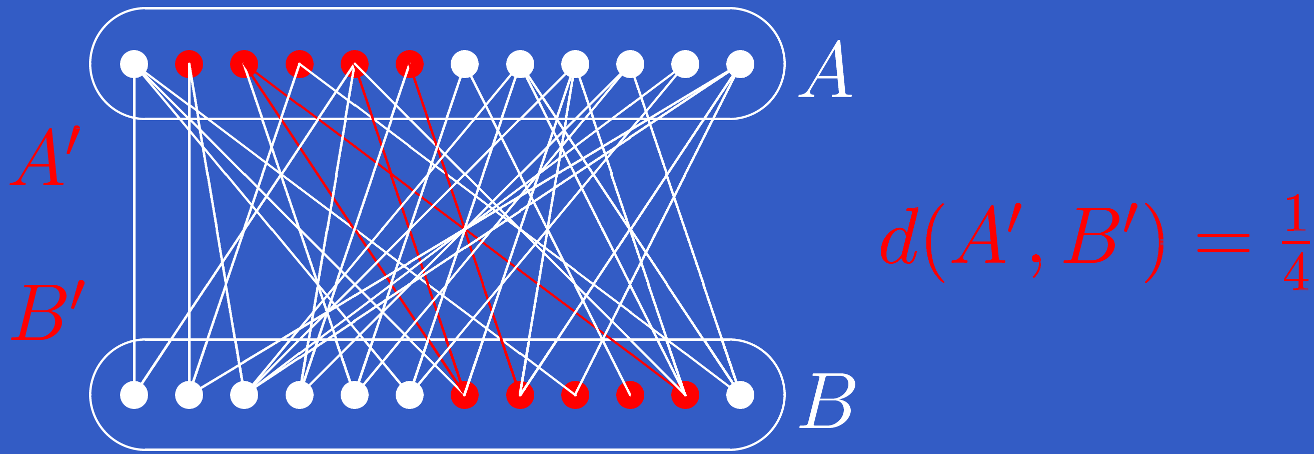
ϵ-regular pairs are tied inexorably to random pairs.



$$d(A, B) = \frac{1}{4}$$

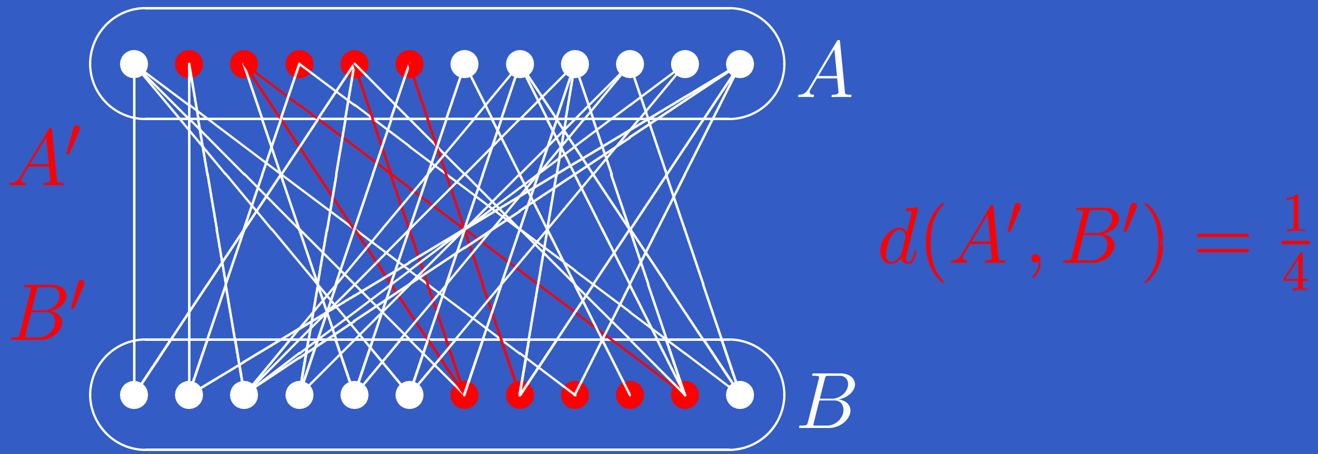
Random graphs

ϵ-regular pairs are tied inexorably to random pairs.



Random graphs

ϵ-regular pairs are tied inexorably to random pairs.



In fact, an ϵ-regular pair has a surprising number of properties that random pairs have.

The Regularity Lemma

The Regularity Lemma ensures that, for any G with enough vertices, $V(G)$ has a partition with:

- a constant number of pieces,

The Regularity Lemma

The Regularity Lemma ensures that, for any G with enough vertices, $V(G)$ has a partition with:

- a constant number of pieces,
- all but one of equal size,

The Regularity Lemma

The Regularity Lemma ensures that, for any G with enough vertices, $V(G)$ has a partition with:

- a constant number of pieces,
- all but one of equal size,
- a dense subgraph G' has **all** pairs ϵ-regular, density either

The Regularity Lemma

The Regularity Lemma ensures that, for any G with enough vertices, $V(G)$ has a partition with:

- a constant number of pieces,
- all but one of equal size,
- a dense subgraph G' has **all** pairs ϵ-regular, density either
    - zero, or

The Regularity Lemma

The Regularity Lemma ensures that, for any G with enough vertices, $V(G)$ has a partition with:

- a constant number of pieces,
- all but one of equal size,
- a dense subgraph G' has **all** pairs ϵ-regular, density either
  - zero, or
  - bounded below by a constant.

The Regularity Lemma

The Regularity Lemma ensures that, for any G with enough vertices, $V(G)$ has a partition with:

- a constant number of pieces,
- all but one of equal size,
- a dense subgraph G' has **all** pairs ϵ-regular, density either
  - zero, or
  - bounded below by a constant.

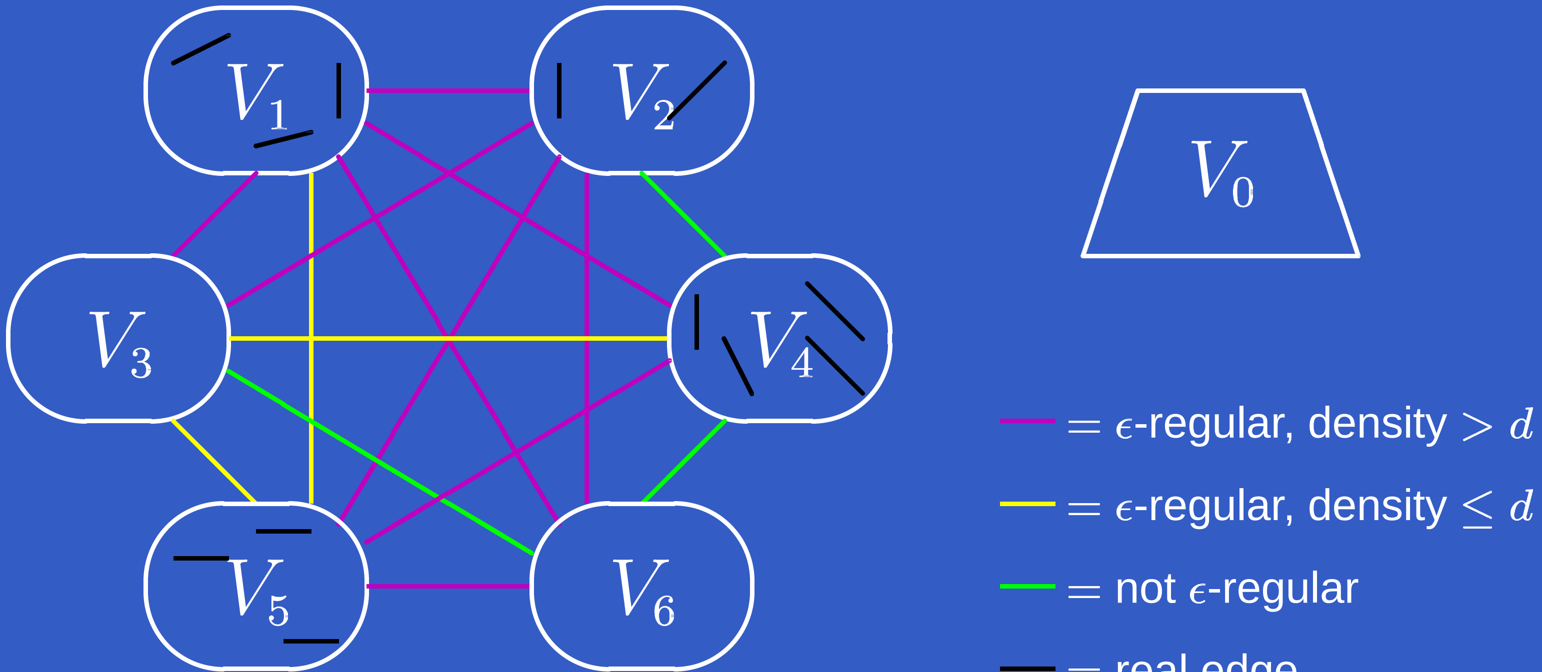
It's easier to see in a picture.

