The fundamental idea of the **probabilistic method** is to prove the existence of desired structures by showing the *probability* of their existence to be positive. More formally, we generally try to express the existence of desired objects as a "good" event $\overline{\bigcup_{i\in I} A_i}$, where $A_i, i \in I$ is a family of "bad" events. Then by proving $\sum_{i\in I} \mathbb{P}(A_i) < 1$, we get

$$\mathbb{P}(\bigcup_{i\in I} A_i) \overset{\text{union bound}}{\leq} \sum_{i\in I} \mathbb{P}(A_i) < 1 \Rightarrow \mathbb{P}(\overline{\bigcup_{i\in I} A_i}) = 1 - \mathbb{P}(\bigcup_{i\in I} A_i) > 0 \tag{1}$$

and hence the good event $\overline{\bigcup_{i\in I} A_i}$ is non-empty.

*Remark* (Drawback of the Probabilistic Method). Note that probabilistic arguments generally do **not** yield explicit constructions.

*Remark* (Philosophy behind the Probabilistic Method). In discrete settings, probabilistic arguments can generally be replaced by counting arguments. However, the probabilistic method is often more elegant or simple. For example, similarly, geometric arguments can often be replaced by analytic ones (e.g. calculating the intersection of lines as functions rather than talking about geometric objects).

*Example* 0.1 (Ramsey numbers). For $n \in \mathbb{N}$, the Ramsey numbers are defined as

$$R(n) := \min\{N \in \mathbb{N} \mid \text{any 2-coloring of the edges of } K_N \text{ contains a monochromatic } K_n\}$$

It is known that $R(n) \leq 4^n$ (and, since 2024, that $R(n) \leq 3.8^n$).

A trivial lower bound $(n-1)^2 < R(n)$ can be shown via the following explicit construction of a 2-coloring of $K_{(n-1)^2}$ that contains no monochromatic $K_n$. **TODO** image

Proposition 0.2 uses the probabilistic method to show a more involved lower bound. The proof is structured as follows:

- in any graph $K_m$ with $m$ smaller than the lower bound
- there exists a 2-coloring of the edges without a monochromatic $K_n$
- because the probability that such a 2-coloring randomly occurs is positive

**Proposition 0.2** (Erdös 1947)**.** *If $\binom{m}{n} \cdot 2^{1-\binom{m}{2}} < 1$, then $R(n) > m$.*

*Proof.* Assume $N \leq m$. We want to show that there exists a 2-coloring of the edges of $K_N$ that contains a mono. $K_n$. To apply the probabilistic method, we choose a **random** edge 2-coloring, i.e. each edge is colored red or blue independently with probability $1/2$.

For $S \in \binom{[N]}{n}$, the **bad events** are

$$A_S := \text{the subgraph of } K_N \text{ induced by } S \text{ is a monochromatic } K_n$$

Then $\mathbb{P}(A_S) = 2 \cdot 2^{-\binom{n}{2}} = 2^{1-\binom{n}{2}}$ because $S$ can be either red or blue (the first factor of 2) and the probability that all $\binom{N}{n}$ edges of $S$ are of the same fixed color is $2^{-\binom{n}{2}}$.

Hence $\sum_{S\subseteq\binom{[N]}{n}} \mathbb{P}(A_S) = \binom{N}{n} \cdot 2^{1-\binom{n}{2}} < 1$ and $\overline{\bigcup_{S\subseteq\binom{[N]}{n}} A_S}$ is the **good event** that no copy $S$ of $K_n$ in $K_N$ is monochromatic. Applying Equation (1) then proves the proposition. $\square$

We can improve the lower bound from Proposition 0.2 using the same proof idea, but better approximations of the probabilities.

**Corollary 0.3** (Improved Lower Bounds for $R(n)$)**.** *It is* $R(n) > \frac{n}{\sqrt{2}e} \cdot \sqrt{2}^n$. *(A more recent improvement of this result yields* $R(n) > \frac{n}{e} \cdot \sqrt{2}^{n+1}$*.)*

*Proof.* Let $N \leq m := \lfloor \frac{n}{e} \cdot \sqrt{2}^{n+1} \rfloor$. Then analogous to the proof of Proposition 0.2, the probability that a random edge 2-coloring of $K_N$ yields at least one monochromatic copy of $K_n$ is

$$\binom{N}{n} \cdot 2^{1-\binom{n}{2}} < \frac{N^n}{n!} \cdot 2^{1-\binom{n}{2}} \overset{\text{Stirling's formula}}{<<} N^n \cdot (\frac{e}{n})^n \cdot 2^{-\binom{n}{2}} = 1$$

where Stirling's formula is $n! \sim \sqrt{2\pi n}(\frac{n}{e})^n >> 2(\frac{e}{n})^n$. □

