"We shall all be changed, in a moment, in the twinkling of an eye"

Paul the Apostle, 1 Corinthians 15:51–2

"Do I dare
Disturb the universe?
In a minute there is time
For decisions and revisions which a minute will reverse.
. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
I have seen the moment of my greatness flicker,
And I have seen the eternal Footman hold my coat, and snicker,
And in short, I was afraid."

T. S. Eliot, "The Love Song of J. Alfred Prufrock"

# 1 Introduction and Old Results

**Definition.** For $n \in \mathbb{N}, p \in [0,1]$, the Erdős-Rényi random graph model $\mathcal{G}(n,p)$ is a probability distribution on the set of graphs on $n$ vertices, where each of the $\binom{n}{2}$ potential edges appears independently with probability $p$.

For any graph property $P$, we denote by $\mathbb{P}_p(P)$ the probability that a graph drawn from $\mathcal{G}(n,p)$ will have the property $P$.

**Example.** Here are several theorems about random graphs having various properties:

1. (Erdős-Rényi 1960) For any $\varepsilon > 0$,

$$\mathbb{P}_p(G \text{ is connected}) \to \begin{cases} 0 & \text{if } p < (1-\varepsilon)\dfrac{\log n}{n} \\ 1 & \text{if } p > (1+\varepsilon)\dfrac{\log n}{n} \end{cases}$$

2. (Erdős-Rényi 1960)

$$\mathbb{P}_p(G \text{ contains a triangle}) \to \begin{cases} 0 & \text{if } p \ll \frac{1}{n} \\ 1 & \text{if } p \gg \frac{1}{n} \end{cases}$$

More generally, for any fixed graph $H$, we *morally* have (the truth is a bit more complicated)

$$\mathbb{P}_p(G \text{ has a copy of } H) \to \begin{cases} 0 & \text{if } p \ll n^{-|V(H)|/|E(H)|} \\ 1 & \text{if } p \gg n^{-|V(H)|/|E(H)|} \end{cases}$$

3. (Erdős-Rényi 1960)

$$\mathbb{P}_p(G \text{ is bipartite}) \to \begin{cases} 0 & \text{if } p \ll \frac{1}{n} \\ 1 & \text{if } p \gg \frac{1}{n} \end{cases}$$

4. (Erdős-Rényi 1960) For any $\varepsilon > 0$,

$$\mathbb{P}_p(G \text{ is planar}) \to \begin{cases} 1 & \text{if } p < (1-\varepsilon)\frac{1}{n} \\ 0 & \text{if } p > (1+\varepsilon)\frac{1}{n} \end{cases}$$

5. (Erdős-Rényi 1966) For any $\varepsilon > 0$,

$$\mathbb{P}_p(G \text{ has a perfect matching}) \to \begin{cases} 0 & \text{if } p < (1-\varepsilon)\dfrac{\log n}{n} \\ \\ 1 & \text{if } p > (1+\varepsilon)\dfrac{\log n}{n} \end{cases}$$

6. (Komlós-Szemerédi 1983, Bollobás 1983) For any $\varepsilon > 0$,

$$\mathbb{P}_p(G \text{ has a Hamiltonian cycle}) \to \begin{cases} 0 & \text{if } p < (1-\varepsilon)\dfrac{\log n + \log\log n}{n} \\ \\ 1 & \text{if } p > (1+\varepsilon)\dfrac{\log n + \log\log n}{n} \end{cases}$$

7. (Achlioptas-Naor 2005) For any integer $k \geq 3$ and any $\varepsilon > 0$,

$$\mathbb{P}_p(G \text{ is } k\text{-colorable}) \to \begin{cases} 1 & \text{if } p < (1-\varepsilon)\dfrac{2k\log k}{n} \\ \\ 0 & \text{if } p > (1+\varepsilon)\dfrac{2k\log k}{n} \end{cases}$$

   In fact, they proved a stronger result: for any $0 < d < \infty$ and $k_d$ the smallest integer $k$ satisfying $d < 2k\log k$, we have that

$$\chi\left(\mathcal{G}\left(n, \frac{d}{n}\right)\right) \in \{k_d, k_d+1\} \text{ with probability} \to 1 \text{ as } n \to \infty$$

There are several remarks to be made about these examples.

