

# Combinatorics

## Ramsey Theory

尹一通 Nanjing University, 2026 Spring

# Ramsey's Theorem



Frank P. Ramsey  
(1903-1930)

*“In any party of six people, either at least three of them are mutual strangers or at least three of them are mutual acquaintances”*

Color edges of  $K_6$  with 2 colors.  
There must be a **monochromatic**  $K_3$ .

ON A PROBLEM OF FORMAL LOGIC

*By* F. P. RAMSEY.

[Received 28 November, 1928.—Read 13 December, 1928.]

This paper is primarily concerned with a special case of one of the leading problems of mathematical logic, the problem of finding a regular procedure to determine the truth or falsity of any given logical formula\*. But in the course of this investigation it is necessary to use certain theorems on combinations which have an independent interest and are most conveniently set out by themselves beforehand.



Frank P. Ramsey  
(1903-1930)

$R(k,l) \triangleq$  the smallest integer satisfying:

if  $n \geq R(k,l)$ , then no matter how to color edges of  $K_n$  with ■ and ■, there must exist a red  $K_k$  or a blue  $K_l$ .

2-coloring of  $K_n$

$$f : \binom{[n]}{2} \rightarrow \{\text{red, blue}\}$$



**Ramsey Theorem**

$R(k, l)$  is finite.

$$R(3,3) = 6$$



Frank P. Ramsey  
(1903-1930)

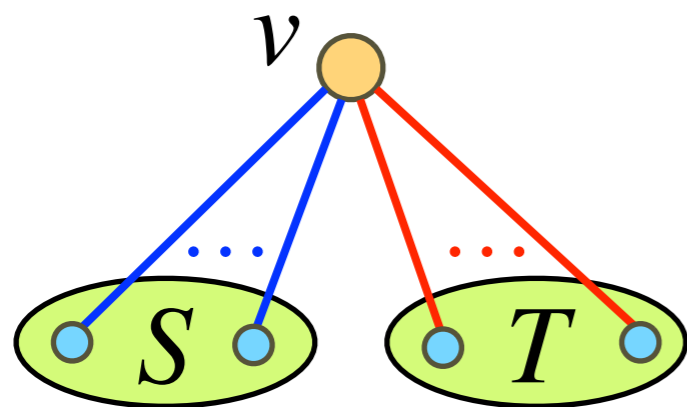
if  $n \geq R(k,l)$ , then no matter how to color edges of  $K_n$  with  and , there must exist a red  $K_k$  or a blue  $K_l$ .

$$R(k,2) = k ; \quad R(2,l) = l ;$$

$$R(k,l) \leq R(k, l-1) + R(k-1, l)$$

if  $n \geq R(k, l)$ , then no matter how to color edges of  $K_n$  with ■ and ■, there must exist a red  $K_k$  or a blue  $K_l$ .

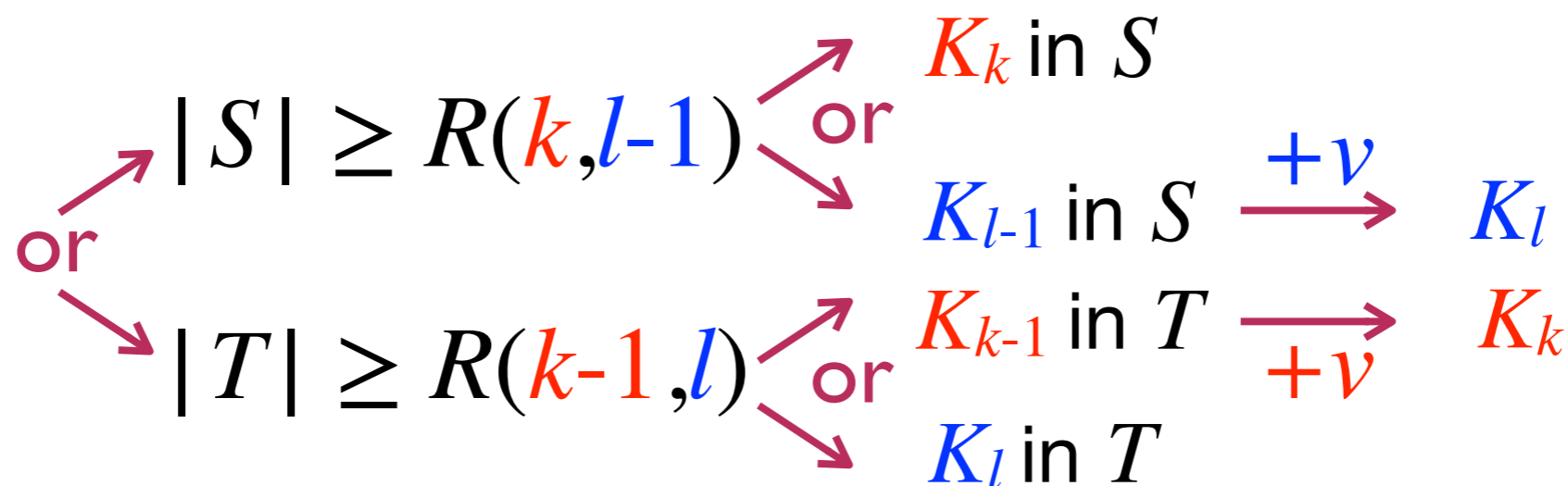
$$R(k, l) \leq R(k, l - 1) + R(k - 1, l)$$





take  $n = R(k, l - 1) + R(k - 1, l)$

arbitrary vertex  $v$

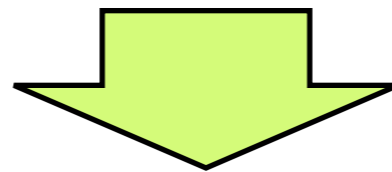
$$|S| + |T| + 1 = n = R(k, l - 1) + R(k - 1, l)$$



if  $n \geq R(k,l)$ , then no matter how to color edges of  $K_n$  with  and , there must exist a red  $K_k$  or a blue  $K_l$ .

$$R(k,2) = k ; \quad R(2,l) = l ;$$

$$R(k, l) \leq R(k, l - 1) + R(k - 1, l)$$



**Ramsey Theorem**  $R(k, l)$  is finite.

By induction:

$$R(k, l) \leq \binom{k + l - 2}{k - 1}$$

$$R(k, k) \geq n$$

“ $\exists$  a 2-coloring of  $K_n$  with no **monochromatic**  $K_k$ .”

a random 2-coloring of  $K_n$ :

$\forall \{u, v\} \in K_n$ , uniformly and independently  $\begin{cases} uv \\ uv \end{cases}$

$\forall S \in \binom{[n]}{k}$  event  $A_S$ :  $S$  is a **monochromatic**  $K_k$

$$\Pr[A_S] = 2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$$

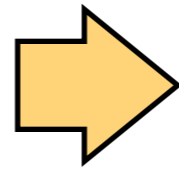
$A_S, A_T$  dependent  $\longleftrightarrow |S \cap T| \geq 2$

max degree of dependency graph  $d \leq \binom{k}{2} \binom{n}{k-2}$

To prove:  $\Pr \left[ \bigwedge_{S \in \binom{[n]}{k}} \overline{A_S} \right] > 0$

## Lovász Local Lemma

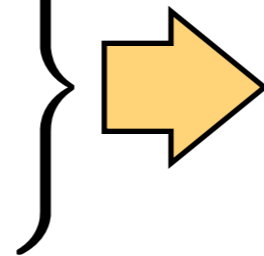
- $\forall i, \Pr[A_i] \leq p$
- $ep(d+1) \leq 1$



$$\Pr \left[ \bigwedge_{i=1}^n \overline{A_i} \right] > 0$$

$$\Pr[A_S] = 2^{1 - \binom{k}{2}}$$

$$d \leq \binom{k}{2} \binom{n}{k-2}$$



for some  $n = ck2^{k/2}$   
with constant  $c$

$$e2^{1 - \binom{k}{2}} (d+1) \leq 1$$

To prove:  $\Pr \left[ \bigwedge_{S \in \binom{[n]}{k}} \overline{A_S} \right] > 0$

$$R(k, k) \geq n = \Omega(k2^{k/2})$$



# Multicolor

if  $n \geq R(k,l)$ , for any 2-coloring of edges of  $K_n$ ,  
there exists a red  $K_k$  or a blue  $K_l$ .

$$R(r; k_1, k_2, \dots, k_r)$$

if  $n \geq R(r; k_1, k_2, \dots, k_r)$ ,

for any  $r$ -coloring of edges of  $K_n$ , for some  $i \in [r]$   
there exists a  $k_i$ -clique monochromatic with color  $i$ .

$$R(r; k_1, \dots, k_{r-2}, k_{r-1}, k_r) \leq R(r-1; k_1, \dots, k_{r-2}, \underbrace{R(2; k_{r-1}, k_r)})$$

the mixing color trick:

color 

# Multicolor

if  $n \geq R(k, l)$ , for any 2-coloring of edges of  $K_n$ ,  
there exists a red  $K_k$  or a blue  $K_l$ .

$R(r; k_1, k_2, \dots, k_r)$

if  $n \geq R(r; k_1, k_2, \dots, k_r)$ ,

for any  $r$ -coloring of edges of  $K_n$ , for some  $i \in [r]$   
there exists a  $k_i$ -clique monochromatic with color  $i$ .

## Ramsey Theorem

$R(r; k_1, k_2, \dots, k_r)$  is finite.

# Hypergraph

if  $n \geq R(r; k_1, k_2, \dots, k_r)$ ,

for any  $r$ -coloring of edges of  $K_n$ , for some  $i \in [r]$   
there exists a  $k_i$ -clique monochromatic with color  $i$ .

complete  $t$ -uniform hypergraph  $\binom{[n]}{t}$

$r$ -coloring  $f : \binom{[n]}{t} \rightarrow \{1, 2, \dots, r\}$

# Hypergraph

if  $n \geq R_t(r; k_1, k_2, \dots, k_r)$ ,

for any  $r$ -coloring of  $\binom{[n]}{t}$ , there exists  
a monochromatic  $\binom{S}{t}$  with color  $i$  and

$|S| = k_i$  for some  $i \in \{1, 2, \dots, r\}$

complete  $t$ -uniform hypergraph  $\binom{[n]}{t}$

$r$ -coloring  $f : \binom{[n]}{t} \rightarrow \{1, 2, \dots, r\}$

# Partition of Set Family

If  $n \geq R_t(r; k_1, k_2, \dots, k_r)$ ,

for any  $r$ -partition of  $\binom{[n]}{t} = C_1 \cup \dots \cup C_r$ ,  
there exists an  $S \subseteq [n]$  such that  $|S| = k_i$   
and  $\binom{S}{t} \subseteq C_i$  for some  $i \in \{1, 2, \dots, r\}$ .