What this does

Before Regularity:

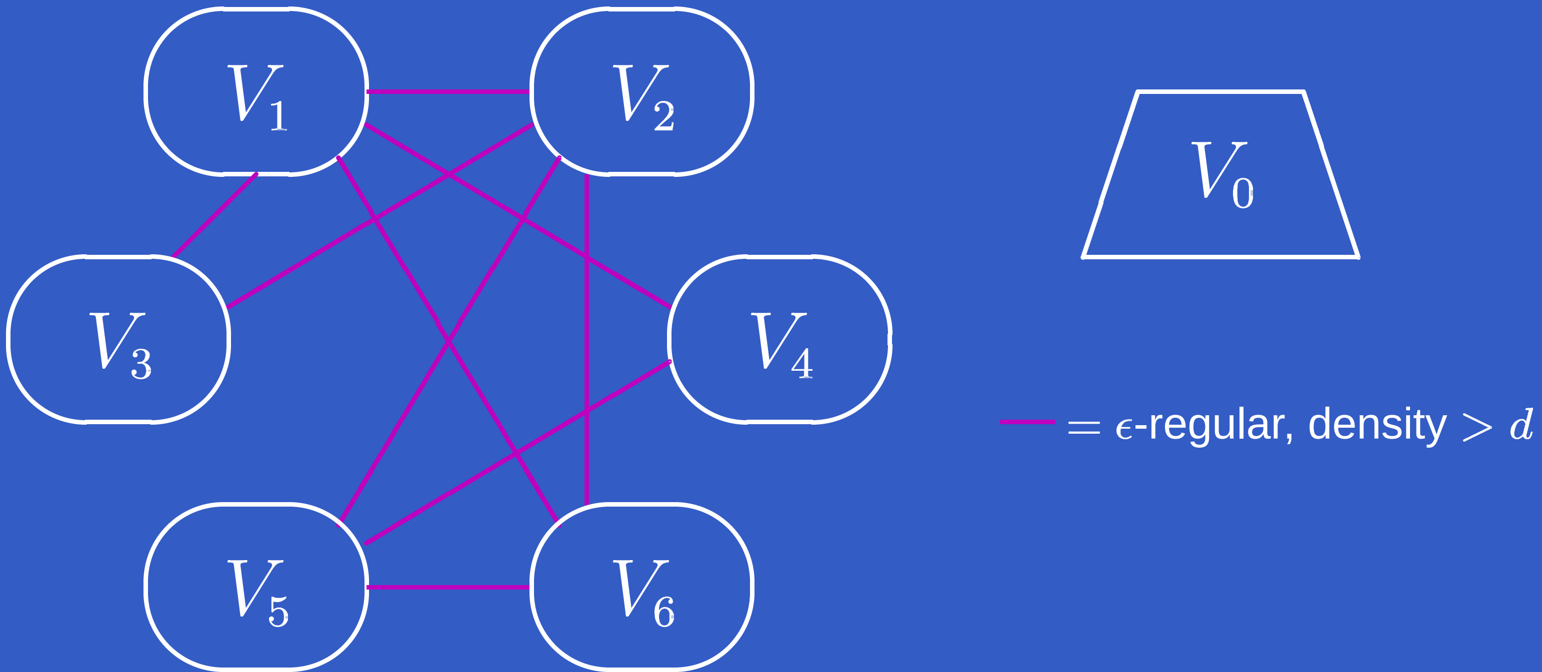
What this does

After Regularity, the graph G:



What this does

After Regularity, the graph G':



G' is very applicable.

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

- $k \leq M$

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

- $k \leq M$
- $|V_0| \leq \epsilon |V|$ and $|V_1| = \cdots = |V_k|$

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

- $k \leq M$
- $|V_0| \leq \epsilon |V|$ and $|V_1| = \cdots = |V_k|$
- $\exists$ subgraph $G' \subset G$

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

- $k \leq M$
- $|V_0| \leq \epsilon |V|$ and $|V_1| = \cdots = |V_k|$
- $\exists$ subgraph $G' \subset G$
- $\deg_{G'}(v) > \deg_G(v) - (d + \epsilon)|V|$, $\forall v \in V$

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

- $k \leq M$
- $|V_0| \leq \epsilon |V|$ and $|V_1| = \cdots = |V_k|$
- $\exists$ subgraph $G' \subset G$
- $\deg_{G'}(v) > \deg_G(v) - (d + \epsilon)|V|$, $\forall v \in V$
- no G' edges inside a V_i, $i = 1, \ldots, k$

Formal definition

THEOREM [Szemerédi, 1978] (Degree Form)

$\forall \epsilon > 0$, $\exists M = M(\epsilon)$ such that,
if $G = (V, E)$ and $d \in [0, 1]$, then
$V = V_0 + V_1 + \cdots + V_k$ (a partition) with

- $k \leq M$
- $|V_0| \leq \epsilon |V|$ and $|V_1| = \cdots = |V_k|$
- $\exists$ subgraph $G' \subset G$
- $\deg_{G'}(v) > \deg_G(v) - (d + \epsilon)|V|$, $\forall v \in V$
- no G' edges inside a V_i, $i = 1, \ldots, k$
- **all** pairs $G'(V_i, V_j)$, ($i, j \geq 1$) are ϵ-regular, each with a density **either** 0 **or** $> d$.

Geometry applications

One application is to look at distances in the plane.
Let us be given n points in the plane.

Geometry applications

One application is to look at distances in the plane.
Let us be given n points in the plane.

There are

$$\binom{n}{2} = \frac{n(n-1)}{2} \approx \frac{n^2}{2}$$

pairs of points.

Getting large distances

Let n be a large number divisible by 2.
Consider the following subset of $\mathbb{Z}^2$:

Getting large distances

Let n be a large number divisible by 2.
Consider the following subset of $\mathbb{Z}^2$:

$$P = \left\{ (in^2, j) \mid i \in \{0, 1\}, j \in \{1, 2, \ldots, n/2\} \right\}$$

Getting large distances

$$P = \left\{ (in^2, j) \mid i \in \{0,1\}, j \in \{1,2,\ldots,n/2\} \right\}$$

Consider two points, (in^2, j) and $(i'n^2, j')$ and compute their distance:

$$d = \sqrt{(i-i')^2 n^4 + (j-j')^2}$$

Getting large distances

$$P = \left\{(in^2, j) \mid i \in \{0,1\}, j \in \{1,2,\ldots,n/2\}\right\}$$

Consider two points, (in^2, j) and $(i'n^2, j')$ and compute their distance:

$$|i-i'|n^2 \le d = \sqrt{(i-i')^2 n^4 + (j-j')^2} \le |i-i'|n^2+1$$

as long as $i \ne i'$.

Getting large distances

$$P = \left\{(in^2, j) \mid i \in \{0,1\}, j \in \{1, 2, \ldots, n/2\}\right\}$$

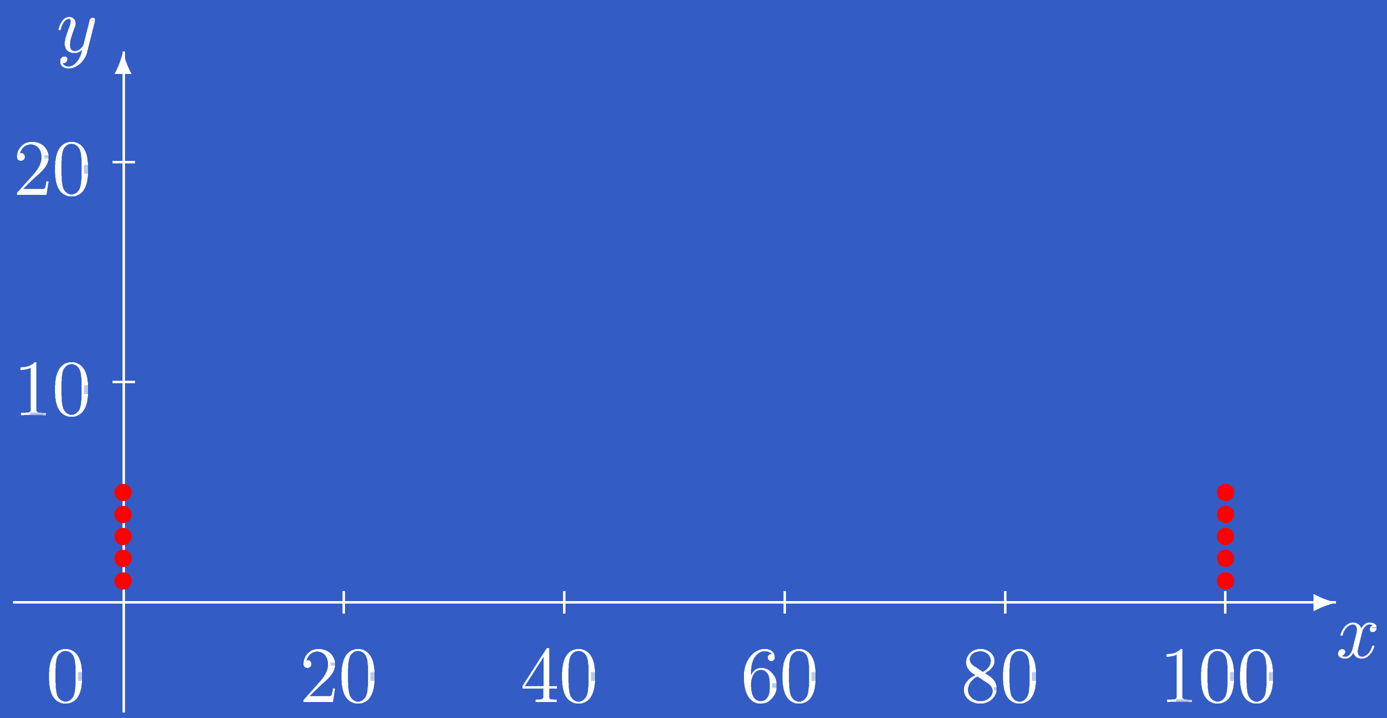
Consider two points, (in^2, j) and $(i'n^2, j')$ and compute their distance:

$$|i-i'|n^2 \leq d = \sqrt{(i-i')^2 n^4 + (j-j')^2} \leq |i-i'|n^2 + 1$$

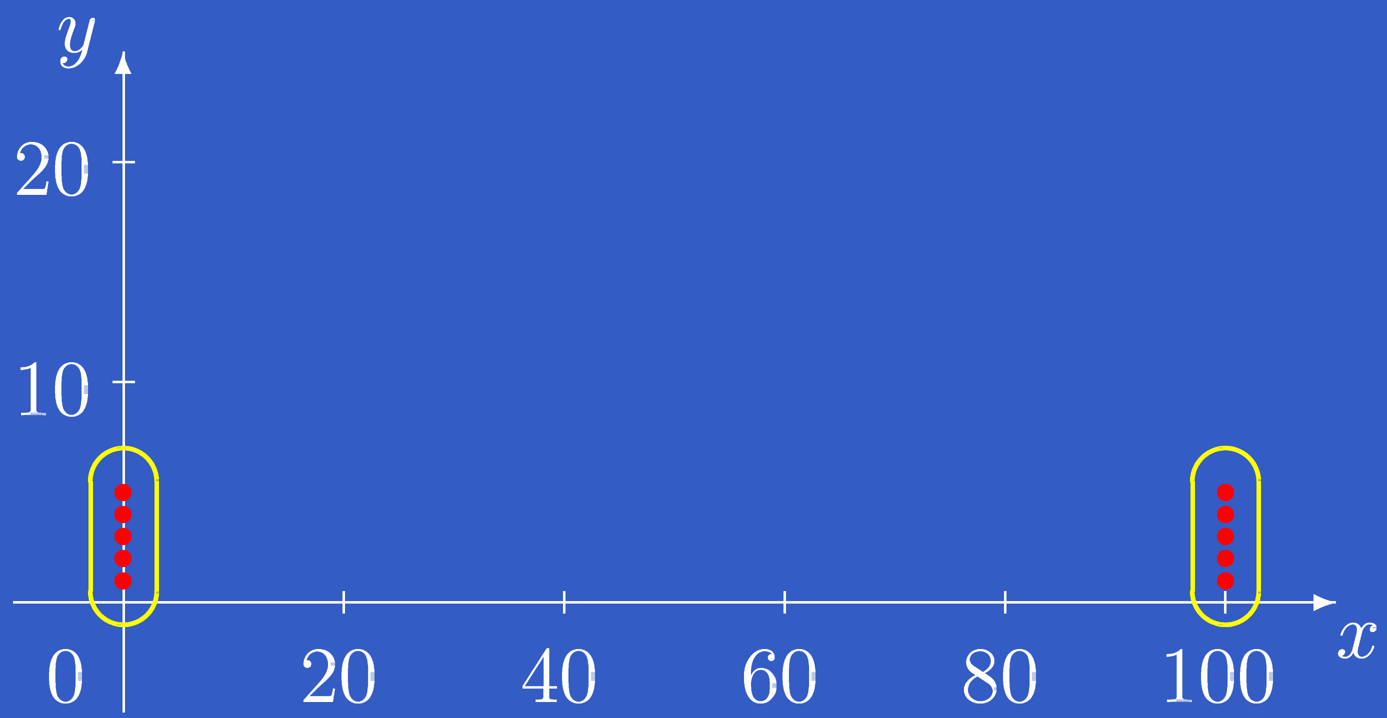
as long as $i \neq i'$.

Even if $i = i'$, the distance is at least 1.

A picture of this family



A picture of this family



Distances are restricted

Thus, P has at least

$$\left(1-\frac{1}{2}\right)\frac{n^2}{2}=\frac{n^2}{4}$$

pairs with distances in

$$[n^2, n^2+1].$$

Best possible, asymptotically

THEOREM [Erdős-Makai-Pach-Spencer, 1991]
Given $\epsilon > 0$, there is a c and a positive integer n_0 satisfying the following condition:

Best possible, asymptotically

THEOREM [Erdős-Makai-Pach-Spencer, 1991]

Given $\epsilon > 0$, there is a c and a positive integer n_0 satisfying the following condition:

For any set $\{p_1, p_2, \ldots, p_n\} \subseteq \mathbb{R}^2$ ($n \geq n_0$) with minimal distance at least 1 and for any real t , the number of pairs $\{p_i, p_j\}$ whose distance is

$$\|p_i - p_j\| \in [t, t + c\sqrt{n}]$$

Best possible, asymptotically

THEOREM [Erdős-Makai-Pach-Spencer, 1991]
Given $\epsilon > 0$, there is a c and a positive integer n_0 satisfying the following condition:
For any set $\{p_1, p_2, \ldots, p_n\} \subseteq \mathbb{R}^2$ ($n \geq n_0$) with minimal distance at least 1 and for any real t , the number of pairs $\{p_i, p_j\}$ whose distance is

$$||p_i - p_j|| \in [t, t + c\sqrt{n}]$$

is at most $n^2/4$

Best possible, asymptotically

THEOREM [Erdős-Makai-Pach-Spencer, 1991]
Given $\epsilon > 0$, there is a c and a positive integer n_0 satisfying the following condition:
For any set $\{p_1, p_2, \ldots, p_n\} \subseteq \mathbb{R}^2$ $(n \geq n_0)$ with minimal distance at least 1 and for any real t , the number of pairs $\{p_i, p_j\}$ whose distance is

$$||p_i - p_j|| \in [t, t + c\sqrt{n}]$$

is at most $n^2/4$

There is a generalization.

Best possible, asymptotically

THEOREM [Erdős-Makai-Pach, 1993]
Given $k \in \mathbb{Z}^+$ and $\epsilon > 0$, there is a c and a positive integer n_0 satisfying the following condition:
For any set $\{p_1, p_2, \ldots, p_n\} \subseteq \mathbb{R}^2$ $(n \geq n_0)$ with minimal distance at least 1 and for any reals $t_1, t_2, \ldots, t_k$, the number of pairs $\{p_i, p_j\}$ whose distance is

$$\|p_i - p_j\| \in \bigcup_{\ell=1}^{k} [t_\ell, t_\ell + c\sqrt{n}]$$

is at most $(1 - 1/(k+1) + \epsilon)\,(n^2/2)$

Graph theory applications

Most applications are in graph theory.

Graph theory applications

Most applications are, of course, in graph theory.

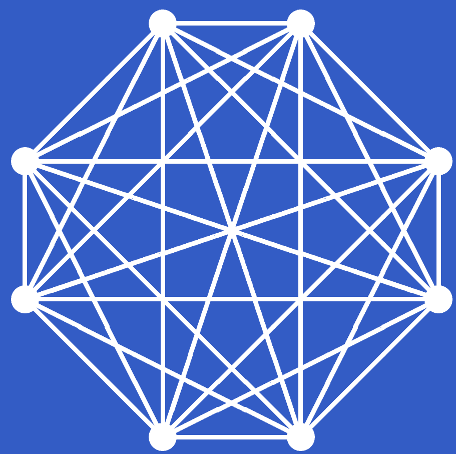
Graph theory applications

Most applications are, of course, in graph theory.

Define a complete graph on p vertices, K_p, to be the graph with p vertices where there is an edge between each pair of vertices.

Graph theory applications

Define a complete graph on p vertices, K_p, to be the graph with p vertices where there is an edge between each pair of vertices.



K_8

Theorem of Turán

Turán's theorem is a classical result. Here is a special case.

Theorem of Turán

Turán's theorem is a classical result. Here is a special case.

THEOREM [Turán, 1941]

- Let K_3 be the complete graph on 3 vertices.
- Let G_n, on n vertices, contain no K_3 as a subgraph.

Theorem of Turán

Turán's theorem is a classical result. Here is a special case.

THEOREM [Turán, 1941]

- Let K_3 be the complete graph on 3 vertices.
- Let G_n, on n vertices, contain no K_3 as a subgraph. Then,

$$|E(G_n)| \leq \frac{n^2}{4}$$

Consequence of Regularity Lemma

THEOREM

Let $\beta > 0$ be given and write $\epsilon = (\beta/6)^3$.

Consequence of Regularity Lemma

THEOREM

Let $\beta > 0$ be given and write $\epsilon = (\beta/6)^3$.
There is an $M(\epsilon)$ such that if n is large enough and a graph G_n has

$$|E(G_n)| > \left(\frac{1}{2} + \beta\right) \frac{n^2}{2},$$

Consequence of Regularity Lemma

THEOREM

Let $\beta > 0$ be given and write $\epsilon = (\beta/6)^3$.
There is an $M(\epsilon)$ such that if n is large enough and a graph G_n has

$$|E(G_n)| > \left(\frac{1}{2} + \beta\right)\frac{n^2}{2},$$

then G_n contains at least $\left(\frac{\epsilon}{M(\epsilon)}n\right)^3$ copies of K_3.

Zero to many

There are graphs G_n such that

$$|E(G_n)| = \left(\frac{1}{2}\right)\frac{n^2}{2} - o(n^2)$$

but G_n contains no K_3.

Zero to many

There are graphs G_n such that

$$|E(G_n)| = \left(\frac{1}{2}\right)\frac{n^2}{2} - o(n^2)$$

but G_n contains no K_3.

So, just a few edges takes us from zero copies of K_3 to a large number:

$$\left(\frac{\epsilon n}{M(\epsilon)}\right)^3$$

Zero to many

There are graphs G_n such that

$$|E(G_n)| = \left(1 - \frac{1}{p-1}\right)\frac{n^2}{2} - o(n^2)$$

but G_n contains no K_p.

Just a few edges takes us from zero copies of K_p to a large number:

$$\left(\frac{\epsilon n}{M(\epsilon)}\right)^p$$

Zero to many

There are graphs G_n such that

$$|E(G_n)| = \left(1 - \frac{1}{p-1}\right)\frac{n^2}{2} - o(n^2)$$

but G_n contains no K_p.

Just a few edges takes us from zero copies of K_p to a large number:

$$\left(\frac{\epsilon n}{M(\epsilon)}\right)^p$$

Note that $M(\epsilon)$ is a constant.

Size of $M(\epsilon)$

Question: How big is this $M(\epsilon)$?

Size of $M(\epsilon)$

Question: How big is this $M(\epsilon)$?

Answer:

Size of $M(\epsilon)$

Question: How big is this $M(\epsilon)$?

Answer: Huge.

OK, smart guy, how huge?

OK, smart guy, how huge?

Applying the Regularity Lemma to a graph requires it to have many vertices.

The proof of the Regularity Lemma gives a tower function for $M(\epsilon)$.

OK, smart guy, how huge?

Applying the Regularity Lemma to a graph requires it to have many vertices.

The proof of the Regularity Lemma gives a tower function for $M(\epsilon)$.

What the proof gives is:

$$4^{4^{4^{\cdot^{\cdot^{\cdot^{4^{4^{(1/\epsilon)}}}}}}}} \leq M(\epsilon)$$

OK, smart guy, how huge?

Applying the Regularity Lemma to a graph requires it to have many vertices.

The proof of the Regularity Lemma gives a tower function for $M(\epsilon)$.