**Remark 1.** If we imagine, for fixed $n$, plotting $\mathbb{P}_p$ as a function of $p$, then all of these results say that as we let $n \to \infty$, our plot looks more and more like that of a step function, namely it's constant for most of the interval with a big jump somewhere in the middle. This is very reminiscient of the Kolmogorov Zero-One Law, which roughly says that in any infinite analogue of these, wherein we are looking at a property that is independent of any finite collection of edges, its plot will actually be a step function. So it's often insightful to think of these examples, and what comes next, as a finitary analogue of the Kolmogorov Zero-One Law.

**Remark 2.** Why does water always freeze at 0° C? In theory, since hydrogen bonds between molecules are being formed in a very unpredictable way, we might expect that the order in which they form could affect when crystallization takes place. Results like these, especially the ones concerning connectivity, can be viewed as a mathematical explanation of this physical fact. Moreover, the general result we'll soon state can be used to explain many such "phase transitions" in physics.

**Remark 3.** This random model turns necessary conditions into sufficient ones. Indeed, it can be shown e.g. that if $p < n/\log n$, then $\mathcal{G}(n,p)$ will have an isolated node with high probability. Having no isolated nodes is necessary for being connected, but is in general not sufficient—but in the random model it is. Similarly, a perfect matching appears once every vertex has degree $\geq 1$ and a Hamiltonian cycle appears once every vertex has degree $\geq 2$.

## 2 Friedgut-Kalai

**Definition.** Let $N = \binom{n}{2}$. Then a *graph* on $n$ vertices is a subset of $\{1, 2, \ldots, N\}$, or equivalently an element of $\{0,1\}^N$.

A *graph property* is some set $P$ of graphs on $n$ vertices, or equivalently a subset $P \subseteq \{0,1\}^N$.

**Definition.** A property is called *monotone* if it is upward closed, meaning that adding edges preserves the property. More formally, $P$ is monotone if for all $x, y \in \{0,1\}^N$ with $x \in P$ and $x \leq y$ pointwise, then $y \in P$. Examples include connectivity, non-planarity, having a Hamiltonian cycle, having chromatic number at least 17, etc. A non-example is having an Eulerian cycle.

**Definition.** A property is called *symmetric* if for all $i, j \in \{1, \ldots, N\}$ there is a permutation $\pi \in S_N$ such that $\pi(i) = j$ and $P$ is invariant under $P$ (or, more formally, the induced permutation on $\{0,1\}^N$). If you're studying a non-symmetric property, you're not doing graph theory.

**Theorem** (Friedgut-Kalai, 1996)**.** *Every symmetric monotone property has a threshold. More formally, there is an absolute constant $c$ such that for any $N \in \mathbb{N}$ and any monotone symmetric property $P \subseteq \{0,1\}^N$ with $\mathbb{P}_p(P) \geq \varepsilon$ for $0 < \varepsilon < 1/2$, we have $\mathbb{P}_q(P) > 1 - \varepsilon$ for any*

$$q \geq p + c\frac{\log(1/2\varepsilon)}{\log N}$$

*In other words, the threshold has width $\sim \log(1/2\varepsilon)/\log N$.*

**Remark 4.** This implies all the facts we saw above, although it comes with a very important caveat: the Friedgut-Kalai theorem does not help us locate where the threshold will be, it only tells us that a threshold must exist.

**Remark 5.** In addition to these applications in random graph theory, it also has many important implications in other areas of math; Erik's talk showed us applications in pure probability theory (percolation), and Kumar-Pfister and independently Kudekar-Mondelli-Šašoğlu-Urbanke (2015) used it as a key tool in proving that Reed-Muller codes achieve capacity on the BEC.

# 3 Interlude: Voting

Suppose we have $N$ voters who each cast a ballot for one of two candidates, and we have some system according to which a winner is selected. More formally, we simply have a function $f : \{0,1\}^N \to \{0,1\}$, where we think of an element of $\{0,1\}^N$ as the sequence of votes by the $N$ voters. Here are some examples:

- A democracy: $f$ is the majority function (for simplicity, say $N$ is odd)
- A dictatorship: voter $i$ decides, namely $f(x) = x_i$.
- A rigged system: in this case, $f$ would be just a constant function, and the votes don't matter.
- The electoral college: the input is partitioned into blocks, there is a majority function within each block, and the final result is a weighted majority of the winners in the blocks.
- Insanity:
$$f(x) = \sum_{i=1}^{N} x_i \mod 2$$

One natural notion is that of the *influence* of the $i$th voter, which is intuitively how likely the $i$th voter is to swing the election. In a rigged system, everyone's influence is zero. In a dictatorship, everyone has zero influence except the dictator, who has influence 1. In a democracy, everyone has a pretty small (and equal) influence (in fact, $\Theta(1/\sqrt{n})$), whereas in the electoral college system some voters have a greater influence than others (as we know very well from real life). In the insane system, everyone has influence 1, since changing your vote will definitely flip the outcome.