Erdős-Rado **partition arrow**

$$n \rightarrow (k_1, k_2, \dots, k_r)^t$$

If  $n \geq R_t(r; k_1, k_2, \dots, k_r)$ ,

for any  $r$ -partition of  $\binom{[n]}{t} = C_1 \cup \dots \cup C_r$ ,  
there exists an  $S \subseteq [n]$  such that  $|S| = k_i$   
and  $\binom{S}{t} \subseteq C_i$  for some  $i \in \{1, 2, \dots, r\}$ .

mixing color:

$$R_t(r; k_1, \dots, k_{r-2}, k_{r-1}, k_r) \leq R_t(r-1; k_1, \dots, k_{r-2}, R_t(2; k_{r-1}, k_r))$$

$$R_t(k, l) \leq R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

$$n = R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

goal:  $\forall f : \binom{[n]}{t} \rightarrow \{\text{red}, \text{blue}\}$   
 $\exists \binom{X}{t}, |X| = k$  or  $\binom{Y}{t}, |Y| = l$

$$R_t(k, l) \leq R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

$$n = R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

$$\forall f : \binom{[n]}{t} \rightarrow \{\text{red}, \text{blue}\}$$

remove  $n$  from  $[n]$ , consider  $\binom{[n-1]}{t-1}$   
 ( $[n] = \{1, 2, \dots, n\}$ )

define  $f' : \binom{[n-1]}{t-1} \rightarrow \{\text{red}, \text{blue}\}$

$$\forall A \in \binom{[n-1]}{t-1}, f'(A) = f(A \cup \{n\})$$

$$n-1 = R_{t-1}(R_t(k-1, l), R_t(k, l-1))$$

or  $\left\{ \begin{array}{l} \exists S \subseteq [n-1], |S| = R_t(k-1, l), \binom{S}{t-1} \text{ by } f' \\ \exists T \subseteq [n-1], |T| = R_t(k, l-1), \binom{T}{t-1} \text{ by } f' \end{array} \right.$

$$R_t(k, l) \leq R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

$$n = R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

$$\forall f : \binom{[n]}{t} \rightarrow \{\text{red}, \text{blue}\}$$

$$\text{define } f' : \binom{[n-1]}{t-1} \rightarrow \{\text{red}, \text{blue}\}$$

$$\forall A \in \binom{[n-1]}{t-1}, \quad f'(A) = f(A \cup \{n\})$$

by  
symmetry

$$\exists S \subseteq [n-1], |S| = R_t(k-1, l), \binom{S}{t-1} \text{ by } f'$$

$$\text{or } \left\{ \begin{array}{l} \exists X \subseteq S, |X| = k-1, \binom{X}{t} \text{ by } f \\ \exists Y \subseteq S, |Y| = l, \binom{Y}{t} \text{ by } f \quad \checkmark \end{array} \right.$$

$$\checkmark \binom{X \cup \{n\}}{t} \text{ by } f$$

If  $n \geq R_t(r; k_1, k_2, \dots, k_r)$ ,

$\forall r$ -partition of  $\binom{[n]}{t} = C_1 \cup \dots \cup C_r$

$\exists i \in [r]$  and  $S \subseteq [n]$  with  $|S| = k_i$   
such that  $\binom{S}{t} \subseteq C_i$

mixing color:

$$R_t(r; k_1, \dots, k_{r-2}, k_{r-1}, k_r) \leq R_t(r-1; k_1, \dots, k_{r-2}, R_t(2; k_{r-1}, k_r))$$

$$R_t(k, l) \leq R_{t-1}(R_t(k-1, l), R_t(k, l-1)) + 1$$

**Theorem** (Ramsey 1930)

$R_t(r; k_1, k_2, \dots, k_r)$  is finite.

# Ramsey Theorem

If  $n \geq R_t(r; k_1, k_2, \dots, k_r)$ ,

$\forall r$ -partition of  $\binom{[n]}{t} = C_1 \cup \dots \cup C_r$

$\exists i \in [r]$  and  $S \subseteq [n]$  with  $|S| = k_i$   
such that  $\binom{S}{t} \subseteq C_i$

**Theorem** (Ramsey 1930)

$R_t(r; k_1, k_2, \dots, k_r)$  is finite.

# Ramsey Theorem

If  $n \geq R_t(r; k_1, k_2, \dots, k_r)$ ,

$\forall r$ -coloring  $f: \binom{[n]}{t} \rightarrow [r]$

$\exists i \in [r]$  and  $S \subseteq [n]$  with  $|S| = k_i$

such that  $f\left(\binom{S}{t}\right) = \{i\}$

**Theorem** (Ramsey 1930)

$R_t(r; k_1, k_2, \dots, k_r)$  is finite.

# Ramsey Theorem (diagonal)

If  $n \geq R_t(r; k) \triangleq R_t(r; \underbrace{k, k, \dots, k}_r)$ ,

$\forall r$ -coloring  $f: \binom{[n]}{t} \rightarrow [r]$

$\exists$  a monochromatic  $\binom{S}{t}$  with  $S \in \binom{[n]}{k}$

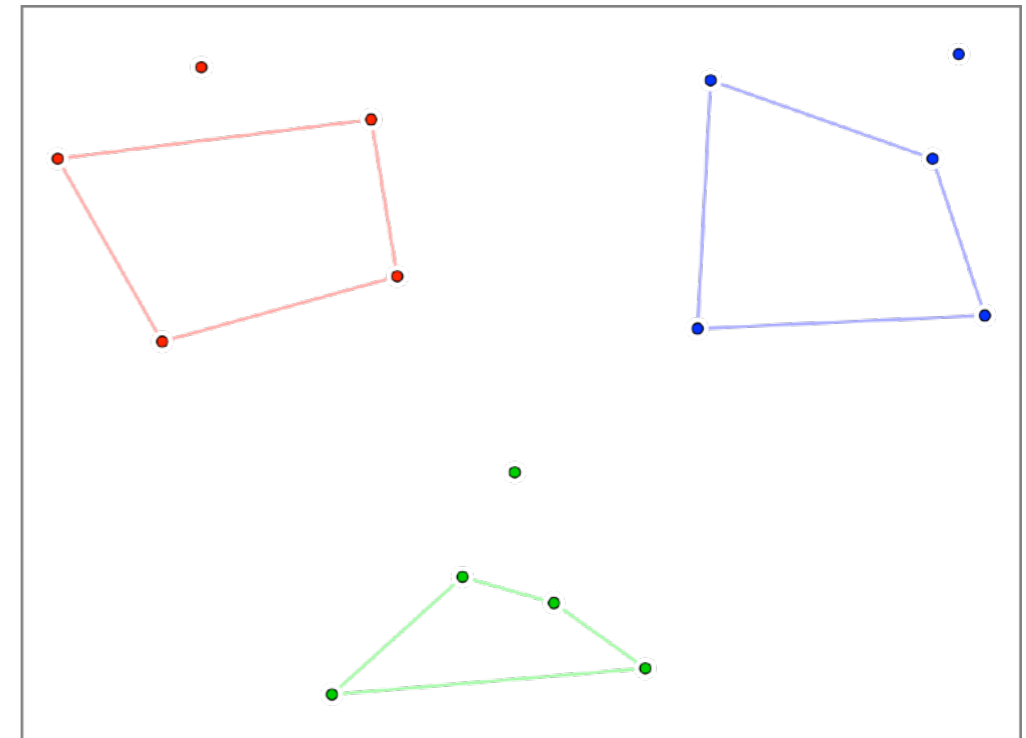
**Theorem** (Ramsey 1930)

$R_t(r; k)$  is finite.

# Applications of Ramsey's Theorem

# Happy Ending Problem

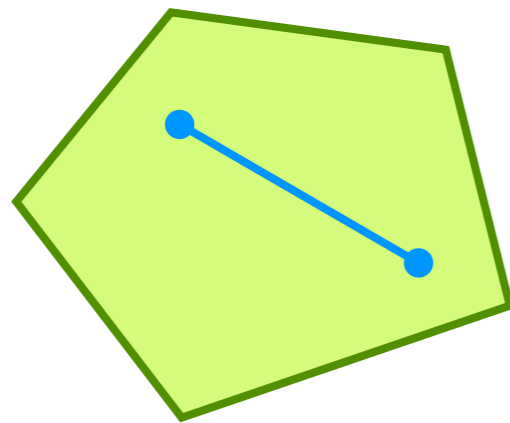
Any 5 points in the plane, no three on a line, has a subset of 4 points that form a convex quadrilateral.



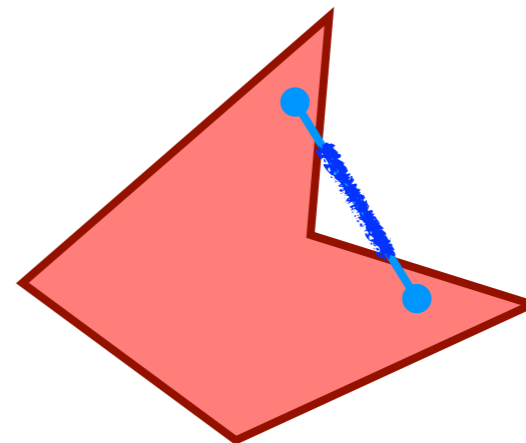
## **Theorem** (Erdős-Szekeres 1935)

$\forall m \geq 3, \exists N(m)$  such that any set of  $n \geq N(m)$  points in the plane, no three on a line, contains  $m$  points that are the vertices of a convex  $m$ -gon.

Polygon:



convex

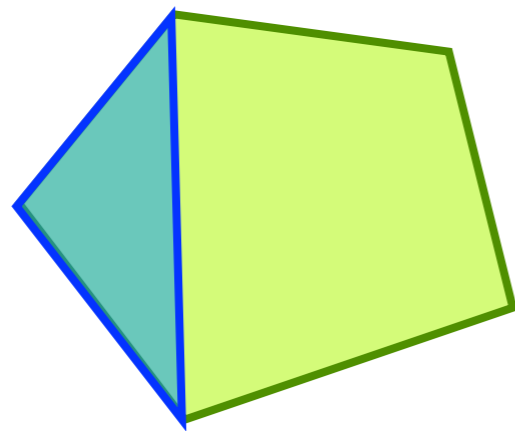


concave

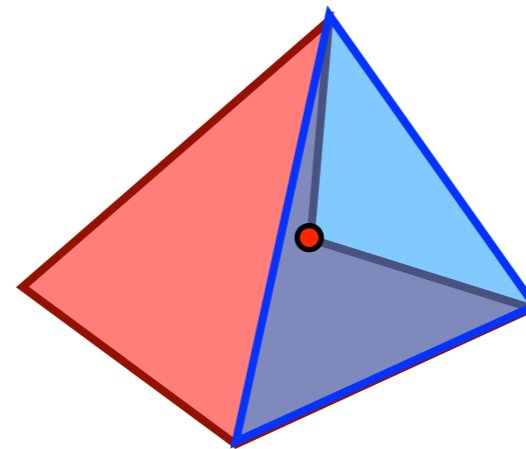
**Theorem** (Erdős-Szekeres 1935)

$\forall m \geq 3, \exists N(m)$  such that any set of  $n \geq N(m)$  points in the plane, no three on a line, contains  $m$  points that are the vertices of a convex  $m$ -gon.

Polygon:



convex



concave

## Theorem (Erdős-Szekeres 1935)

$\forall m \geq 3, \exists N(m)$  such that any set of  $n \geq N(m)$  points in the plane, no three on a line, contains  $m$  points that are the vertices of a convex  $m$ -gon.