What the proof gives is:

$$4^{4^{4^{\cdot^{\cdot^{\cdot^{4^{4^{(1/\epsilon)}}}}}}}} \le M(\epsilon) \le 5^{5^{5^{\cdot^{\cdot^{\cdot^{5^{5^{(1/\epsilon)}}}}}}}}$$

OK, smart guy, how huge?

Applying the Regularity Lemma to a graph requires it to have many vertices.

The proof of the Regularity Lemma gives a tower function for $M(\epsilon)$.

What the proof gives is:

$$\left\lceil \frac{10}{\epsilon^5} \right\rceil \left\{ 4^{4^{4^{\cdot^{\cdot^{\cdot^{4^{4^{(1/\epsilon)}}}}}}}} \leq M(\epsilon) \leq 5^{5^{5^{\cdot^{\cdot^{\cdot^{5^{5^{(1/\epsilon)}}}}}}}} \right\} \left\lceil \frac{10}{\epsilon^5} \right\rceil$$

Does it need to be so big?

Some kind of tower function is necessary.

THEOREM [Gowers, 1997] For any $\epsilon > 0$, there is a graph so that any application of the Regularity Lemma requires that the number of clusters is at least a number which is a tower of twos of height proportional to $\log(1/\epsilon)$.

Results are still satisfying

Despite the size requirement, there are still pretty results.

- A diameter 2 graph is one that has a path of length 2 between any pair of nonadjacent vertices.

Results are still satisfying

Despite the size requirement, there are still pretty results.

- A diameter 2 graph is one that has a path of length 2 between any pair of nonadjacent vertices.
- A minimal diameter 2 graph is a diameter 2 graph but $G \setminus e$ is **not**, $\forall e \in E(G)$.

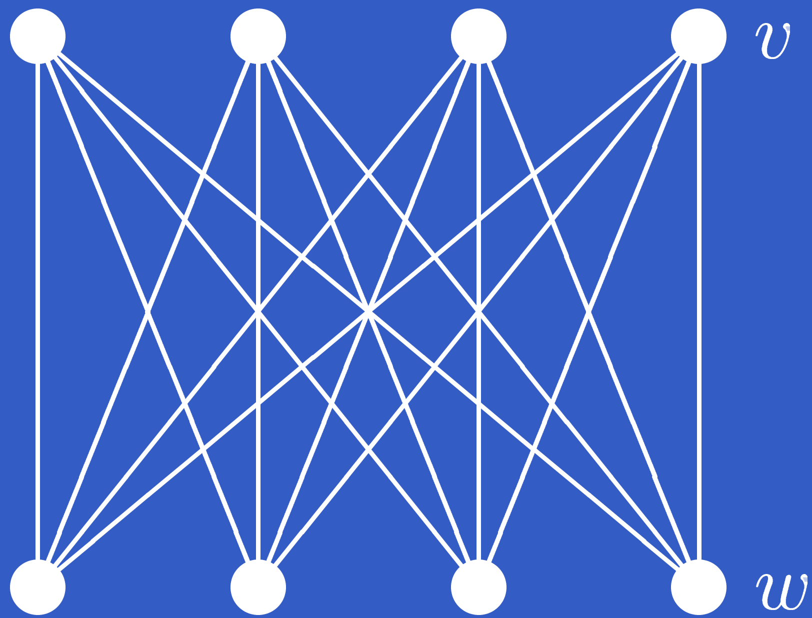
Results are still satisfying

Despite the size requirement, there are still pretty results.

- A diameter 2 graph is one that has a path of length 2 between any pair of nonadjacent vertices.
- A minimal diameter 2 graph is a diameter 2 graph but $G \setminus e$ is **not**, $\forall e \in E(G)$.
- Any complete bipartite graph is minimal diameter 2.

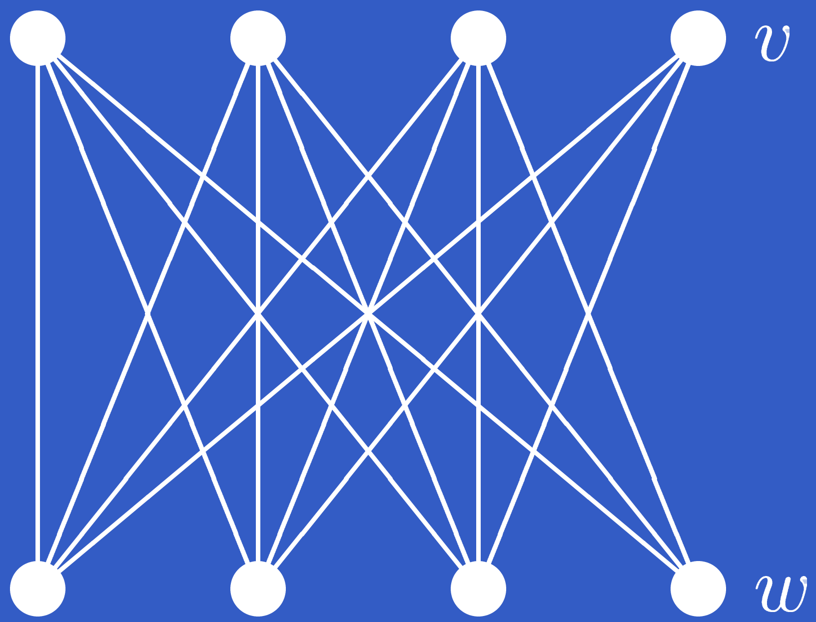
Complete bipartite graphs

Consider a complete bipartite graph $K_{4,4}$:



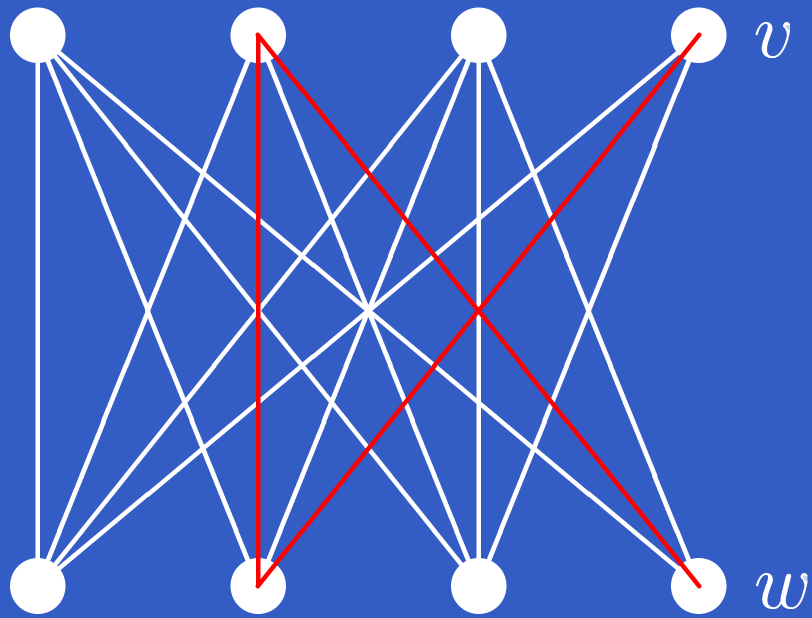
Complete bipartite graphs

Delete an edge vw.



Complete bipartite graphs

Distance between v and w is 3.



Minimal diameter 2 theorem

THEOREM [Füredi, 1992]
There is a n_0 such that if $n \geq n_0$ and G_n is a minimal graph of diameter 2, then

$$|E(G_n)| \leq \left\lfloor \frac{n^2}{4} \right\rfloor$$

Minimal diameter 2 theorem

THEOREM [Füredi, 1992]
There is a n_0 such that if $n \geq n_0$ and G_n is a minimal graph of diameter 2, then

$$|E(G_n)| \leq \left\lfloor \frac{n^2}{4} \right\rfloor$$

Furthermore, equality occurs if and only if

$$G = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}.$$

Loebl's Conjecture

CONJECTURE [Loebl]

If G_n is a graph on n vertices, and at least $n/2$ vertices have degrees at least $n/2$, then G contains, as subgraphs, all trees with at most $n/2$ edges.

Loebl's Conjecture

CONJECTURE [Loebl]
If G_n is a graph on n vertices, and at least $n/2$ vertices have degrees at least $n/2$, then G contains, as subgraphs, all trees with at most $n/2$ edges.

THEOREM [Ajtai-Komlós-Szemerédi, 1994]
For every $\epsilon > 0$, there is a n_0 such that if G_n has at least $(1+\epsilon)n/2$ vertices of degree at least $(1+\epsilon)n/2$, then G_n contains, as subgraphs, all trees with at most $n/2$ edges.

Loebl's Conjecture

CONJECTURE [Loebl]
If G_n is a graph on n vertices, and at least $n/2$ vertices have degrees at least $n/2$, then G contains, as subgraphs, all trees with at most $n/2$ edges.

THEOREM [Zhao, 2003?]
There is a constant n_0 so that Loebl's Conjecture holds for $n \geq n_0$.

Hajnal-Szemerédi

There is a classical theorem, not proven by regularity, that gives a condition for which a graph can be covered by copies of K_p.

Define $\delta(G) = \min\{\deg(v) : v \in V(G)\}$.

Hajnal-Szemerédi

There is a classical theorem, not proven by regularity, that gives a condition for which a graph can be covered by copies of K_p.

Define $\delta(G) = \min\{\deg(v) : v \in V(G)\}$.

THEOREM [Hajnal-Szemerédi, 1969]
If $\delta(G_n) \geq (1 - 1/p)n$, then G_n contains $\lfloor n/p \rfloor$ vertex-disjoint copies of K_p.

Hajnal-Szemerédi

There is a classical theorem, not proven by regularity, that gives a condition for which a graph can be covered by copies of K_p.

Define $\delta(G) = \min\{\deg(v) : v \in V(G)\}$.

THEOREM [Hajnal-Szemerédi, 1969]
If $\delta(G_n) \geq (1 - 1/p)n$, then G_n contains $\lfloor n/p \rfloor$ vertex-disjoint copies of K_p.

Let's just deal with $p = 3$.

Hajnal-Szemerédi

There is a classical theorem, not proven by regularity, that gives a condition for which a graph can be covered by copies of K_p.