**Definition.** Fix $f : \{0,1\}^N \to \{0,1\}$. For any input $x \in \{0,1\}^N$ and an index $i \in \{1, \dots, N\}$, we say that $i$ is *pivotal* for $x$ if

$$f(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_N) \neq f(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_N)$$

Note that this notion depends both on the voter $i$ and on how everyone else votes, namely $x$.

For $p \in [0,1]$, we define the *influence* of voter $i$ as

$$\beta_i(f) = \mathbb{P}_p(i \text{ is pivotal})$$

In other words, we make every other voter vote for 1 with probability $p$, and then we ask for the probability that voter $i$ affects the outcome.

Intuitively, the smaller everyone's influence is, the "fairer" the voting system is. Then a natural question is how fair we can get. Of course, this notion is imperfect: in the "rigged" system, everyone has influence zero, and indeed we need to include the probability that each candidate wins.

**Theorem** (Kahn-Kalai-Linial, 1988)**.** *For any function $f : \{0,1\}^N \to \{0,1\}$ and any $p \in [0,1]$, we have*

$$\max_{1 \le i \le N} \beta_i(f) \geq cm \frac{\log N}{N}$$

*for some universal constant $c$ and*

$$m = \min\{\mathbb{P}_p(f = 0), \mathbb{P}_p(f = 1)\}$$

*namely the probability of an upset. This is tight up to constants (thanks to the "tribes" example).*

This was later generalized to an arbitrary probability spaces by Bourgain-Kahn-Kalai-Katznelson-Linial (1992). The proof of both of these uses some discrete Fourier analysis and hypercontractivity inequalities.

# 4 Proof

We will need the following simple lemma, due to Margulis (1974) and again Russo (1981).

**Lemma.** *Let $P \subseteq \{0,1\}^N$ be a monotone property. Then*

$$\frac{\mathrm{d}}{\mathrm{d}p} \mathbb{P}_p(P) = \sum_{i=1}^N \beta_i(P)$$

*where by $\beta_i(P)$, we mean $\beta_i(\chi_P)$, the influence of the characteristic function of $P$.*

*Proof sketch.* Let $\underline{p} = (p_1, \dots, p_N)$ be a vector of probabilities, and let $\mathbb{P}_{\underline{p}}(P)$ denote the probability of $P$ where the $i$th index is 1 with probability $p_i$. Then since $P$ is monotone, we have

$$\begin{aligned}
\mathbb{P}_{\underline{p}+\Delta p_i}(P) - \mathbb{P}_{\underline{p}}(P) &= \mathbb{P}(\text{increasing } p_i \text{ by } \Delta p_i \text{ moves from out of } P \text{ into } P) \\
&= \mathbb{P}(\text{increasing } p_i \text{ changes coordinate } i) \mathbb{P}(\text{chaning coordinate } i \text{ changes the outcome}) \\
&= \Delta p_i \beta_i(P)
\end{aligned}$$

Therefore,

$$\frac{\partial}{\partial p_i} \mathbb{P}_{\underline{p}}(P) = \beta_i(P)$$

So by the multivariate chain rule, setting $\underline{p} = (p, \dots, p)$, we get that

$$\frac{\mathrm{d}}{\mathrm{d}p} \mathbb{P}_p(P) = \sum_{i=1}^N \beta_i(P)$$

□

*Proof of Friedgut-Kalai.* Since $P$ is symmetric, we get that $\beta_i(P) = \beta_j(P)$ for all $i, j$. So the Margulis-Russo formula becomes, for any $r \in [0, 1]$

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{P}_r(P) = N\beta(P)$$

where $\beta(P)$ is the influence of any coordinate. For any $r$ such that $\mathbb{P}_r(P) \leq 1/2$, Kahn-Kalai-Linial tells us that

$$\beta(P) \geq c\mathbb{P}_r(P)\frac{\log N}{N}$$

since

$$m = \min(\mathbb{P}_r(P), 1 - \mathbb{P}_r(P)) = \mathbb{P}_r(P)$$

Combining these, we get

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{P}_r(P) \geq c\mathbb{P}_r(P)\log N$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}r}\log\mathbb{P}_r(P) = \frac{\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{P}_r(P)}{\mathbb{P}_r(P)} \geq \frac{c\mathbb{P}_r(P)\log N}{\mathbb{P}_r(P)} = c\log N$$