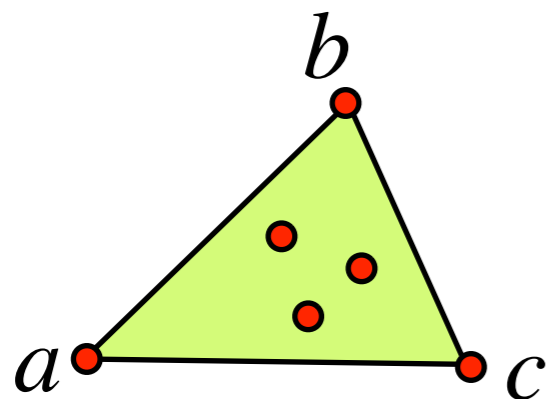
$$N(m) = R_3(2; m, m)$$

$$|X| \geq N(m)$$

$$\forall f : \binom{X}{3} \rightarrow \{0, 1\} \quad \exists S \subseteq X, |S| = m$$

monochromatic  $\binom{S}{3}$

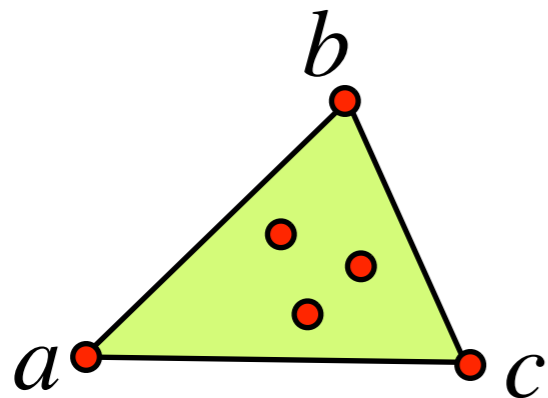
$X$ : set of points in the plane, no 3 on a line



$\forall a, b, c \in X, \triangle_{abc}$ : points in triangle  $abc$

$$f(\{a, b, c\}) = |\triangle_{abc}| \bmod 2$$

$X$ : set of points in the plane, no 3 on a line



$\forall a, b, c \in X, \triangle_{abc}$ : points in triangle  $abc$

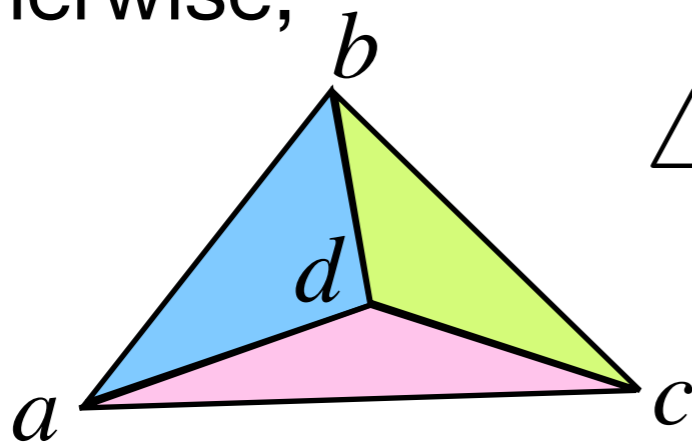
$$f(\{a, b, c\}) = |\triangle_{abc}| \bmod 2$$

$$|X| \geq R_3(2; m, m) \quad \forall f : \binom{X}{3} \rightarrow \{0, 1\}$$

$$\exists S \subseteq X, |S| = m \quad \text{monochromatic} \binom{S}{3}$$

$S$  is a convex  $m$ -gon

Otherwise,



disjoint union:

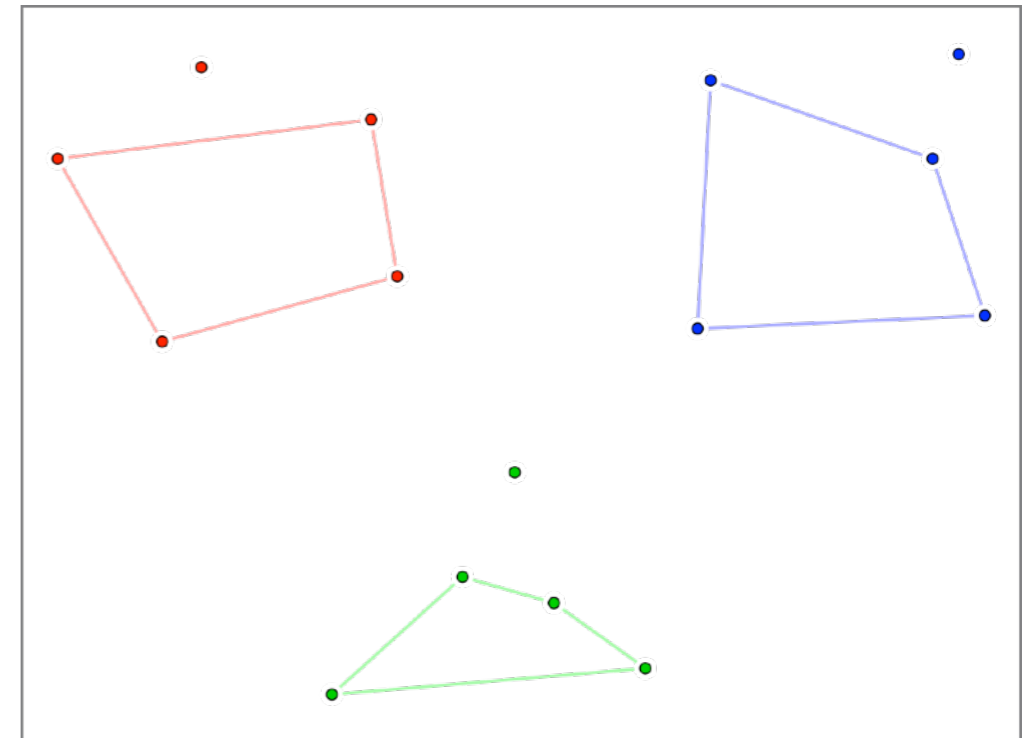
$$\triangle_{abc} = \triangle_{abd} \cup \triangle_{acd} \cup \triangle_{bcd} \cup \{d\}$$

$$f(abc) = f(abd) + f(acd) + f(bcd) + 1$$

**Contradiction!**

# Happy Ending Problem

Any 5 points in the plane, no three on a line, has a subset of 4 points that form a convex quadrilateral.



# Data Structures

Problem:

“Is  $x \in S$ ?”

data set  $S \in \binom{[N]}{n}$  key  $x \in [N]$  data universe  $[N]$

Solution:

Data structure:

sorted table

Search Alg:

~~binary search~~

Complexity:

$\geq \log_2 n$

memory accesses  
in the worst-case

“Is  $x \in S$ ?”  $x \in [N]$   $S \in \binom{[N]}{n}$

**Theorem** (Yao 1981)

If  $N \geq 2n$ , **on sorted table**, any search Alg requires  $\Omega(\log n)$  accesses in the worst-case.

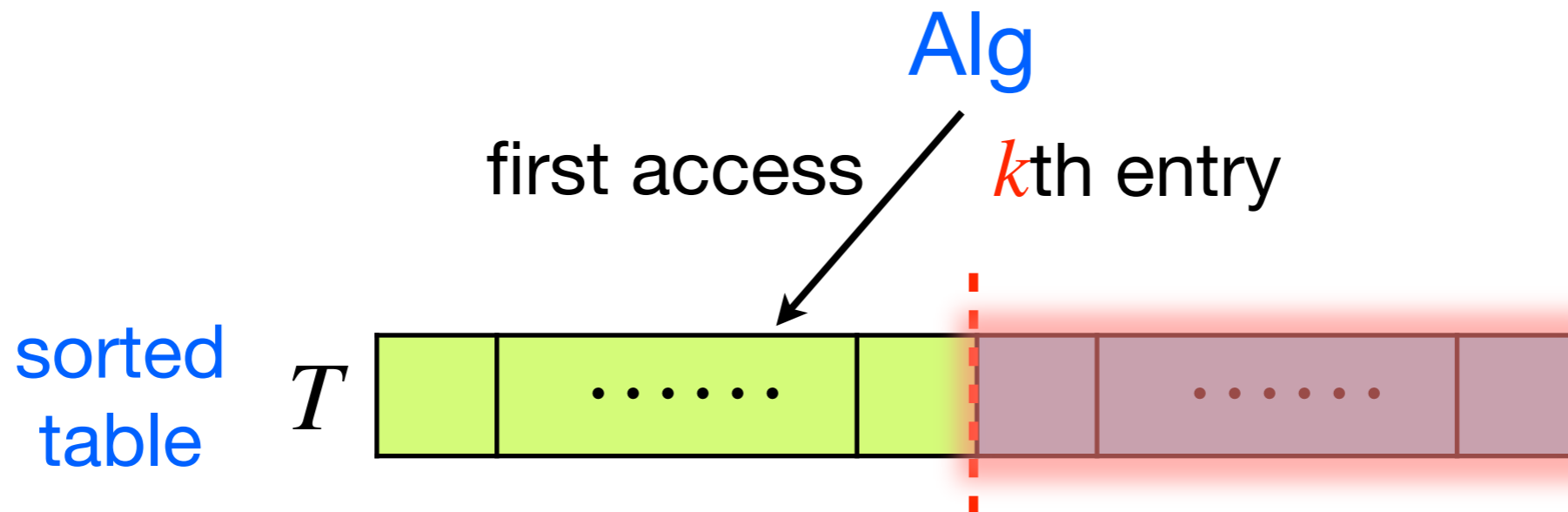
Problem: “Is  $n \in S$ ?”  $\forall S \in \binom{[N]}{n}$   $N \geq 2n$

Induction on  $n$ ,  $n = 2$ , trivial

Suppose it is true for any smaller  $n$ .

**adversarial argument + self-reduction**

Problem: “Is  $n \in S$ ?”  $\forall S \in \binom{[N]}{n}$   $N \geq 2n$



$$T[k] = \begin{cases} k & \text{if } k \leq \frac{n}{2} \\ N - (n - k) & \text{if } k > \frac{n}{2} \end{cases}$$

$$\binom{\{\frac{n}{2}, \dots, N - \frac{n}{2}\}}{\frac{n}{2}} \subseteq \binom{\{\frac{n}{2}, \dots, N\}}{\frac{n}{2}} \subseteq \text{possible} \{T[\frac{n}{2} + 1], \dots, T[n]\}$$

$$n' = \frac{n}{2} \quad N' = |\{\frac{n}{2}, \dots, N - \frac{n}{2}\}| \geq n \geq 2n'$$