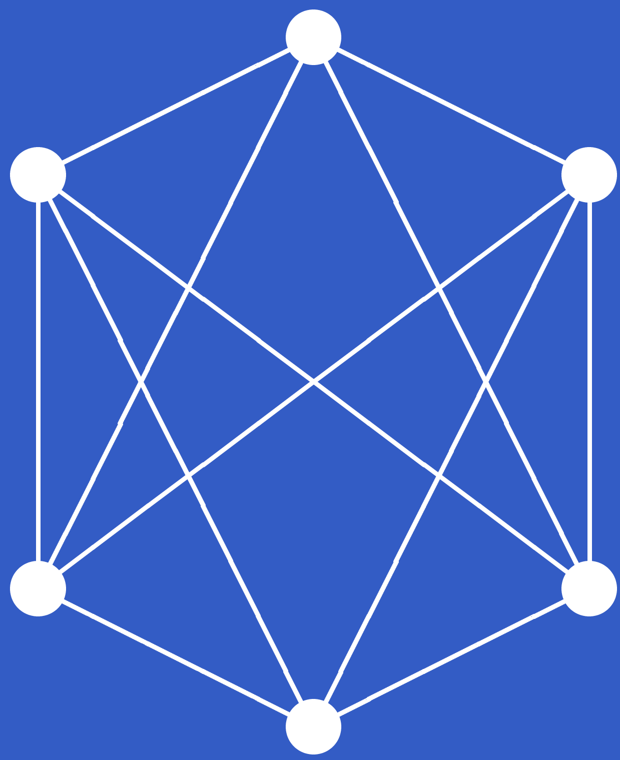
Define $\delta(G) = \min\{\deg(v) : v \in V(G)\}$.

THEOREM [Corrádi-Hajnal, 1963]
If $\delta(G_n) \geq (2/3)n$, then G_n contains $\lfloor n/3 \rfloor$ vertex-disjoint copies of K_3.

Let's just deal with $p = 3$.

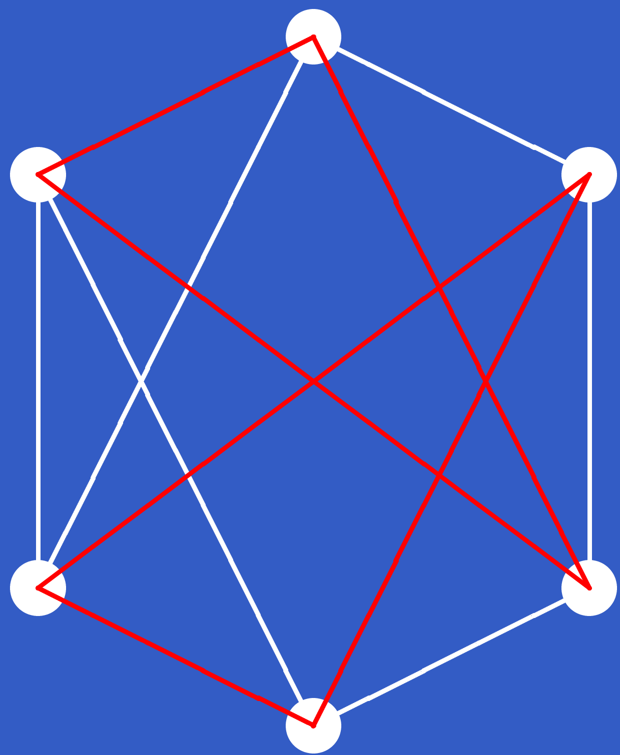
A small example

Here the minimum degree is
$4 = (2/3) \times 6$.



A small example

Hajnal-Szemerédi says that it can be covered by triangles (K_3's).



Tripartite version

The following conjecture is a natural extension of Corrádi-Hajnal, but not a consequence:

PROBLEM

Let G be a graph that is

- 3-partite (tripartite),

Tripartite version

The following conjecture is a natural extension of Corrádi-Hajnal, but not a consequence:

PROBLEM

Let G be a graph that is

- 3-partite (tripartite),
- has N vertices in each partition,

Tripartite version

The following conjecture is a natural extension of Corrádi-Hajnal, but not a consequence:

PROBLEM

Let G be a graph that is

- 3-partite (tripartite),
- has N vertices in each partition,
- has each vertex adjacent to $> \frac{2}{3}N$ vertices in each of the other classes.

Tripartite version

The following conjecture is a natural extension of Corrádi-Hajnal, but not a consequence:

PROBLEM

Let G be a graph that is

- 3-partite (tripartite),
- has N vertices in each partition,
- has each vertex adjacent to $> \frac{2}{3}N$ vertices in each of the other classes.

Then G contains N vertex-disjoint copies of K_3.

Tripartite version

PROBLEM

Let G be a graph that is

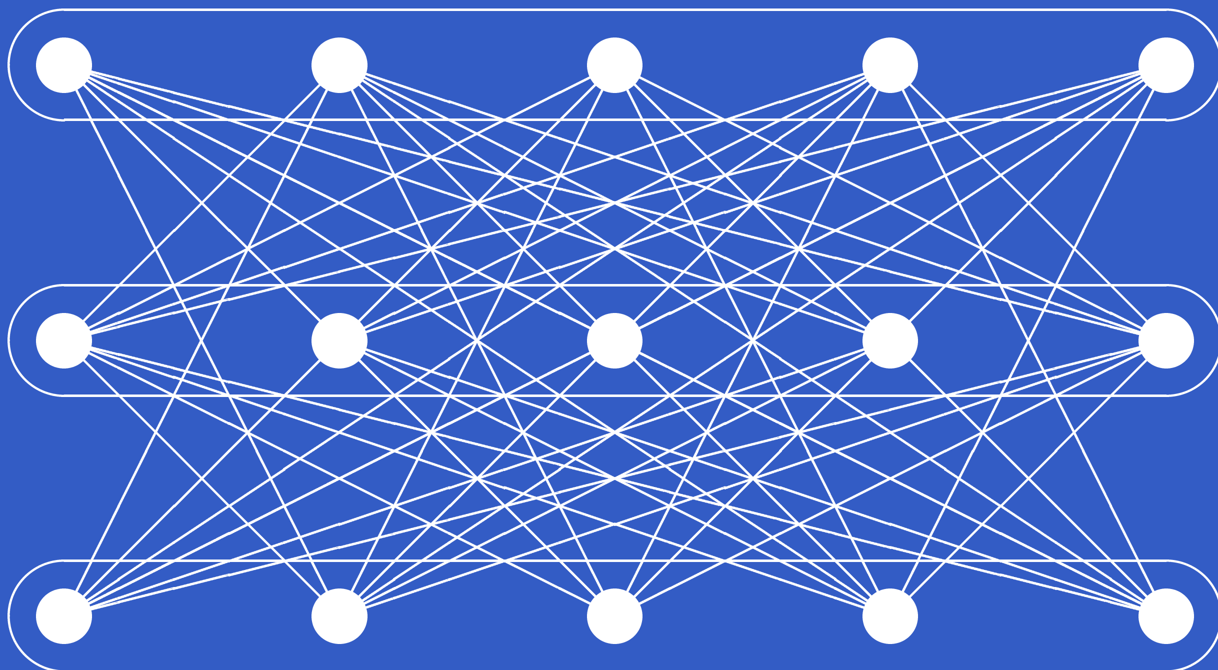
- 3-partite (tripartite),
- has N vertices in each partition,
- has each vertex adjacent to $> \frac{2}{3}N$ vertices in each of the other classes.

Then G contains N vertex-disjoint copies of K_3.

In fact, "$>$" can be replaced by "$\geq$", but there is one exceptional case.

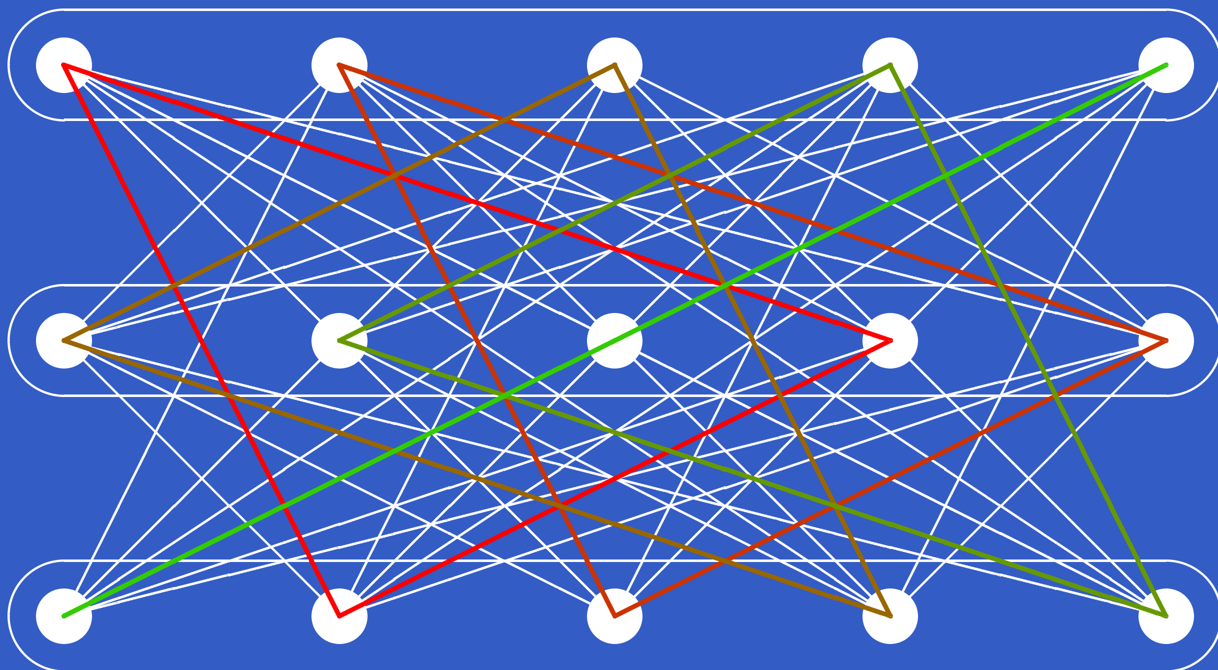
An example

Each vertex is adjacent to at least $4 > (2/3) \times 5$ vertices in each of the other classes.