Now, suppose $p$ is such that $\mathbb{P}_p(P) \geq \varepsilon$. Set

$$q_1 = p + \frac{\log(1/2\varepsilon)}{c\log N}$$

Then integrating the above inequality from $p$ to $q_1$ gives

$$\begin{aligned}
\log\mathbb{P}_{q_1}(P) - \log\mathbb{P}_p(P) &= \int_p^{q_1}\frac{\mathrm{d}}{\mathrm{d}r}\log\mathbb{P}_r(P)\mathrm{d}r \\
&\geq \int_p^{q_1} c\log N\mathrm{d}r \\
&= (q_1 - p)c\log N \\
&= \log(1/2\varepsilon)
\end{aligned}$$

Therefore,

$$\log\mathbb{P}_{q_1}(P) \geq \log\mathbb{P}_p(P) + \log(1/2\varepsilon) \geq \log(\varepsilon) + \log(1/2\varepsilon) = \log(1/2)$$

and thus $\mathbb{P}_{q_1}(P) \geq 1/2$. In other words, by increasing $p$ a little bit, we've increased the probability of $P$ from $\varepsilon$ to $1/2$.

We now repeat this argument, with a slight variation. For $r$ such that $\mathbb{P}_r(P) \geq 1/2$, Kahn-Kalai-Linial gives us

$$\beta(P) \geq c(1 - \mathbb{P}_r(P))\frac{\log N}{N}$$

Then repeating all the above gives

$$\frac{\mathrm{d}}{\mathrm{d}r}\mathbb{P}_r(P) = N\beta(P) \geq c(1 - \mathbb{P}_r(P))\log N$$

and so

$$\frac{\mathrm{d}}{\mathrm{d}r}\log(1 - \mathbb{P}_r(P)) = \frac{\frac{\mathrm{d}}{\mathrm{d}r}(1 - \mathbb{P}_r(P))}{(1 - \mathbb{P}_r(P))} \leq -\frac{c(1 - \mathbb{P}_r(P))\log N}{(1 - \mathbb{P}_r(P))} = -c\log N$$

We set

$$q = q_1 + \frac{\log(1/2\varepsilon)}{c\log N}$$

and integrate from $q_1$ to $q$ and get

$$\begin{aligned}\log(1-\mathbb{P}_q(P)) - \log(1-\mathbb{P}_{q_1}(P)) &= \int_{q_1}^{q} \frac{\mathrm{d}}{\mathrm{d}r} \log(1-\mathbb{P}_r(P))\mathrm{d}r \\ &\le \int_{q_1}^{q} (-c\log N)\mathrm{d}r \\ &= -c\log N \frac{\log(1/2\varepsilon)}{c\log N} \\ &= -\log(1/2\varepsilon)\end{aligned}$$

Thus,

$$\log(1-\mathbb{P}_q(P)) \le \log(1-\mathbb{P}_{q_1}(P)) - \log(1/2\varepsilon) \le \log(1/2) - \log(1/2\varepsilon) = \log(\varepsilon)$$

which gives

$$\mathbb{P}_q(P) \ge 1-\varepsilon$$

as desired. □

# 5 Coarse vs. Sharp Thresholds

If we return to the set of examples at the beginning, we see that there seems to be a pretty important qualitative difference: on the one hand, the results about containing a subgraph and being bipartite seem sort of fluffy, whereas the other ones seem much more precise. We can formalize this as follows:

**Definition.** A monotone property $P$ has a *sharp threshold* if there is a function $f^*(n)$ such that for all $\varepsilon > 0$,

$$\mathbb{P}_p(P) \to \begin{cases} 0 & \text{if } p < (1-\varepsilon)f^*(n) \\ 1 & \text{if } p > (1+\varepsilon)f^*(n) \end{cases}$$

If this doesn't happen, the threshold is called *coarse*.