relative key in  $[N']$ :  $n - \frac{n}{2} = n'$

Problem: “Is  $n \in S$ ?”  $\forall S \in \binom{[N]}{n}$   $N \geq 2n$

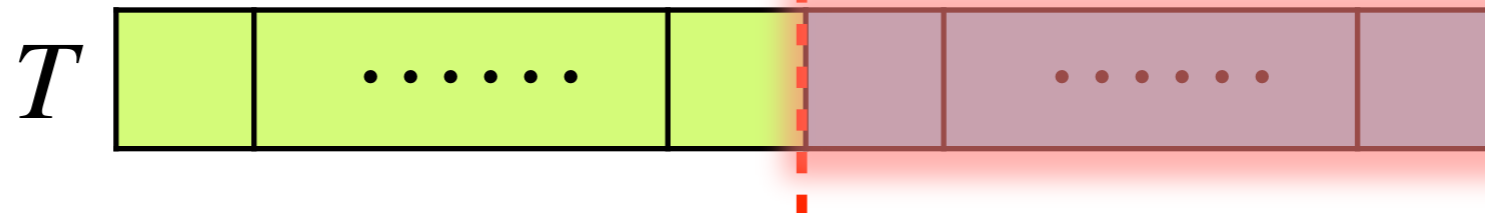
$$\geq 1 + \log \frac{n}{2} = \log n$$

Alg

first access

$k$ th entry

sorted  
table

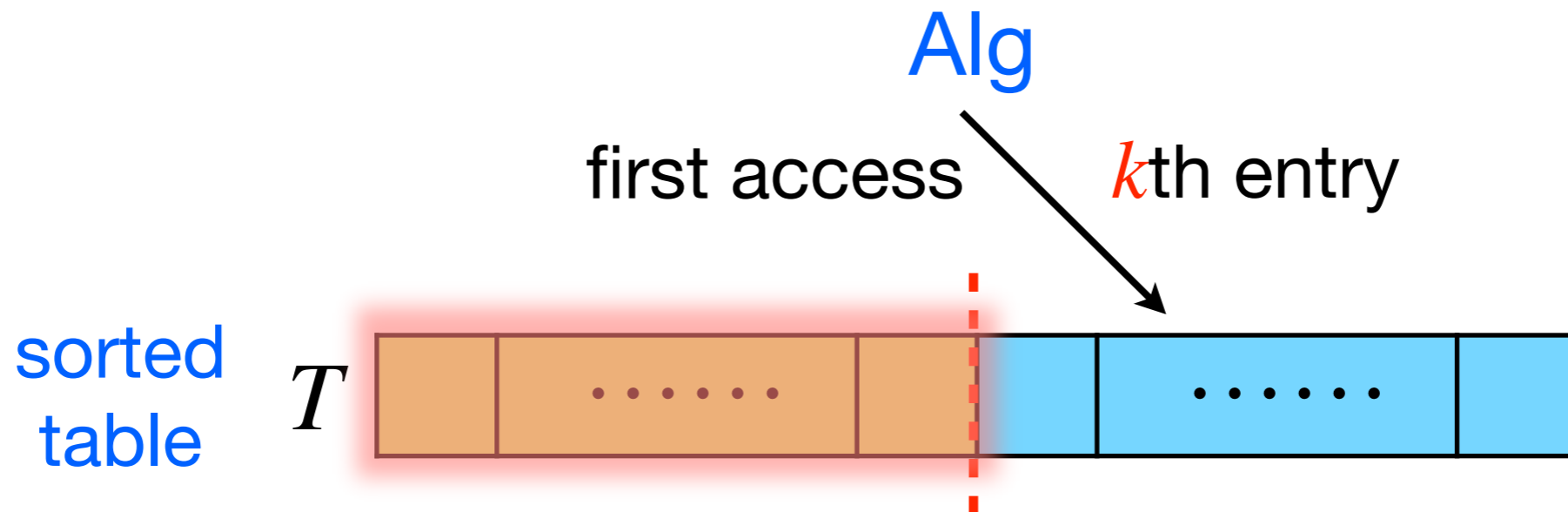



$$T[k] = \begin{cases} k & \text{if } k \leq \frac{n}{2} \\ N - (n - k) & \text{if } k > \frac{n}{2} \end{cases}$$

$n' = \frac{n}{2}$  “Is  $n' \in S$ ?”  $\forall S' \in \binom{[N']}{n'}$   $N' \geq 2n'$

I.H. require  $\log \frac{n}{2}$  memory accesses

Problem: "Is  $n \in S$ ?"  $\forall S \in \binom{[N]}{n}$   $N \geq 2n$



  $T[k] = \begin{cases} k & \text{if } k \leq \frac{n}{2} \\ N - (n - k) & \text{if } k > \frac{n}{2} \end{cases}$

$\binom{\{\frac{n}{2}, \dots, N - \frac{n}{2}\}}{\frac{n}{2}} \subseteq \binom{\{1, \dots, N - \frac{n}{2}\}}{\frac{n}{2}} \subseteq \text{possible } \{T[1], \dots, T[\frac{n}{2}]\}$

$n' = \frac{n}{2}$   $N' = |\{\frac{n}{2}, \dots, N - \frac{n}{2}\}| \geq 2n'$

relative key in  $[N']$ :  $n - \frac{n}{2} = n'$

“Is  $x \in S$ ?”  $x \in [N]$   $S \in \binom{[N]}{n}$

### Theorem (Yao 1981)

If  $N \geq 2n$ , **on sorted table**, any search Alg requires  $\Omega(\log n)$  accesses in the worst-case.

implicit data structure:

each  $S \in \binom{[N]}{n}$  is stored as a permutation of  $S$

$$f : \binom{[N]}{n} \rightarrow [n!]$$

$f(S) = \pi$   
dataset  $S: x_1 < \dots < x_n$   
table:  $(x_{\pi(1)}, \dots, x_{\pi(n)})$

$\binom{[N]}{n}$  is mapped to the same  $\pi$   **same as**

$\exists X \subseteq [N], |X| \geq 2n, \binom{X}{n}$  **monochromatic**  **sorted**

“Is  $x \in S$ ?”  $x \in [N]$   $S \in \binom{[N]}{n}$

**Theorem** (Yao 1981)

For sufficiently large  $N$ , on any implicit data structure, any search Alg requires  $\Omega(\log n)$  accesses in the worst-case.

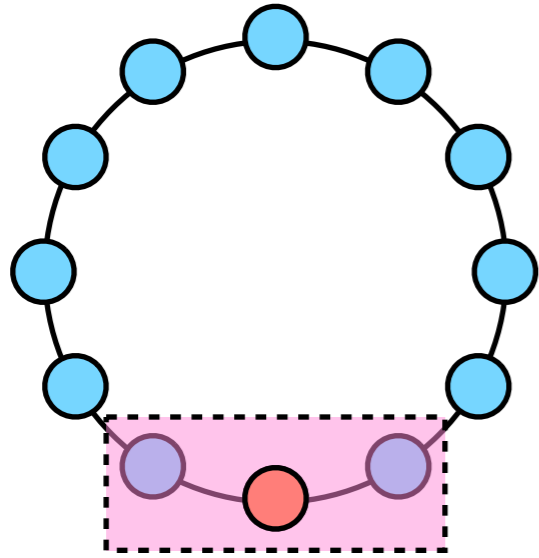
implicit data structure:

$$f : \binom{[N]}{n} \rightarrow [n!]$$

$$N \geq R_n(n!; \underbrace{2n, \dots, 2n}_{n!}) \quad \text{or equivalently} \quad N \rightarrow \underbrace{(2n, \dots, 2n)_{n!}}^n$$

$\exists X \subseteq [N], |X| \geq 2n, \binom{X}{n}$  monochromatic  $\rightarrow \geq \log n$  accesses

# Local Computation



distributed computing in a ring

$n$  nodes, ID from  $[n]$

**maximal** independent set (MIS)

for **any** input ring,  
**locally** compute the MIS

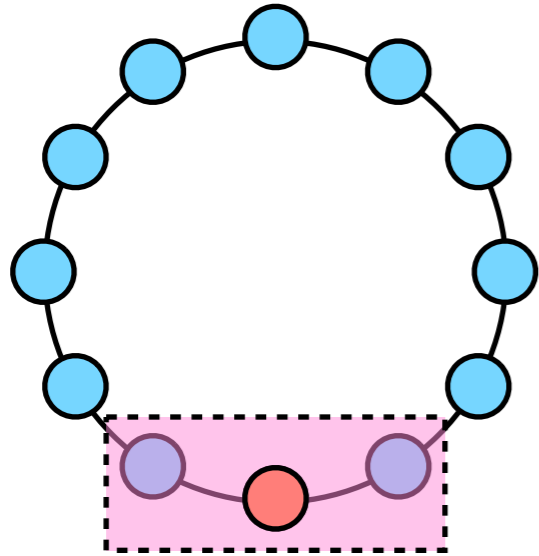
$t$ -local algorithm:

$$\mathcal{L} : \{t\text{-permutations of } [n]\} \rightarrow \{0, 1\}$$

➔  $f : \binom{[n]}{t} \rightarrow \{0, 1\}$

$$f(\{a_1, \dots, a_t\}) = \mathcal{L}(a_1, \dots, a_t)$$
$$a_1 < a_2 < \dots < a_t$$

# Local Computation



distributed computing in a ring

$n$  nodes, ID from  $[n]$

**maximal** independent set (MIS)

$t$ -local algorithm:

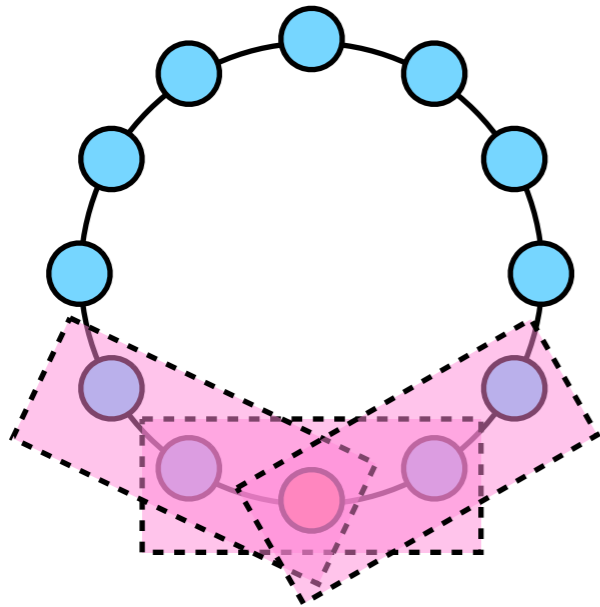
$$f : \binom{[n]}{t} \rightarrow \{0, 1\}$$

$$n \geq R_t(2; t + 2, t + 2)$$

$$\exists \text{ a monochromatic } \binom{S}{t} \quad |S| = t + 2$$

$$S = \{a_1, \dots, a_t, a_{t+1}, a_{t+2}\} \quad a_1 < \dots < a_t < a_{t+1} < a_{t+2}$$

# Local Computation



distributed computing in a ring

$n$  nodes, ID from  $[n]$

**maximal** independent set (MIS)

$t$ -local algorithm:

$$\mathcal{L} : \{t\text{-permutations of } [n]\} \rightarrow \{0, 1\}$$

$$n \geq R_t(2; t+2, t+2) \implies \exists S = \{a_1, \dots, a_t, a_{t+1}, a_{t+2}\}$$

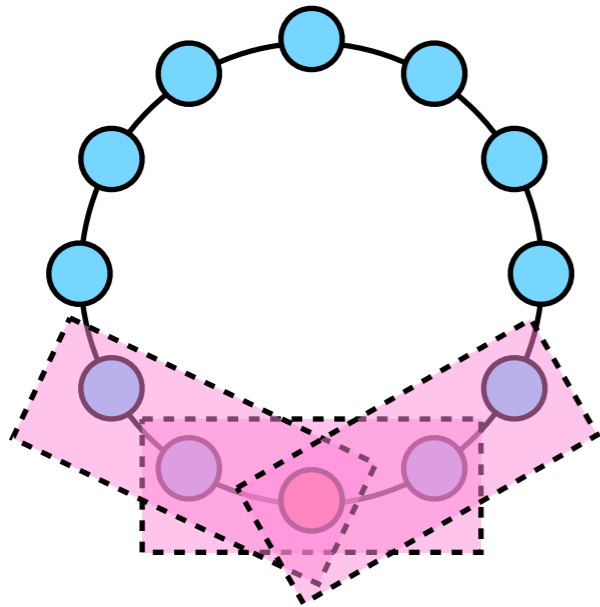
$$\mathcal{L}(a_1, \dots, a_t) = \mathcal{L}(a_2, \dots, a_{t+1}) = \mathcal{L}(a_3, \dots, a_{t+2})$$

construct a bad ring starting with

$$(a_1, a_2, \dots, a_t, a_{t+1}, a_{t+2})$$

**Contradiction!**

# Local Computation



distributed computing in a ring

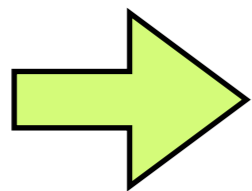
$n$  nodes, ID from  $[n]$

**maximal** independent set (MIS)

$t$ -local algorithm:

$$\mathcal{L} : \{t\text{-permutations of } [n]\} \rightarrow \{0, 1\}$$

$$n < R_t(2; t+2, t+2) \leq \underbrace{2^{2^{\cdot 2^{ct}}}}_t$$



$$t = \Omega(\log^* n)$$

# Ramsey Theory

# Arithmetic Progression

2-coloring of 1 to 12

1 2 3 4 5 6 7 8 9 10 11 12

monochromatic arithmetic progression

Can you give a 2-coloring of  $\mathbb{N}$ , no infinite monochromatic arithmetic progression?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 ...