*Remark* ("Construction" using the Probabilistic Method)*.* While probabilistic arguments do not yield explicit constructions, they do yield insight into the likelihood of randomly constructing a desired object. For instance, Corollary 0.3 shows that the probability of randomly constructing an edge 2-coloring of $K_N, N \leq \frac{n}{\sqrt{2}e} \cdot \sqrt{2}^n$ that does not contain any monochromatic copy of $K_n$ is close to 1. In other words, the probability of the bad events $A_S$ is very small, so we can construct a good event at random.

*Example* 0.4 (Geometric Example(?), Dr. Arsenii Sagdeev)*.* For $v \in \{0,1,2\}^n$, let

$$T(v) := v + \{0,1\}^n + 3\mathbb{Z}^n$$

denote the translates by steps of 3 in any direction of the $n-$dimensional integer-hypercube with lower-left corner $v$. With this, let

$$f(n) := \min\{N \in \mathbb{N} \mid \exists v_1, \ldots, v_N \in \{0,1,2\}^n \text{ s.t. } \bigcup_{i \in [N]} T(v_i) = \mathbb{Z}^n\}$$

be the minimum amount of vertices $v_i$ needed so that their $T(v_i)$ cover the whole $n-$dimensional integer lattice.

For instance, $f(1) = 2$ with $T(0) \cup T(2) = \mathbb{Z}$. Moreover, $f(2) = 3$ with $T(v_1) \cup T(v_2) \cup T(v_3) = \mathbb{Z}^2$ where $v_1 = (0,0), v_2 = (1,1), v_3 = (2,2)$. **TODO** image Note that $f(2) > 2$ since, for any $v, w \in \{0,1,2\}^2$, we get

$$|(T(v) \cup T(w)) \cap \{0,1,2\}^2| \leq 8 < 9 = |\{0,1,2\}^2|$$

so $T(v) \cup T(w)$ cannot even cover the whole square $\{0,1,2\}^2$. More generally, this argument yields a lower bound $f(n) \geq (\frac{3}{2})^n$ because for $N < (\frac{3}{2})^n$, it is

$$|(\bigcup_{i \in [N]}) \cap \{0,1,2\}^n\}| \leq N \cdot 2^n < 3^n = |\{0,1,2\}^n\}|$$

.

Proposition 0.5 provides an upper bound for $f(n)$ that can be shown using the probabilistic method.

**Proposition 0.5.** *If* $3^n \cdot (1 - \frac{2}{3}^n)^N < 1$*, then* $f(n) \leq N$.

*Proof.* Goal: To obtain an upper bound $f(n) \leq N$, we need to prove that there exist $N$ vertices $v_i \in \{0,1,2\}^n, i \in [N]$ whose translates $\bigcup_{i\in[N]} T(v_i)$ cover $\mathbb{Z}^n$. We employ the probabilistic method by choosing $N$ distinct $v_i \in \{0,1,2\}^n$ independently at random with probability $\frac{1}{3^n}$ (the $v_i$ need not be distinct).

For a fixed $w \in \{0,1,2\}^n$, let $A_{w,i} := w \notin T(v_i)$ be the event that $v_i$ is chosen s.t. $T(v_i)$ does not contain $w$. The probability $\mathbb{P}(\overline{A_{w,i}})$ that $w$ *is* covered by $T(v_i)$ is $\frac{1}{3^n} \cdot 2^n$ since, taken mod 3, each of the $n$ coordinates of $v_i$ must either equal the corresponding coordinate of $w$ or be 1 smaller, so there are $2^n$ possibilities to choose $v_i$ s.t. $T(v_i)$ contains $w$. Thus

$$\mathbb{P}(A_{w,i}) = 1 - (\frac{2}{3})^n$$

Let $A_w := \bigcap_{i\in[N]} A_{w,i}$ be the **bad event** that $w$ is not covered by any $T(v_i)$. Then, because the $v_i$ are chosen independently and hence the $A_{w,i}$ are independent for a fixed $w$:

$$\mathbb{P}(A_w) = \prod_{i\in[N]} \mathbb{P}(A_{w,i}) = (1 - (\frac{2}{3})^n)^N$$

Note that $\overline{\bigcup_{w\in\{0,1,2\}^n} A_w}$ is the **good event** that all $w \in \{0,1,2\}^n$ and thus the whole lattice $\mathbb{Z}^n$ is covered by $\bigcup_{i\in[N]} T(v_i)$ and we can hence apply Equation (1) with

$$\mathbb{P}(\bigcup_{w\in\{0,1,2\}^n} A_w) \leq \sum_{w\in\{0,1,2\}^n} \mathbb{P}(A_w) = 3^n \cdot (1 - (\frac{2}{3})^n)^N < 1$$

□

In analogy with Corollary 0.3, the upper bound from Proposition 0.5 can be improved using better approximations of the probability of bad events.