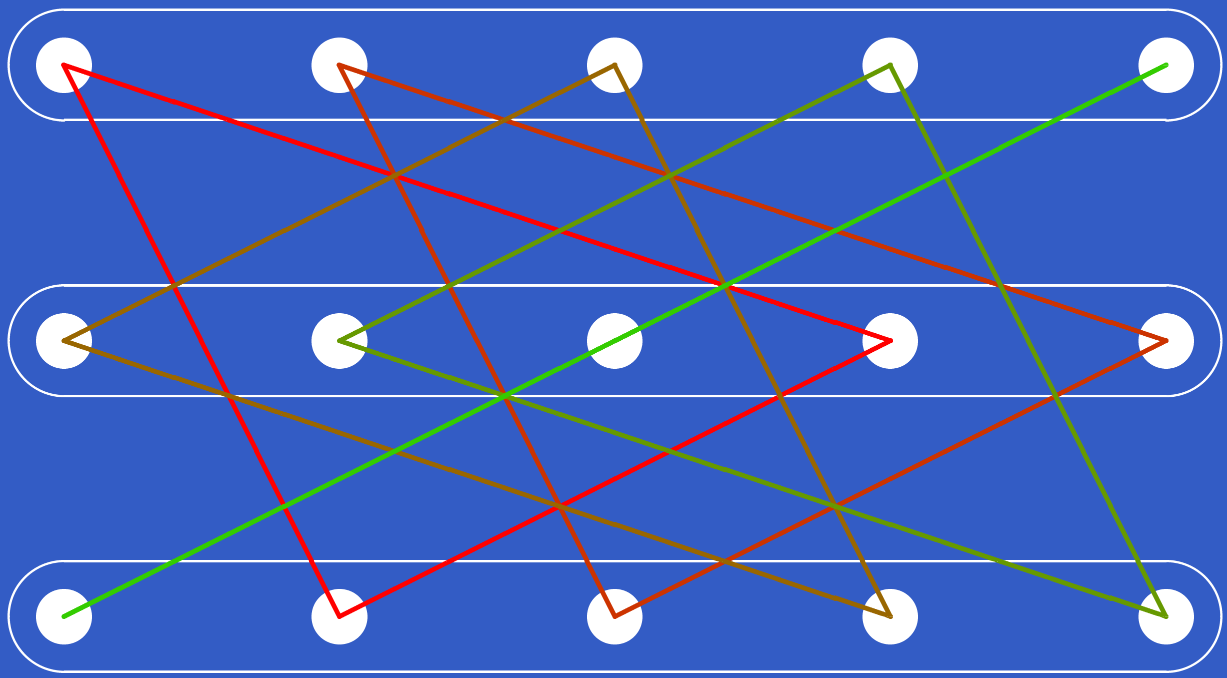
An example

Each vertex is adjacent to at least $4 > (2/3) \times 5$ vertices in each of the other classes.



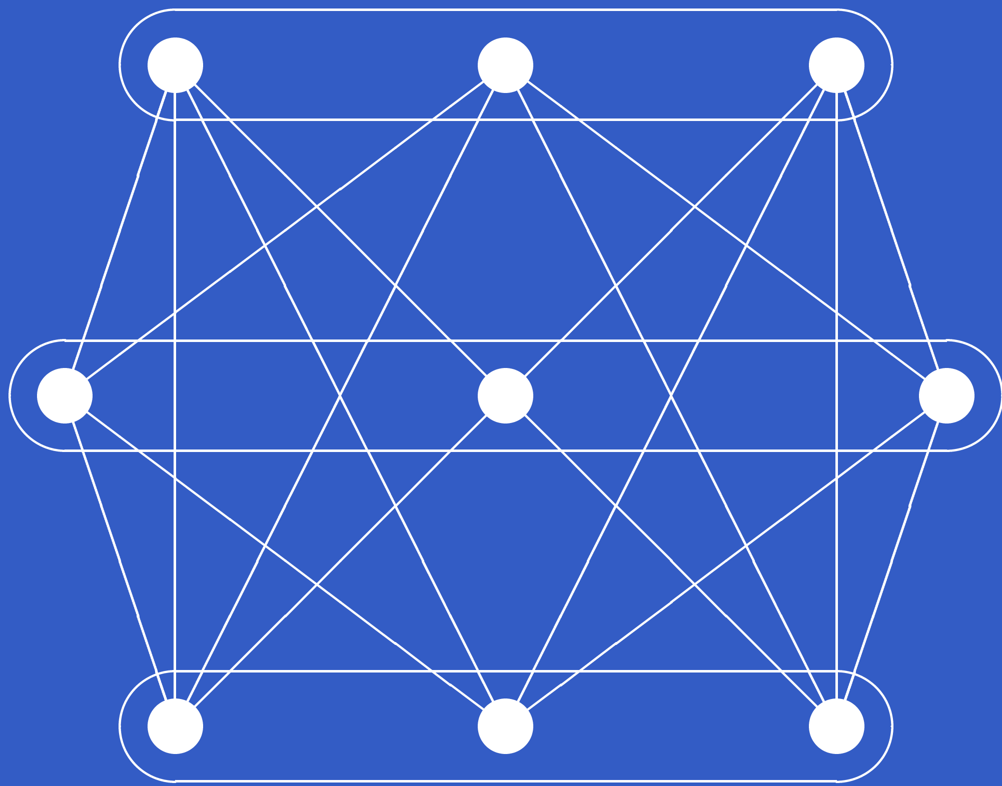
An example

Each vertex is adjacent to at least $4 > (2/3) \times 5$ vertices in each of the other classes.



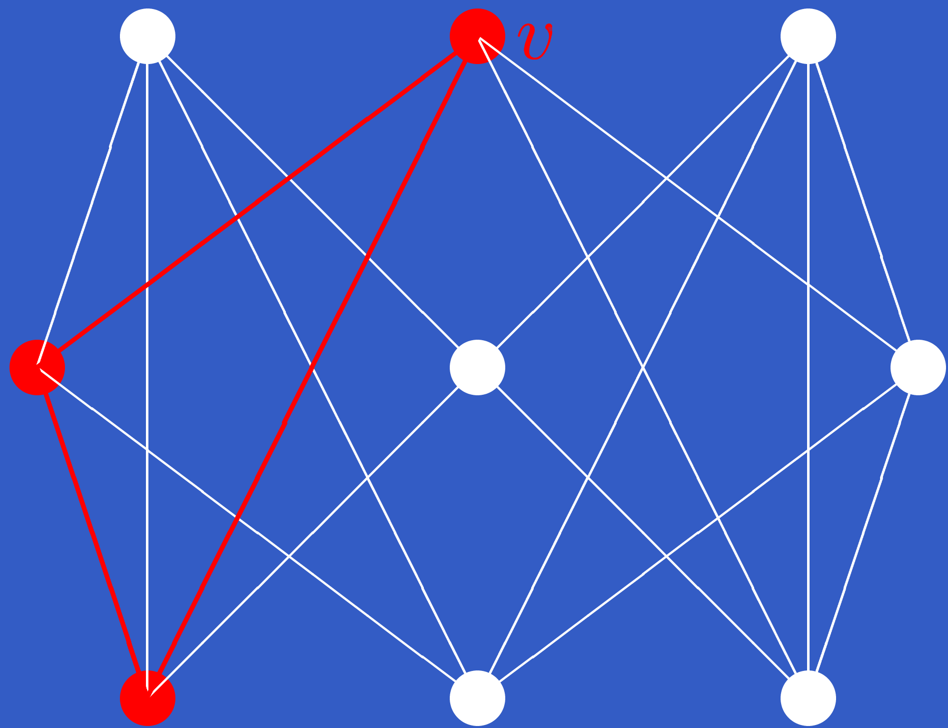
The result is tight

Consider the following example. Each vertex is adjacent to 2 in each other piece.



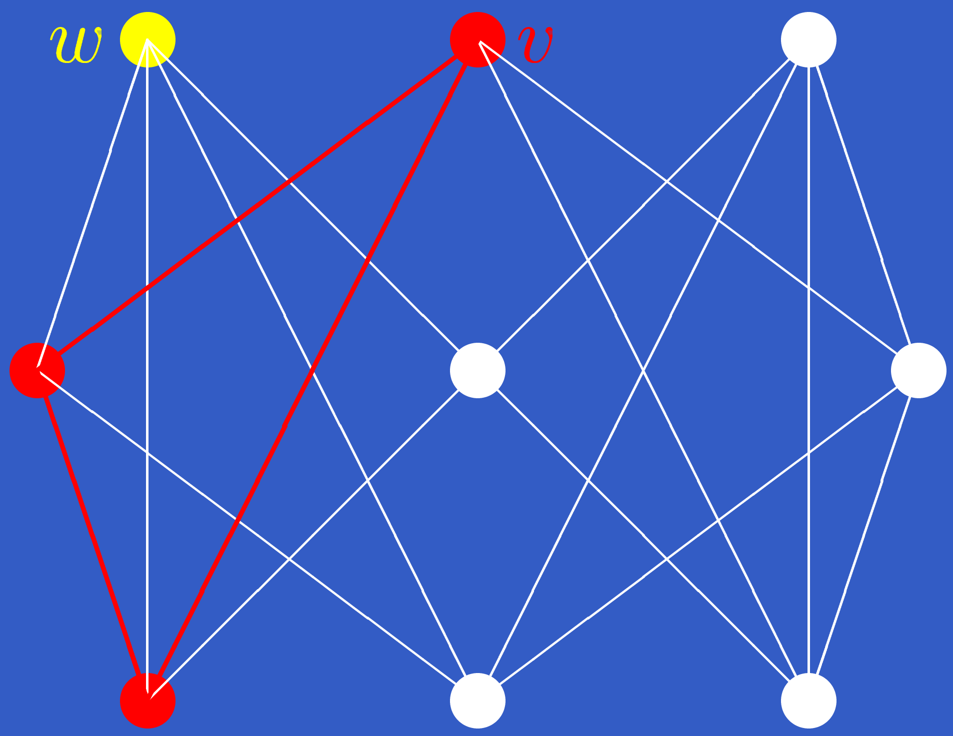
The result is tight

Without loss of generality, v must be in the red triangle.