# Van der Waerden Theorem

$W(r,k) \triangleq$  the smallest integer satisfying:

if  $n \geq W(r,k)$ , for any  $r$ -coloring of  $[n]$ , there exists a **monochromatic arithmetic progression** of length  $k$

## **VdW Theorem**

(Van der Waerden 1927)

$W(r,k)$  is finite.



Bartel Leendert van der Waerden  
(1903-1996)

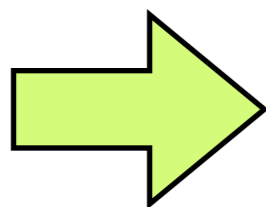
# Van der Waerden Theorem

$W(r,k) \triangleq$  the smallest integer satisfying:

if  $n \geq W(r,k)$ , for any  $r$ -coloring of  $[n]$ , there exists a **monochromatic arithmetic progression** of length  $k$

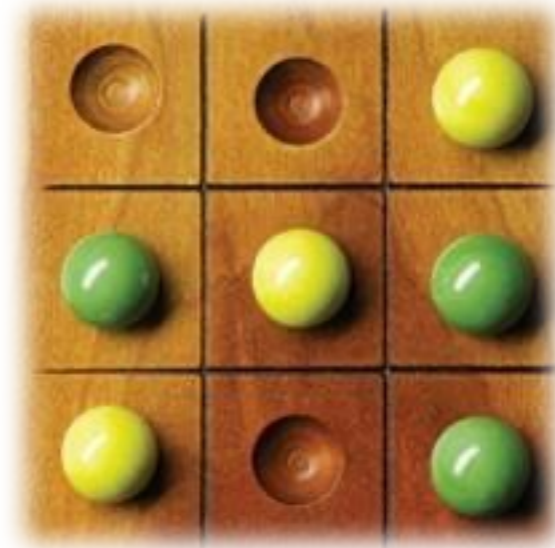
**VdW Theorem**  $W(r,k)$  is finite.

$$\mathbb{N} = C_1 \cup C_2 \cup \dots \cup C_r$$



some  $C_i$  contains **arbitrarily** long **arithmetic progression**.

# Hales-Jewett Theorem



# Hales-Jewett Theorem

$[k] = \{1, \dots, k\}$ : an alphabet of  $k$  symbols

$[k]^n$ :  $n$ -dimensional discrete cube

**combinatorial line:**  $L_\tau = \{\tau(1), \tau(2), \dots, \tau(k)\}$

$\tau \in ([k] \cup \{\star\})^n$      $\tau$  contains “ $\star$ ”

$\forall i \in [k], \tau(i) =$  replacing  $\star$  by  $i$  in  $\tau$

$\tau = 12 \star 3 \star 2$   
 $k = 4$

$L_\tau = \{ 12\color{red}13\color{red}12$   
 $12\color{red}23\color{red}22$   
 $12\color{red}33\color{red}32$   
 $12\color{red}43\color{red}42 \}$

# Hales-Jewett Theorem

$[k] = \{1, \dots, k\}$ : an alphabet of  $k$  symbols

$[k]^n$ :  $n$ -dimensional discrete cube

combinatorial line:

$HJ(r, k) \triangleq$  the smallest integer satisfying:

If  $n \geq HJ(r, k)$ , for every  $r$ -coloring of the cube  $[k]^n$ , there exists a monochromatic combinatorial line.

**Hales-Jewett Theorem**  $HJ(r, k)$  is finite.

# HJT $\implies$ VdW

If  $n \geq HJ(r, k)$ , for every  $r$ -coloring of the cube  $[k]^n$ , there exists a **monochromatic combinatorial line**.

**reduction**  $\phi: [k]^n \rightarrow [N]$

$$\forall x \in [k]^n \quad \phi(x) = x_1 + x_2 + \dots + x_n$$

**combinatorial line**  $L_\tau = \{\tau(1), \dots, \tau(k)\}$

**arithmetic progression**  $\{\phi(\tau(1)), \dots, \phi(\tau(k))\}$

$$\tau = 12 \star 3 \star 2$$

$$L_\tau = \{ 12\color{red}13\color{red}12, 12\color{red}23\color{red}22 \\ 12\color{red}33\color{red}32, 12\color{red}43\color{red}42 \}$$

if  $N \geq W(r, k)$ , for any  $r$ -coloring of  $[N]$ , there exists a **monochromatic arithmetic progression** of length  $k$

# HJT $\implies$ VdW

If  $n \geq HJ(r, k)$ , for every  $r$ -coloring of the cube  $[k]^n$ , there exists a **monochromatic combinatorial line**.

**reduction**  $\phi: [k]^n \rightarrow [N]$

$$\forall x \in [k]^n \quad \phi(x) = x_1 + x_2 + \dots + x_n$$

**combinatorial line**  $L_\tau = \{\tau(1), \dots, \tau(k)\}$

**arithmetic progression**  $\{\phi(\tau(1)), \dots, \phi(\tau(k))\}$

$$f: [N] \rightarrow [r] \quad \implies \quad f': [k]^n \rightarrow [r]$$
$$f'(x) = f(\phi(x))$$

if  $N \geq W(r, k)$ , for any  $r$ -coloring of  $[N]$ , there exists a **monochromatic arithmetic progression** of length  $k$

# Szemerédi's Regularity Lemma

# History

- Ramsey theory
- Van der Waerden's theorem (1927)
- Erdős-Turán conjecture (1936)
- Szemerédi's theorem (1969, 1975)
  - Szemerédi regularity lemma (1975, 1978)

# Arithmetic Progression

2-coloring of 1 to 12

1 2 3 4 5 6 7 8 9 10 11 12

monochromatic arithmetic progression

Can you give a 2-coloring of  $\mathbb{N}$ , no infinite monochromatic arithmetic progression?

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 ...

# Van der Waerden Theorem

$W(r,k) \triangleq$  the smallest integer satisfying:

if  $n \geq W(r,k)$ , for any  $r$ -coloring of  $[n]$ , there exists a **monochromatic arithmetic progression** of length  $k$

## **VdW Theorem**

(Van der Waerden 1927)

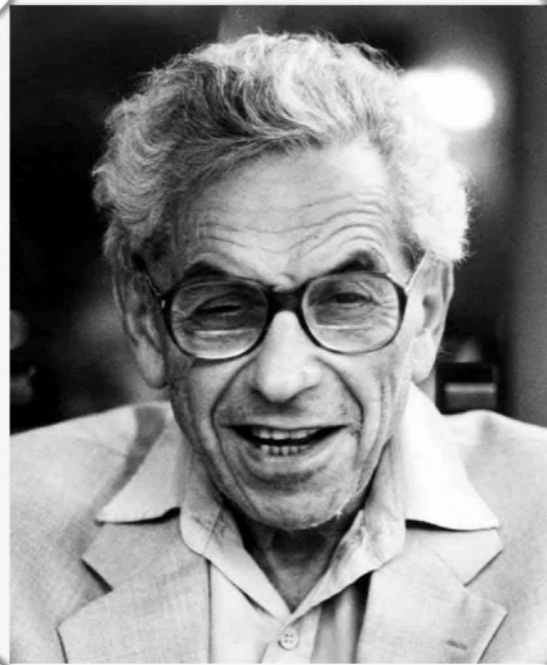
$W(r,k)$  is finite.



Bartel Leendert van der Waerden  
(1903-1996)

# Erdős-Turán Conjecture

if  $n \geq W(r, k)$ , for any  $r$ -coloring of  $[n]$ , there exists a **monochromatic arithmetic progression** of length  $k$



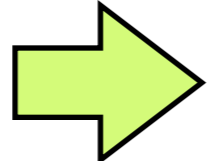
Erdős-Turán (1936): **Coloring is not essential!**

Every **dense enough** integer set has long arithmetic progression.

# Erdős-Turán Conjecture

if  $n \geq W(r, k)$ , for any  $r$ -coloring of  $[n]$ , there exists a **monochromatic arithmetic progression** of length  $k$

$\forall$  density  $0 < \delta < 1$ ,  $\forall$  integer  $k$ ,  $\exists N(\delta, k)$

$n \geq N(\delta, k)$  

every  $\delta n$ -subset  $S$  of  $[n]$  contains a length- $k$  arithmetic progression

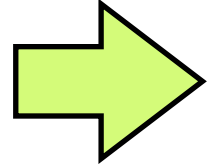
Erdős-Turán (1936): **Coloring is not essential!**

Every **dense enough** integer set has long arithmetic progression.

# Szemerédi's Theorem

**Theorem** (Szemerédi 1975)

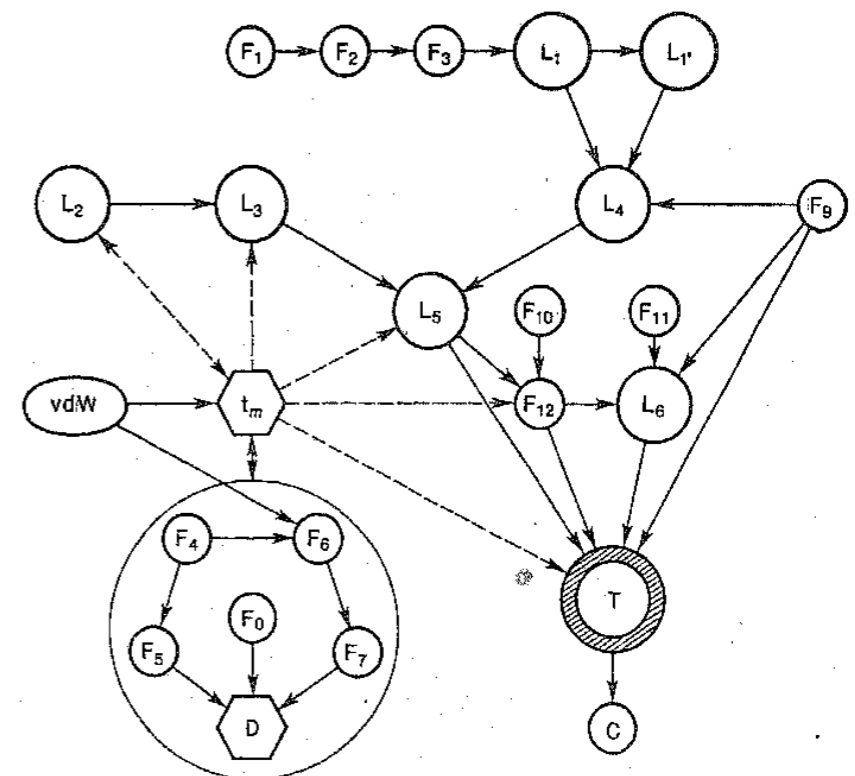
$\forall$  density  $0 < \delta < 1$ ,  $\forall$  integer  $k$ ,  $\exists N(\delta, k)$