**Corollary 0.6.** *It is* $f(n) \leq \ln(3) \cdot n \cdot (\frac{3}{2})^n$.

*Proof.* Let $N := \lfloor \ln(3) \cdot n \cdot (\frac{3}{2})^n \rfloor$. Then the probability $3^n \cdot (1 - (\frac{2}{3})^n)^N$ of the bad events in the above proof is bounded as follows:

$$3^n \cdot (1 - (\frac{2}{3})^n)^N \overset{1-x<e^{-x}}{<} 3^n \cdot e^{-(\frac{2}{3})^n \cdot N} = 1$$

□

*Remark* (Open Question about Example 0.4). The above results give $\Omega(1) = \frac{f(n)}{(3/2)^n} = O(n)$, but it is still open whether the fraction is constant or grows linearly in $n$.

*Example* 0.7 (Tournaments). A tournament is a directed complete graph $K_n = ([n], E), E \subseteq [n] \times [n], n \in \mathbb{N}$. For an edge $(v,w) \in E$, we also say "v dominates w" or "v wins against w" and call $v \in [n]$ a "player" and $(v,w)$ a "game" in the tournament. For fixed edge set $E$ and $k \in \mathbb{N}$, we define the event

$$S_k^E := \forall A \in \binom{[n]}{k} \exists v \in [n] \setminus A \text{ s.t. } v \text{ dominates every vertex from } A \text{ in ([n],E)}$$

which intuitively occurs in any tournament where any $k-$element subset $A$ of $n$ players is dominated by some other player $v$. With this, we let

$$f(k) := \min\{n \in \mathbb{N} \mid \exists \text{ tournament } ([n], E) \text{ for which } S_k^E \text{ occurs}\}$$

be the minimum size $n$ of any tournament (i.e., the number of players) $([n], E)$ where $S_k^E$ occurs. Then Proposition 0.8 yields an upper bound for $f(k)$ via the probabilistic method.

**Proposition 0.8** (Tournament Sizes, Erdös 1963)**.** *If* $\binom{n}{k} \cdot (1-2^{-k})^{n-k} < 1$*, then* $f(k) \leq n$*.*

*Proof.* To show an upper bound $f(k) \leq n$, we prove that there exists a tournament $([n], E)$ on $[n]$ satisfying $S_k^E$. However, we do not explicitly construct $E$. Rather, we employ the probabilistic method by choosing a tournament uniformly at random, i.e. each edge in the underlying complete graph $K_n$ has one of two directions independently with probability $1/2$.

For a fixed $T \in \binom{[n]}{k}$ of size $k$, let

$$A_T := \forall w \in [n] \setminus T : w \text{ does not dominate all of } T$$

be the **bad event** that player set $T$ is not dominated by any single player. Then for a fixed $w \in [n] \setminus T$, the probability of $w$ dominating $T$ is $2^{-k}$ and, correspondingly, the probability of the event $A_{w,T}$ that $w$ *does not* dominate $T$ is $\mathbb{P}(A_{w,T}) = 1 - 2^{-k}$. Hence,

$$\mathbb{P}(A_T) = \mathbb{P}\Big(\bigcap_{w \in [n] \setminus T} A_{w,T}\Big) = \prod_{w \in [n] \setminus T} \mathbb{P}(A_{w,T}) = (1-2^{-k})^{n-k}$$

is the probability of *no* $w \in [n] \setminus T$ dominating $T$ since the edge directions are chosen independently and thus the $A_{w,T}, w \in [n] \setminus T$ are independent for fixed $T$.

The **good event** is then $\overline{\bigcup_{T \in \binom{[n]}{k}} A_T}$ and we can apply Equation (1) again using

$$\mathbb{P}\Big(\bigcup_{T \in \binom{[n]}{k}} A_T\Big) \leq \sum_{\bigcup_{T \in \binom{[n]}{k}} A_T} \mathbb{P}(A_T) = \binom{n}{k}(1-2^{-k})^{n-k} < 1$$

$\square$

Again, similar to Corollary 0.3 and Corollary 0.6, the upper bound from Proposition 0.8 can be improved via better probability approximations.