An equivalent definition is as follows: for any $\varepsilon > 0$, we define the threshold width to be

$$w(\varepsilon) = p(1-\varepsilon) - p(\varepsilon)$$

where

$$p(\varepsilon) = \sup\{p : \mathbb{P}_p(P) < \varepsilon\}$$

Then a sharp threshold is one that satisfies $w(\varepsilon) = o(f^*)$ for any threshold function $f^*$, while a coarse threshold satisfies $w(\varepsilon) = \Theta(f^*)$.

The astonishing observation of Friedgut is that local properties have coarse thresholds, while global properties have sharp thresholds.

**Theorem** (Friedgut, 1999)**.** *For any $\varepsilon > 0, C > 0$, there is some constant $K(C,\varepsilon)$ so that the following holds. For any $n$ and any property $P \subseteq \{0,1\}^{\binom{n}{2}}$ with a coarse threshold, specifically one satisfying*

$$f^*(n) \left.\frac{\mathrm{d}\mathbb{P}_p(P)}{\mathrm{d}p}\right|_{p=f^*(n)} \le C$$

*there exists a finite collection of graphs $H_1, \ldots, H_m$, each having at most $K(C,\varepsilon)$ edges, so that*

$$\mathbb{P}_{f^*(n)}(P \triangle P_{\mathcal{H}}) \le \varepsilon$$

*where $P_{\mathcal{H}}$ is the property of containing a copy of some $H_i$.*

**Corollary** (Friedgut, 1999). *Every coarse graph property has a threshold function given by $n^{-\alpha}$ for some $\alpha \in \mathbb{Q}$. Therefore, any graph property whose threshold function is not a rational power of $n$ must have a sharp threshold.*

*Proof idea.* As we saw in the beginning, containing a copy of $H$ has a threshold function, more or less, of $n^{-|V(H)|/|E(H)|}$. If $P$ had a sharp threshold, by Friedgut's theorem, it would have to have a threshold function of this form. □

# 6 Extensions, Briefly

There has been much work in random graphs after Friedgut-Kalai, most of which I can't get to. One important follow-up paper is due to Bourgain-Kalai, who studied what happens to threshold widths if we assume stronger symmetry conditions on $P$. In particular, observe that any property that is invariant under graph isomorphism is actually stronger than symmetric (as defined above), since $S_n$ permutes the $n$ vertices, which induces a fairly large group action on $N = \binom{n}{2}$. One consequence of their work is the following theorem in this special case:

**Theorem** (Bourgain-Kalai, 1997). *For any $\varepsilon > 0, \eta > 0$, and any monotone property $P$ that is invariant under graph isomorphism, we have that the threshold width satisfies*

$$w(\varepsilon) \leq C(\varepsilon) \frac{1}{(\log N)^{2-\eta}}$$

*for some universal constant $C(\varepsilon)$.*

Recent work, started by Linial-Meshulam, has focused on models of random simplicial complexes, which has proved very fruitful. In particular, there are results about the threshold for the vanishing of $k$th homology, which, much like the random graph case, is an instance of a necessary condition becoming sufficient.

Finally, let me mention a beautiful open problem in this field, known as the Kahn-Kalai conjecture. Let us count the expected number of Hamiltonian cycles in $\mathcal{G}(n,p)$. There are $(n-1)!$ possible Hamiltonian cycles (one for each cyclic permutation of $\{1,\ldots,n\}$), and each one is a Hamiltonian cycle with probability $p^{-n}$, since all $n$ edges must appear. Therefore,

$$\mathbb{E}(\#(\text{Hamiltonian cycles})) = (n-1)!p^n \sim \left(\frac{np}{e}\right)^n$$

Therefore, this expectation becomes $\geq 1$ when $p$ is of the order of $1/n$. However, we know that the actual threshold for Hamiltonicity is at $\log n/n$, so we see a logarithmic gap between the "expectation threshold" and the real threshold. We can do something similar for the number of spanning trees: by Cayley's Theorem,

$$\mathbb{E}(\#(\text{spanning trees})) = n^{n-2}p^{(n-1)} \sim (np)^n$$

so the expectation threshold is again when $p \sim 1/n$. However, a spanning tree exists iff the graph is connected, so again the actual threshold for the existence of a spanning tree is just the connectivity threshold, $p = \log n/n$. An analogous thing happens for the existence of a perfect matching, with the expectation threshold being at $1/n$ despite the real threshold being at $\log n/n$. The Kahn-Kalai conjecture says that these results are general: that for any property, there is at most a logarithmic gap between the expectation threshold and the connectivity threshold.