The result is tight

But then, w cannot be in any triangle.



Example generalizes

This example generalizes.
For N, an odd multiple of 3, there is a graph with

- N vertices in each class,

Example generalizes

This example generalizes.
For N, an odd multiple of 3, there is a graph with

- N vertices in each class,
- each vertex adjacent to exactly $(2/3)N$ vertices in each of the two other parts, but

Example generalizes

This example generalizes.
For N, an odd multiple of 3, there is a graph with

- N vertices in each class,
- each vertex adjacent to exactly $(2/3)N$ vertices in each of the two other parts, but
- **NO** subgraph of N vertex-disjoint copies of K_3.

Example generalizes

This example generalizes.
For N, an odd multiple of 3, there is a graph with

- N vertices in each class,
- each vertex adjacent to exactly $(2/3)N$ vertices in each of the two other parts, but
- **NO** subgraph of N vertex-disjoint copies of K_3.

There is only one such graph (up to isomorphism) and we call it $\Gamma_3(N/3)$.

Tripartite

THEOREM [Magyar-M., 2002]
Let G be a tripartite graph such that

- there are N vertices in each partition

Tripartite

THEOREM [Magyar-M., 2002]

Let G be a tripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(2/3)N$ vertices in each of the other parts, and

Tripartite

THEOREM [Magyar-M., 2002]

Let G be a tripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(2/3)N$ vertices in each of the other parts, and
- $N \geq N_0$ for some absolute constant N_0.

Tripartite

THEOREM [Magyar-M., 2002]

Let G be a tripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(2/3)N$ vertices in each of the other parts, and
- $N \geq N_0$ for some absolute constant N_0.

Then either G has a subgraph which is N vertex-disjoint triangles, or

Tripartite

THEOREM [Magyar-M., 2002]
Let G be a tripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(2/3)N$ vertices in each of the other parts, and
- $N \geq N_0$ for some absolute constant N_0.

Then either G has a subgraph which is N vertex-disjoint triangles, or $G = \Gamma_3(N/3)$ for $N/3$ an odd integer.

What's with N_0?

Yep, that N_0 is the $M(\epsilon)$ from the Regularity Lemma.

Quadripartite?

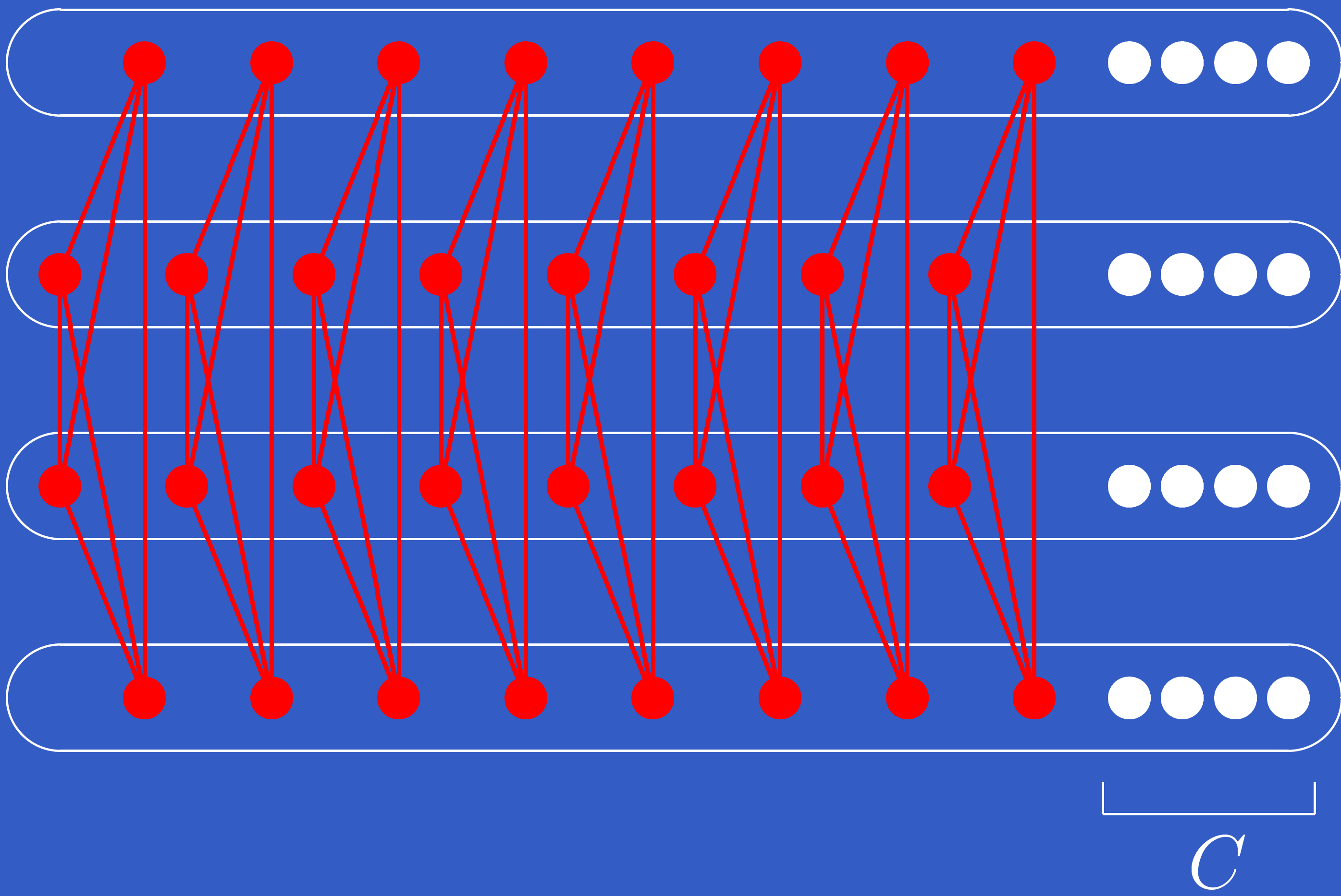
A partial result in the quadripartite case:

THEOREM [Fischer, 1999]
Let G be a quadripartite graph with

- N vertices in each part,
- each vertex adjacent to at least $(3/4)N$ vertices in each of the other three parts,

Then there is an absolute constant C such that G has a subgraph which is a family of $N - C$ vertex-disjoint K_4's.

Diagram of Fischer's result



Quadripartite!

THEOREM [M.-Szemerédi, 200?]

Let G be a quadripartite graph such that

- there are N vertices in each partition

Quadripartite!

THEOREM [M.-Szemerédi, 200?]

Let G be a quadripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(3/4)N$ vertices in each of the other parts, and

Quadripartite!

THEOREM [M.-Szemerédi, 200?]
Let G be a quadripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(3/4)N$ vertices in each of the other parts, and
- $N \geq N_0$ for some absolute constant N_0.

Quadripartite!

THEOREM [M.-Szemerédi, 200?]
Let G be a quadripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(3/4)N$ vertices in each of the other parts, and
- $N \geq N_0$ for some absolute constant N_0.

Then G has a subgraph which is N vertex-disjoint K_4's.

Quadripartite!

THEOREM [M.-Szemerédi, 200?]

Let G be a quadripartite graph such that

- there are N vertices in each partition
- each vertex is adjacent to at least $(3/4)N$ vertices in each of the other parts, and
- $N \geq N_0$ for some absolute constant N_0.

Then G has a subgraph which is N vertex-disjoint K_4's.

This case has no exceptional graph.

A final application

We will begin with a family of graphs:

$$\mathcal{H}(n,d) \stackrel{\text{def}}{=} \{H : \delta(H) \geq dn\}$$

A final application

We will begin with a family of graphs:

$$\mathcal{H}(n,d) \stackrel{\text{def}}{=} \{H : \delta(H) \geq dn\}$$

We will take an **arbitrary** $H \in \mathcal{H}(n,d)$.

A final application

We will begin with a family of graphs:

$$\mathcal{H}(n,d) \stackrel{\text{def}}{=} \{H : \delta(H) \geq dn\}$$

We will take an **arbitrary** $H \in \mathcal{H}(n,d)$.

Then, add m edges at random to H, forming G.

A final application

We will begin with a family of graphs:

$$\mathcal{H}(n,d) \stackrel{\text{def}}{=} \{H : \delta(H) \geq dn\}$$

We will take an **arbitrary** $H \in \mathcal{H}(n,d)$.

Then, add m edges at random to H, forming G.

What is the diameter of the “random” graph G?

Many possibilities

Let's assume d is a small constant.

H could be a variety of possibilites, including

- a traditional random graph,
- an ϵ-regular pair,
- $\lfloor 1/d \rfloor$ disjoint cliques

Diameter 7

- Fix a constant d.

Diameter 7

- Fix a constant d.
- Let H be an arbitrary member of $\mathcal{H}(n, d)$.

Diameter 7

- Fix a constant d.
- Let H be an arbitrary member of $\mathcal{H}(n, d)$.
- Add $m = m(n)$ edges at random to H, forming G.

Diameter 7

- Fix a constant d.
- Let H be an arbitrary member of $\mathcal{H}(n, d)$.
- Add $m = m(n)$ edges at random to H, forming G.
- Let $m \to \infty$ as $n \to \infty$.

Then, $\operatorname{diam}(G) \leq 7$.

The algorithm

Construct $v_1, v_2, \ldots$ greedily such that

$$\mathrm{dist}(v_i, v_{i+1}) = 3.$$

The algorithm

Construct $v_1, v_2, \ldots$ greedily such that

$$\text{dist}(v_i, v_{i+1}) = 3.$$

Since $|N(v_i)| \geq dn$, we stop in $\lfloor 1/d \rfloor$ steps.

Quick facts

- By the partition, every vertex is in the first or second neighborhood of some v_i.