$n \geq N(\delta, k)$  

every  $\delta n$ -subset  $S$  of  $[n]$  contains a length- $k$  arithmetic progression

key ingredient of the proof:

Szemerédi Regularity Lemma



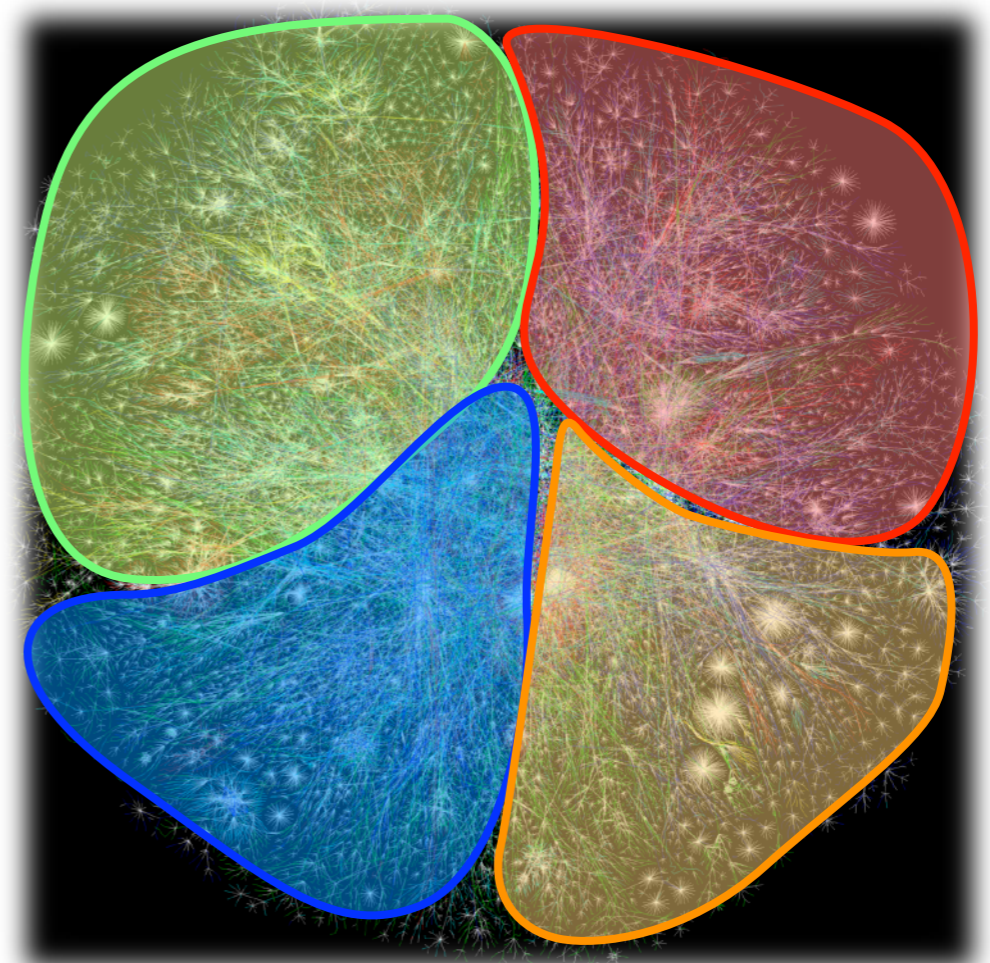
# Szemerédi Regularity Lemma

Every large enough graph can be **approximated** by a simple object of **constant complexity**.

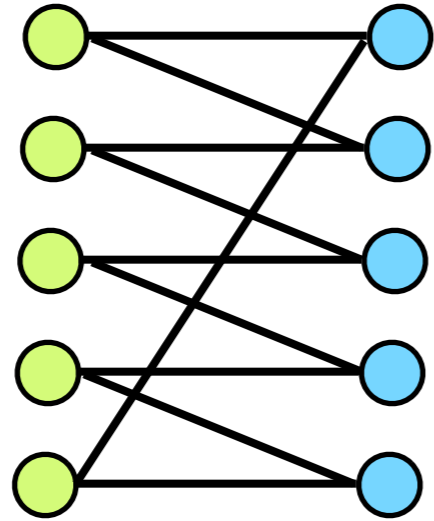
$\forall$  sufficient large graph  
can be partitioned into  
**constant** equal-sized parts

$$V_1, V_2, \dots, V_k$$

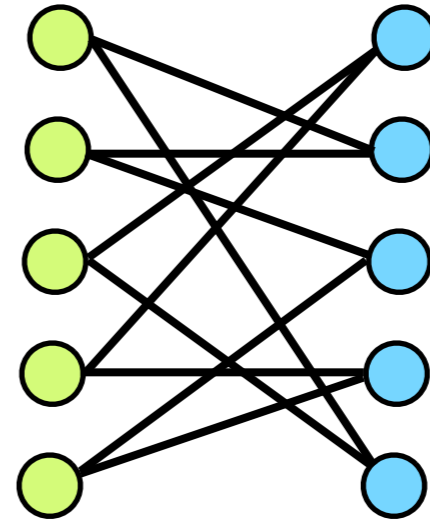
**most  $(V_i, V_j)$  look like  
random bipartite graphs**



# Regularity

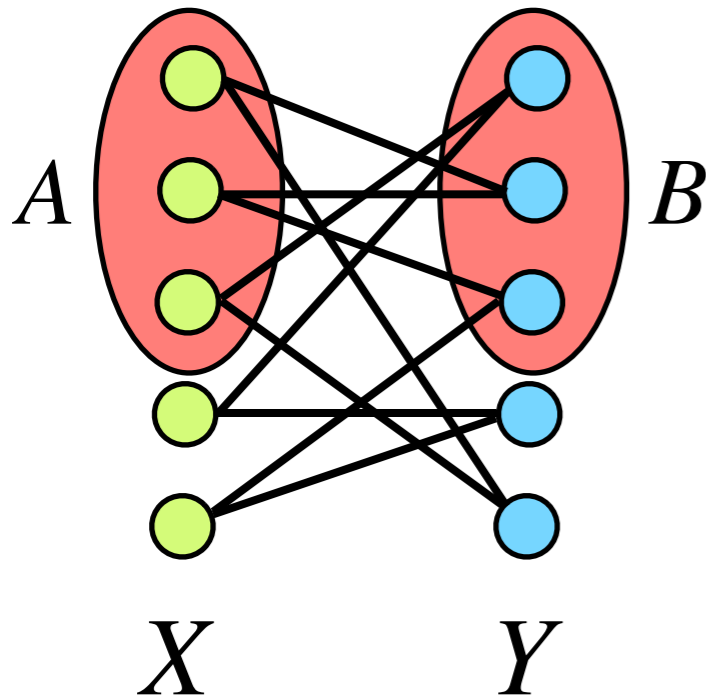


not random



random or is it?

# Regularity



$$E(X, Y) = \{uv \mid u \in X, v \in Y\}$$

density:

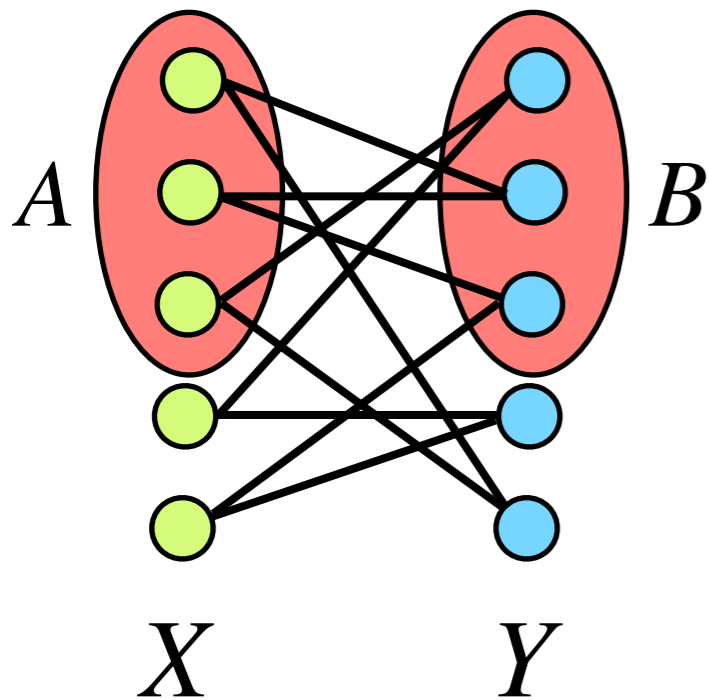
$$d(X, Y) = \frac{E(X, Y)}{|X||Y|}$$

in a **random** graph:  $\forall A \subseteq X, B \subseteq Y$

$$\mathbf{E}[E(A, B)] = |A||B| \cdot d(X, Y)$$

$$\mathbf{E}[d(A, B)] = d(X, Y)$$

# Regularity



$$E(X, Y) = \{uv \mid u \in X, v \in Y\}$$

density:

$$d(X, Y) = \frac{E(X, Y)}{|X||Y|}$$

$\epsilon$ -regularity:  $0 < \epsilon < 1$

$$\forall A \subseteq X, B \subseteq Y \text{ with } |A| \geq \epsilon|X| \text{ and } |B| \geq \epsilon|Y|$$

$$|d(A, B) - d(X, Y)| \leq \epsilon$$

regularity = uniformity

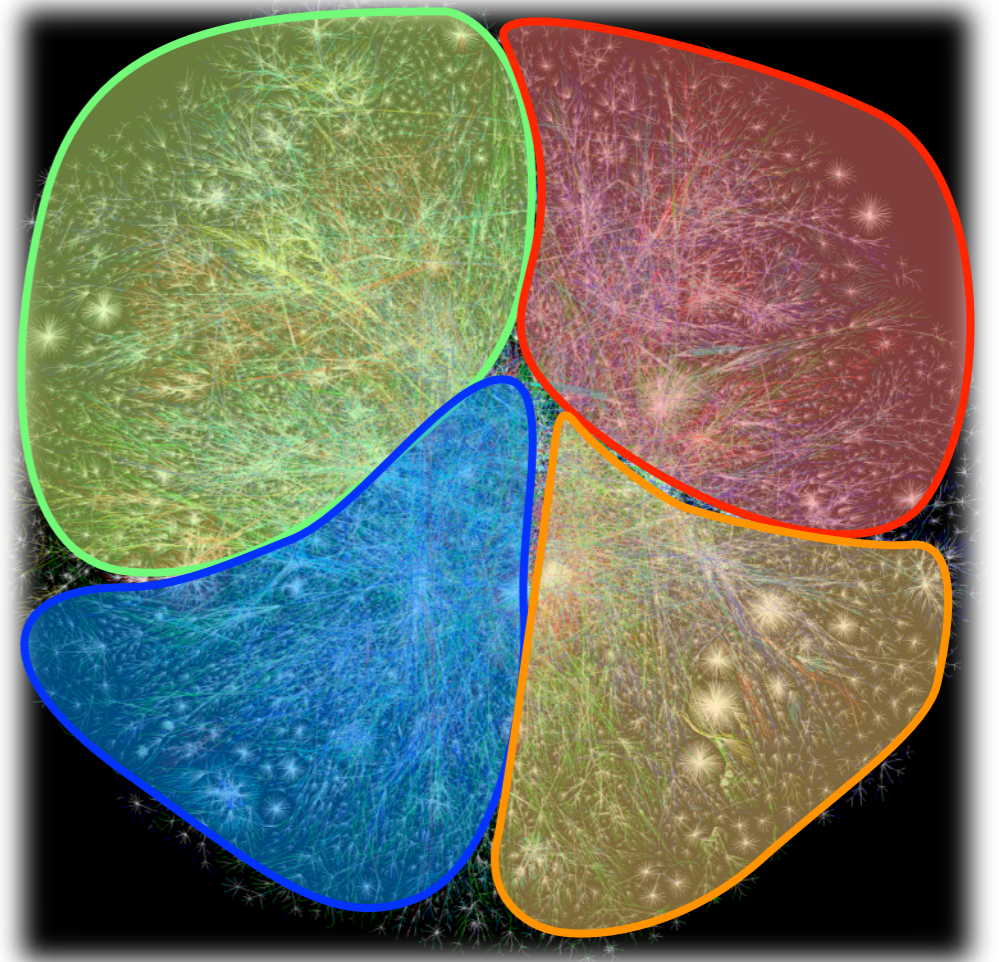
# Regular Partitions

$$G(V, E)$$

$\epsilon$ -regular partition:

$\{R, V_1, V_2, \dots, V_k\}$  of  $V$

1.  $|R| \leq \epsilon|V|$ ; **garbage bin**
2.  $|V_1| = |V_2| = \dots = |V_k|$ ;
3. all but  $\epsilon k^2$  pairs of  $(V_i, V_j)$  are  $\epsilon$ -regular.