**Corollary 0.9.** *It is* $f(k) \leq \ln(2) \cdot k^2 \cdot 2^k$*.*

*Proof.* **TODO** $\square$

To get the lower bound on $f(k)$ from Proposition 0.10, for every tournament smaller than the lower bound, we explicitly choose a set of $k$ players that is not dominated by any single player.

**Proposition 0.10.** *It is* $f(k) \geq 2^k$*. [For an improvement* $f(k) = \Omega(k \cdot 2^k)$*, see the problem classes.]*

*Proof.* To show a lower bound $f(k) \geq 2^k$, we prove that $S_k^E$ does not occur in any tournament $([n], E)$ with $n < 2^k$ players. Fixing such a tournament $([n], E)$, we construct a $T \in \binom{[n]}{k}$ that is not dominated by any vertex in $[n] \setminus T$ inductively as follows:

We call player $v \in [n]$ *strong* if it dominates at least $\frac{n}{2}$ of its $n-1$ games, i.e. $|\{(v,w) \in E \mid w \in [n]\}| \geq \frac{n-1}{2}$. The pigeonhole principle proves that there must be a strong player $v_1 \in [n]$ since $\sum_{v \in [n]} |\{(v,w) \in E \mid w \in [n]\}| = |E| = n \cdot \frac{n-1}{2}$. We add $v_1$ to $T$. Analogously, there must be a strong $v_2 \in V_2 := [n] \setminus (\{v_1\} \cup \{w \in [n] \mid (v_1, w) \in E\})$ w.r.t the tournament $(V_2, E \setminus \{(v,w) \in E \mid w \in [n]\})$ where $|V_2| \leq n - \frac{n-1}{2} - 1 = \frac{n-1}{2} < \frac{n}{2}$. We add $v_2$ to $T$ and repeat this step at most $k$ times (as long as there are vertices left, afterwards we can just arbitrarily fill $T$), yielding $T := \{v_1, v_2, ..., v_k\}$. This $T$ cannot be dominated by any vertex $w \in [n] \setminus T$ because $V_k$ with $|V_k| < \frac{n}{2^k} < 1$, i.e. $V_k = \emptyset$ are the only vertices that can potentially win against all of $T$. **TODO** rewrite, image $\square$

*Example* 0.11 (Hypergraphs). A hypergraph $H = (V, E)$ with $E \subseteq 2^V$ is $n-$uniform if $E \subseteq \binom{V}{n}$, i.e. all edges are of size $n$. For such a hypergraph, we define the chromatic number

$$\chi(H) := \min \text{ \#colors needed in a vertex-coloring of H s.t. no edge in E is monochromatic}$$

We call a vertex-coloring "proper" if it does not yield any monochromatic edge. With this, let

$$m(n) := \min\{|E| \mid \exists n-\text{uniform } H = (V,E) : \chi(H) > 2\}$$

be the minimum number of edges needed in an $n-$uniform hypergraph that does not have a proper vertex 2-coloring.

For instance, $m(n) > 2$ for all $n \in \mathbb{N}$ because a single edge can clearly be properly 2-colored and, for 2 distinct edges, coloring their intersection red and the (non-empty!) rest of their respective vertices blue yields a proper coloring. Moreover, for $n > 2$, it is $m(n) > 3$.

Proposition 0.12 yields a more involved lower bound on $m(n)$ using the probabilistic method.

**Proposition 0.12.** *It is* $m(n) \geq 2^{n-1}$.

*Proof.* We show that any $n-$uniform hypergraph $H = (V, E)$ with less than $2^{n-1}$ edges has a proper 2-coloring, i.e. $\chi(H) \leq 2$. We employ the probabilistic method by choosing a random 2-coloring of $H$ where each vertex is either red or blue independently with probability $1/2$. For $e \in E$, let

$$A_e := e \text{ is monochromatic}$$

be the **bad event** that $e$ is monochromatic. Then $\mathbb{P}(A_e) = 2^{1-|e|} = 2^{1-n}$ and

$$\mathbb{P}(\bigcup_{e \in E} A_e) \leq |E| \cdot 2^{1-|e|} < 2^{n-1} \cdot 2^{1-n} = 1$$

With the **good event** being $\overline{\bigcup_{e \in E} A_e}$, Equation (1) finishes the proof. $\square$