Quick facts

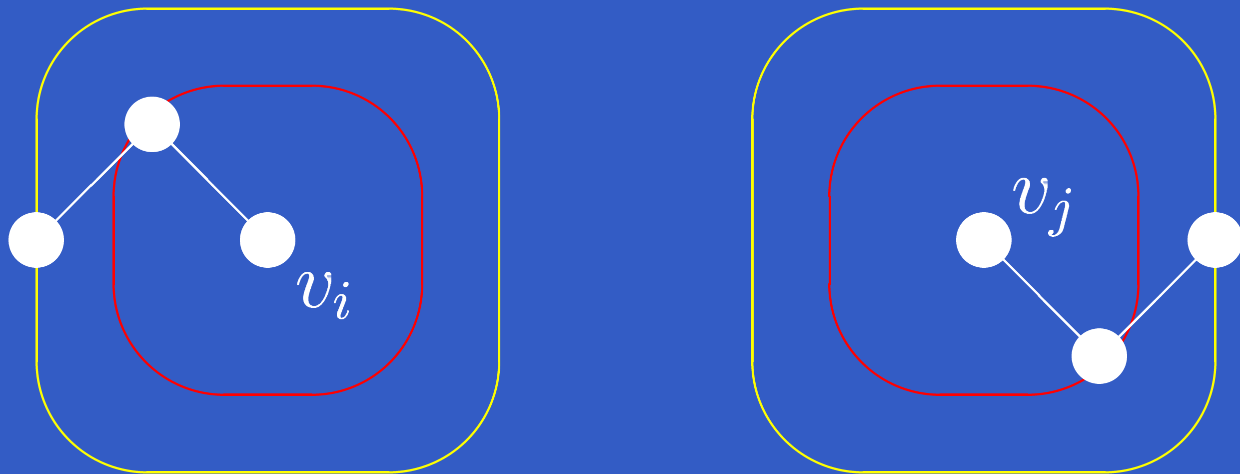
- By the partition, every vertex is in the first or second neighborhood of some v_i.
- There is an edge in any pair $(N(v_i), N(v_j))$ because $m(n) \to \infty$.

Quick facts

- By the partition, every vertex is in the first or second neighborhood of some v_i.
- There is an edge in any pair $(N(v_i), N(v_j))$ because $m(n) \to \infty$.
- So, for vertices u and w, the worst case is if they are in second neighborhoods of different v_i.

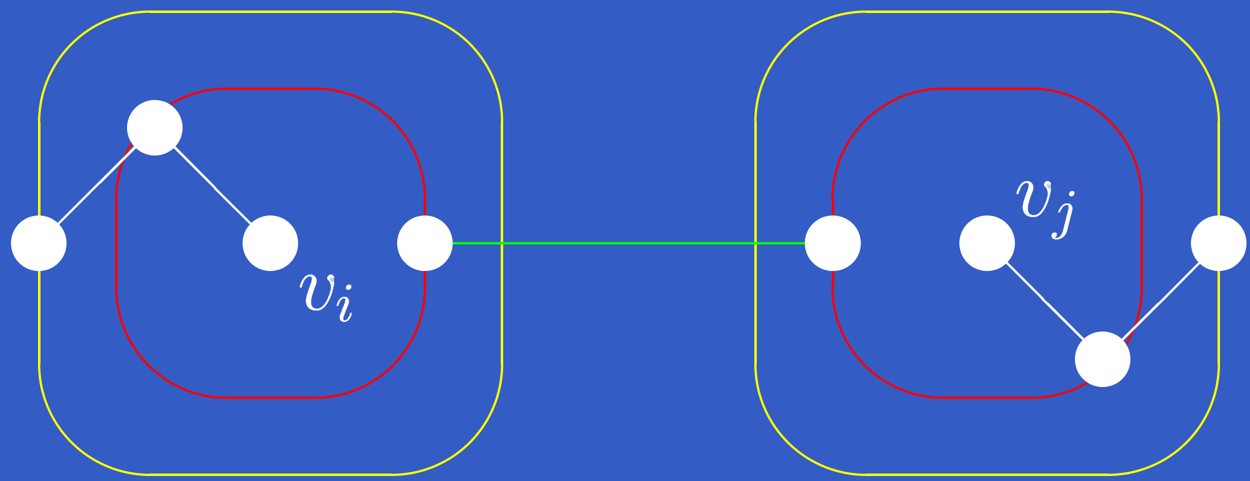
Diagram

First neighborhoods are red ovals.
Second neighborhoods are yellow ovals.



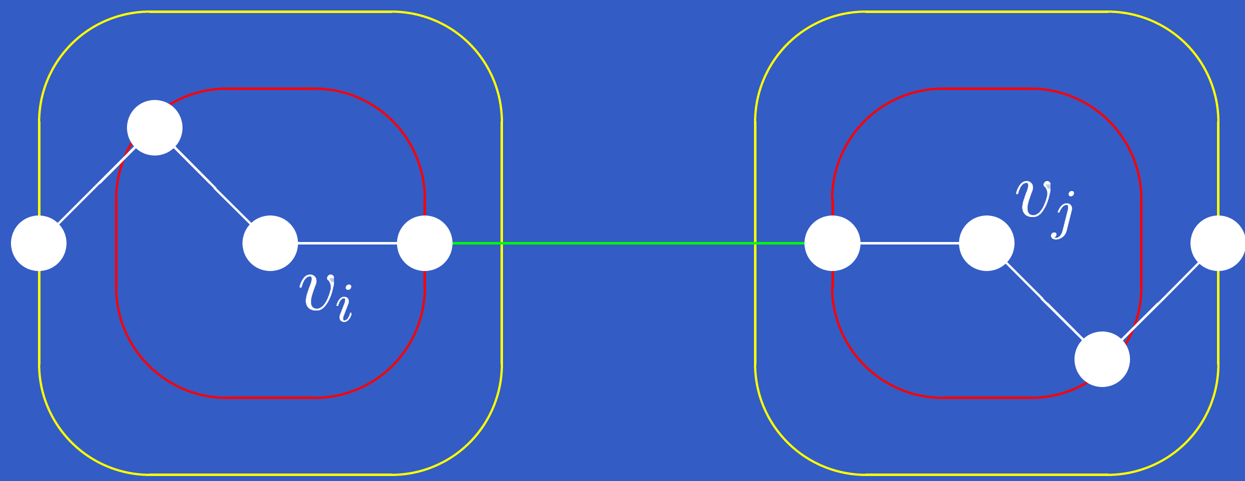
Diagram

A random edge is in $(N(v_i), N(v_j))$.



Diagram

A random edge is in $(N(v_i), N(v_j))$.
Voilá! A path of length 7.



Infinite edges

We can do better with Regularity:

THEOREM [Bohman-Frieze-Krivelevich-M., 200?]
Fix a constant d. Let H be an arbitrary member of $\mathcal{H}(n,d)$. Add m edges at random to H, forming G. If $m \to \infty$ as $n \to \infty$, then

$$\Pr(\mathrm{diam}(G) \leq ?) \to 1$$

as $n \to \infty$.

Infinite edges

We can do better with Regularity:

THEOREM [Bohman-Frieze-Krivelevich-M., 200?]
Fix a constant d. Let H be an arbitrary member of $\mathcal{H}(n, d)$. Add m edges at random to H, forming G. If $m \to \infty$ as $n \to \infty$, then

$$\Pr(\mathrm{diam}(G) \leq 5) \to 1$$

as $n \to \infty$.

Small world problem

In 1967, Stanley Milgram conducted a famous experiment.

Small world problem

In 1967, Stanley Milgram conducted a famous experiment.

He asked 160 families in Omaha, NE, to get a package to a stockbroker in Boston who had a home in Sharon, MA.

Small world problem

In 1967, Stanley Milgram conducted a famous experiment.

He asked 160 families in Omaha, NE, to get a package to a stockbroker in Boston who had a home in Sharon, MA.

It had to be mailed or carried by hand from one acquaintance to the next, until it was delivered in person to the broker. Before the experiment began, Milgram asked his colleagues how many steps they thought it would take for the packages to make the trip.

Small world problem

The consensus was around 100.

Small world problem

The consensus was around 100.

In fact, it took about five or six intermediary transactions for most of the packages to get from Omaha to the broker

Small world problem

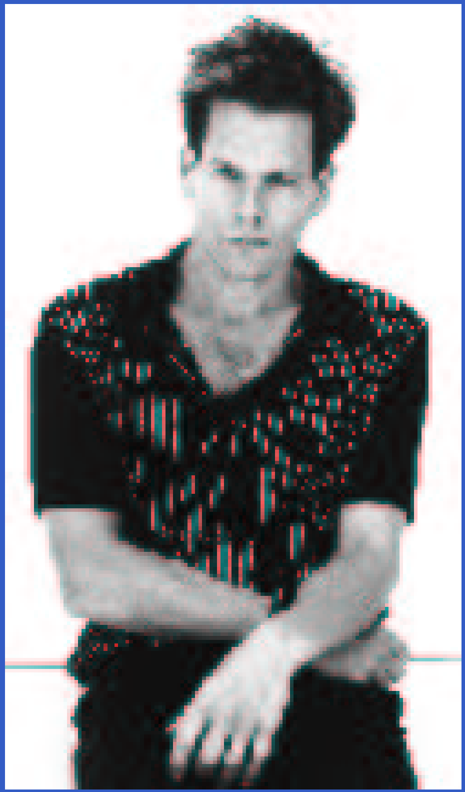
The consensus was around 100.

In fact, it took about five or six intermediary transactions for most of the packages to get from Omaha to the broker

– the "six degrees of separation" that we’re familiar with today.

Popular culture

We've seen similar phenomena in the Kevin Bacon game.



Popular culture

Also in the Erdős number project.

Our answer

So, if familiarity grows proportionally with the size of the population and the amount of randomness in the system grows also, then

Our answer

So, if familiarity grows proportionally with the size of the population and the amount of randomness in the system grows also, then

It's FIVE degrees of separation, not six!

Thanks

Thank you for letting me talk today.

The file for this talk is available online at my website:
http://www.math.cmu.edu/~rymartin
These slides were created by the Prosper document preparation system.