## Theorem (Szemerédi 1978)

$\forall 0 < \epsilon < 1, m \geq 1, \exists M$  and  $N$   
every  $G(V, E)$  with  $|V| \geq N$  has an  $\epsilon$ -partition  
 $\{R, V_1, \dots, V_k\}$  with  $m \leq k \leq M$ .

every large enough graph can be partitioned  
into bounded # of equal-sized parts that  
interact **mostly** in a **regular** way

one may pick the error  $\epsilon$  and the complexity  $m$

Rough structure result for **all** graphs.

One of the most powerful tool in graph theory.

# Proof of Regularity Lemma

## Theorem (Szemerédi 1978)

$\forall 0 < \epsilon < 1, m \geq 1, \exists M$  and  $N$   
every  $G(V, E)$  with  $|V| \geq N$  has an  $\epsilon$ -partition  
 $\{R, V_1, \dots, V_k\}$  with  $m \leq k \leq M$ .

classic proof: potential function + regularization

mean square density partition  $P = \{R, V_1, \dots, V_k\}$

$$q(P) = \sum_{i,j} p_i p_j d^2(V_i, V_j) \quad p_i = \frac{|V_i|}{|V|}$$

$$q(P) \in [0, 1] \quad q(P') \geq q(P) \quad \text{if } P' \text{ refines } P$$

# Proof of Regularity Lemma

## Theorem (Szemerédi 1978)

$\forall 0 < \epsilon < 1, m \geq 1, \exists M$  and  $N$   
every  $G(V, E)$  with  $|V| \geq N$  has an  $\epsilon$ -partition  
 $\{R, V_1, \dots, V_k\}$  with  $m \leq k \leq M$ .

classic proof: potential function + regularization

mean square density partition  $P = \{R, V_1, \dots, V_k\}$

$q(P) \in [0, 1]$   $q(P') \geq q(P)$  if  $P'$  refines  $P$

if  $P$  is not an  $\epsilon$ -regular partition

$\exists$  a refinement  $P'$  of  $P$  to  $\leq k2^k$  parts, s.t.  $q(P') \geq q(P) + c\epsilon^5$

# Proof of Regularity Lemma

## **Theorem (Szemerédi 1978)**

$\forall 0 < \epsilon < 1, m \geq 1, \exists M$  and  $N$   
every  $G(V, E)$  with  $|V| \geq N$  has an  $\epsilon$ -partition  
 $\{R, V_1, \dots, V_k\}$  with  $m \leq k \leq M$ .

a new proof by Tao (2007):

structure + pseudorandomness

## Theorem (Szemerédi 1978)

$\forall 0 < \epsilon < 1, m \geq 1, \exists M$  and  $N$

every  $G(V, E)$  with  $|V| \geq N$  has a partition of  $V$  into  $R, V_1, \dots, V_k$  with  $m \leq k \leq M$ , satisfying:

1.  $|R| \leq \epsilon|V|$ ;
2.  $|V_1| = |V_2| = \dots = |V_k|$ ;
3. all but  $\epsilon k^2$  pairs of  $(V_i, V_j)$  are  $\epsilon$ -regular.

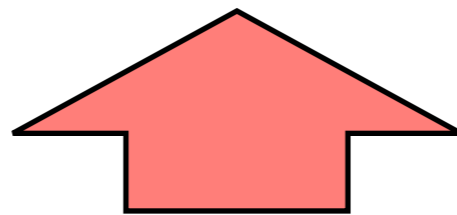
every large enough graph can be partitioned into bounded # of equal-sized parts that interact **mostly** in a **regular** way

one may pick the error  $\epsilon$  and the complexity  $m$

# Applying Regularity Lemma

**Goal:**

a theorem holding for  
all large enough graphs



the theorem holds  
for random graphs

every large enough  
graph looks random  
in some way

# Extremal Graph Theory

Fix a graph  $H$ .

$$\text{ex}(n, H)$$

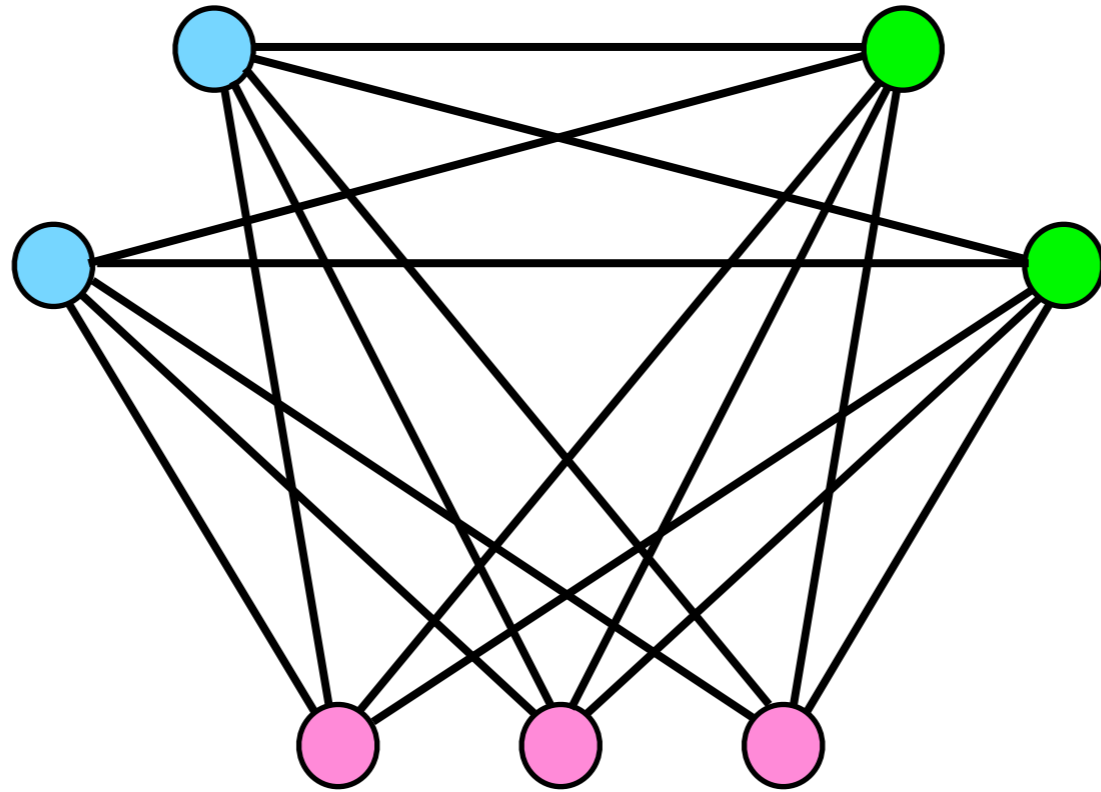
largest possible number of edges  
of  $G \not\supseteq H$  on  $n$  vertices

$$\text{ex}(n, H) = \max_{\substack{G \not\supseteq H \\ |V(G)|=n}} |E(G)|$$

$|E| > \text{ex}(n, H)$  forces subgraph  $H$  in  $G([n], E)$

# Complete multipartite graph $K_{n_1, n_2, \dots, n_r}$

$K_{2,2,3}$



Turán graph  $T(n, r)$

$$T(n, r) = K_{n_1, n_2, \dots, n_r}$$

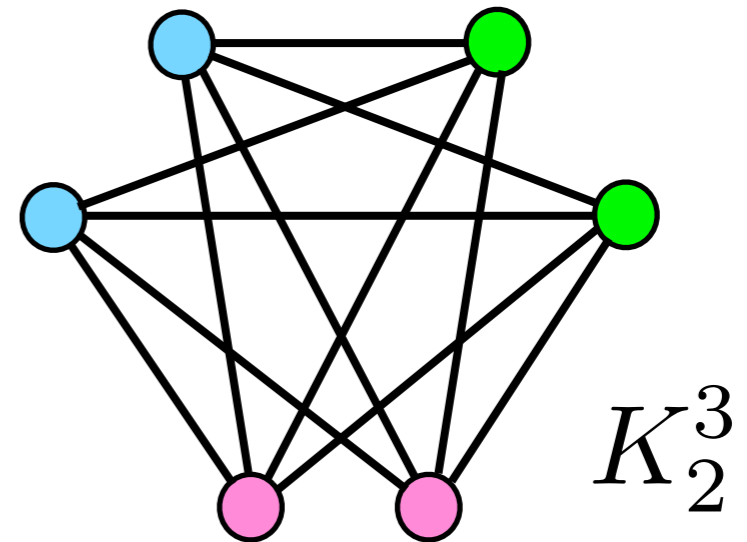
$$n_1 + n_2 + \dots + n_r = n \quad n_i \in \left\{ \left\lfloor \frac{n}{r} \right\rfloor, \left\lceil \frac{n}{r} \right\rceil \right\}$$

# Erdős–Stone theorem

(Fundamental theorem of extremal graph theory)

$$K_s^r = K_{\underbrace{s, s, \dots, s}_r} = T(rs, r)$$

complete  $r$ -partite graph  
with  $s$  vertices in each part



Theorem (Erdős–Stone 1946)

$$\text{ex}(n, K_s^r) = \left( \frac{r-2}{2(r-1)} + o(1) \right) n^2$$

## Theorem (Erdős–Stone 1946)

$$\text{ex}(n, K_s^r) = \left( \frac{r-2}{2(r-1)} + o(1) \right) n^2$$

$\text{ex}(n, H) / \binom{n}{2}$  **extremal density** of subgraph  $H$

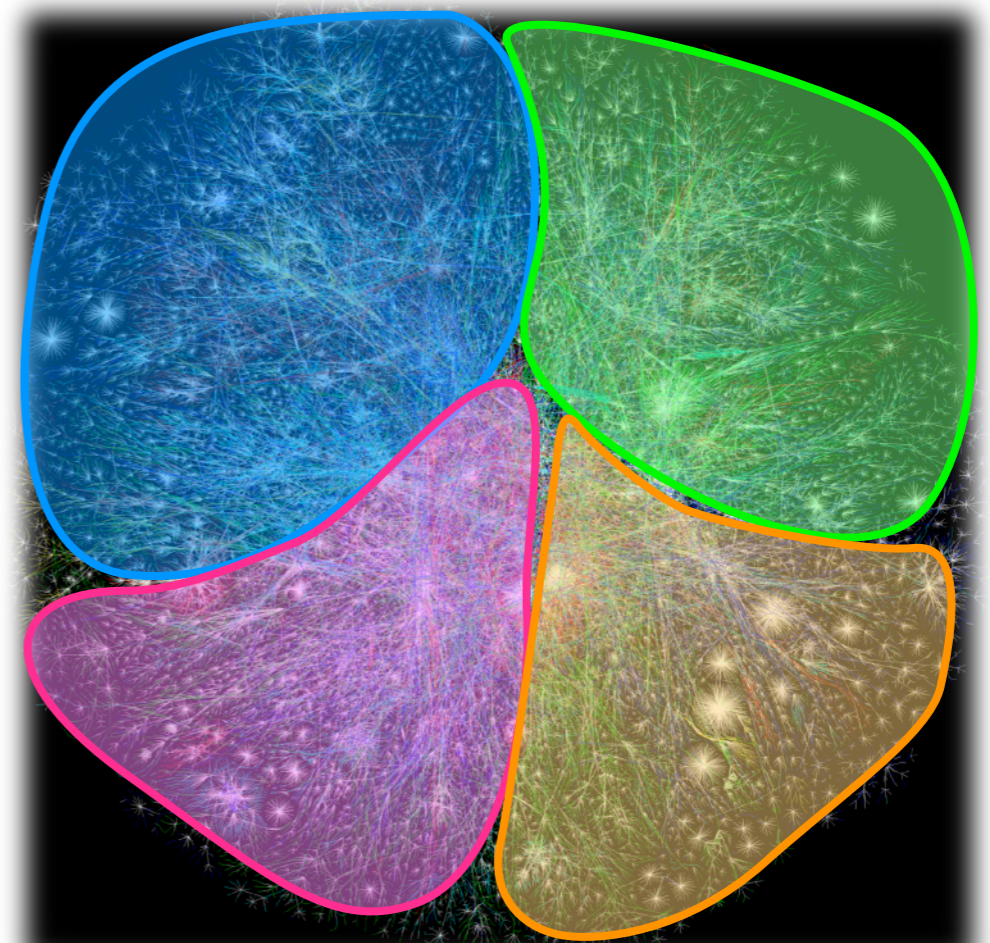
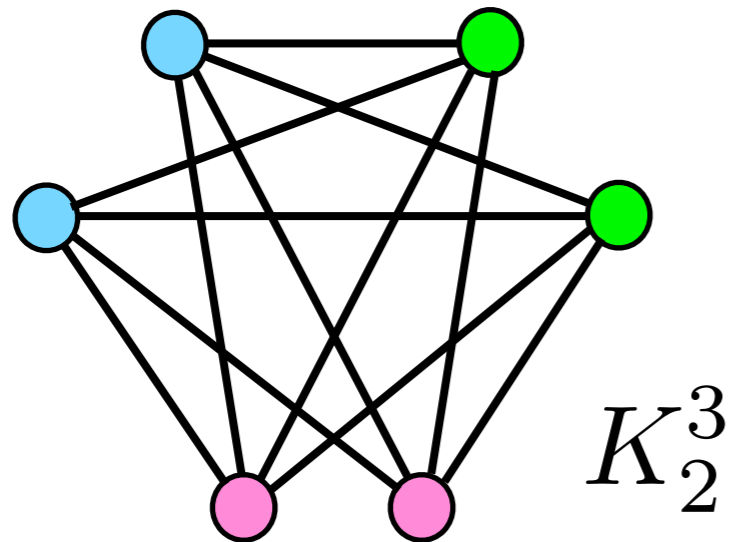
## Corollary

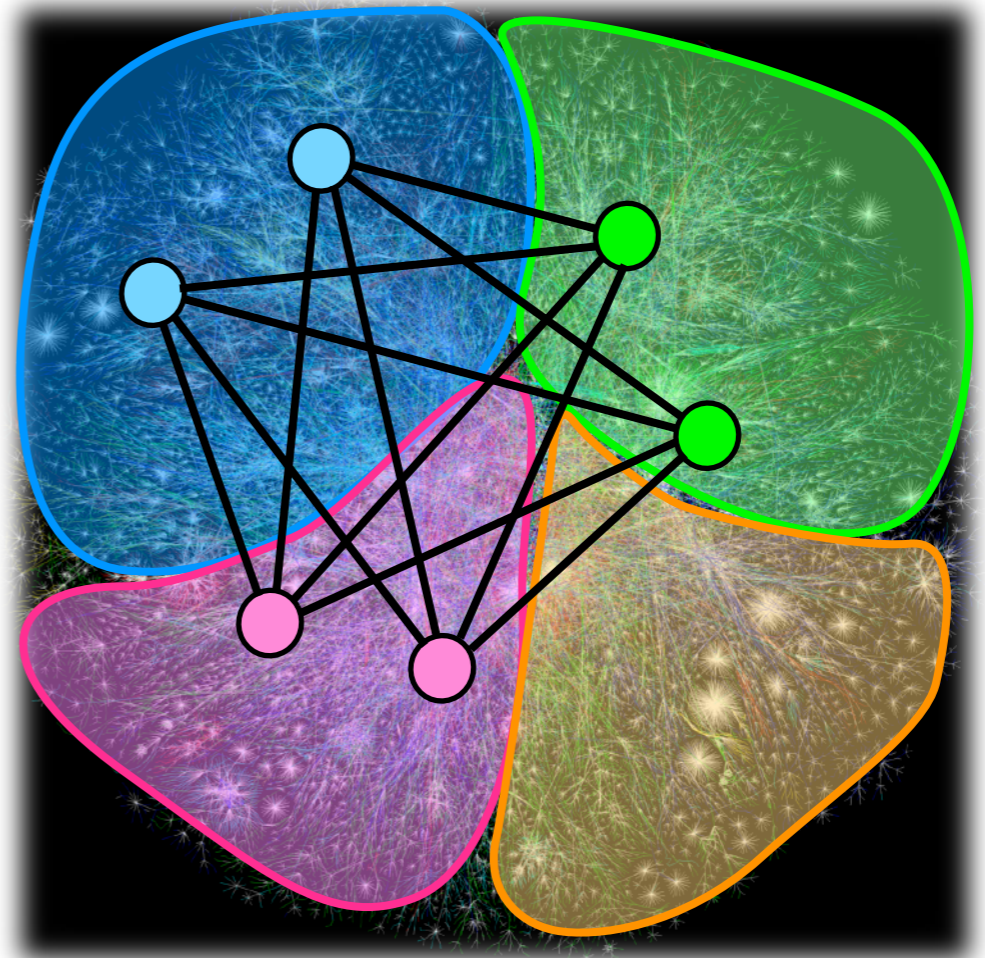
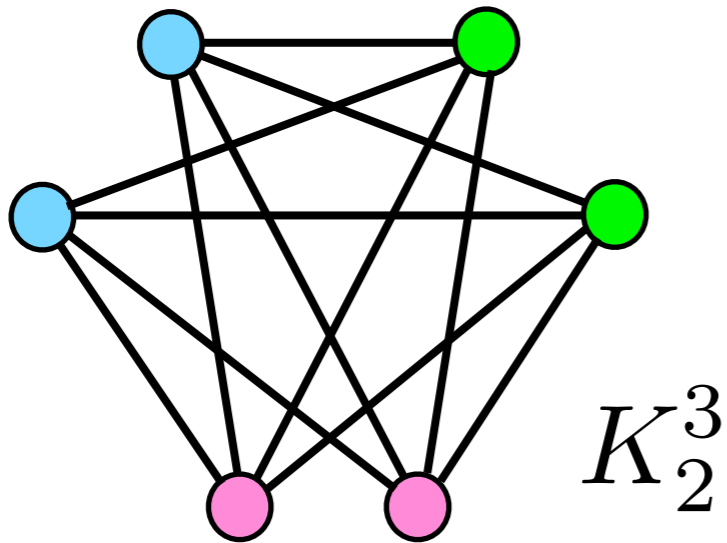
For any nonempty graph  $H$

$$\lim_{n \rightarrow \infty} \frac{\text{ex}(n, H)}{\binom{n}{2}} = \frac{\chi(H) - 2}{\chi(H) - 1}$$

# Theorem (Erdős–Stone 1946)

$$\text{ex}(n, K_s^r) = \left( \frac{r-2}{2(r-1)} + o(1) \right) n^2$$

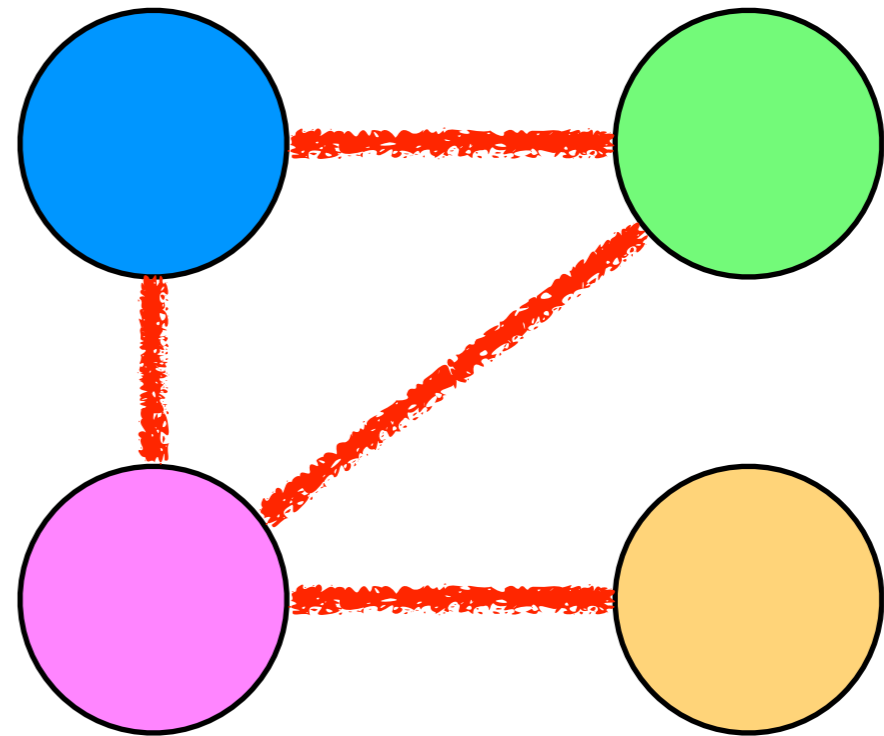




reduced graph  
(cluster graph):

vertex  $\Leftrightarrow$  part

edge  $\Leftrightarrow$  positive density  
 $\varepsilon$ -regular



with enough edge density  $\exists K_r$  in the reduced graph

# Applying Regularity Lemma

certain edge density forces subgraph  $H$   
in large enough graphs

- Edge density of  $G$  implies edge density of the reduced graph  $G^R$ .
- Apply extremal argument to the reduced graph  $G^R$  to locate a subgraph  $H^R$ .
- Embed  $H$  in  $H^R$  with guaranteed density and regularity.