# Phase transitions in Bernoulli percolation

Talk by Byron Chin

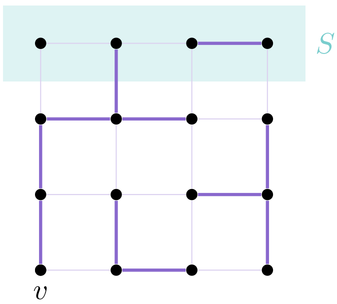
Notes by Sanjana Das

October 27, 2023

## §1 Introduction

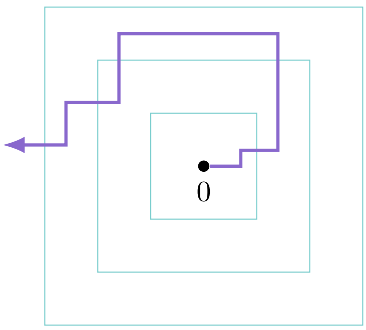
Let $G$ be a locally finite vertex-transitive infinite graph. Throughout this talk we'll work with the square lattice — so we'll take $G = \mathbb{Z}^2$ — but many of these results hold in greater generality. We'll consider the *Bernoulli percolation measure* $\mathbb{P}_p$ where we mark each edge *open* with probability $p$ (independently). We're interested in what the structure of these open edges looks like. Specifically, we can imagine deleting all the closed edges and seeing which things are still connected in the graph.

**Definition 1.1.** We say a vertex $v$ is connected to a vertex subset $S$, denoted $v \leftrightarrow S$, if there is a path consisting of open edges from $v$ to $S$.




**Definition 1.2.** We say $0 \leftrightarrow \infty$ if $0 \leftrightarrow \overline{\Lambda_n}$ for all $n$, where $\Lambda_n = [-n, n]^2$ (and $\overline{\Lambda_n}$ is its complement).

So we're looking at boxes $\Lambda_n$ centered at the origin, and we want to know whether there's a path from the origin that escapes all of these boxes — equivalently, whether there's an infinite cluster at the origin.



**Definition 1.3.** The critical parameter is defined as $p_c = \inf\{p \mid \mathbb{P}_p[0 \leftrightarrow \infty] > 0\}$.

In words, $p_c$ is the first probability at which we see an infinite cluster at the origin with positive probability. We actually know the exact value of $p_c$.

**Theorem 1.4** (Kesten)
We have $p_c(\mathbb{Z}^2) = \frac{1}{2}$.

So this tells us *where* the transition happens — i.e., where the probability of having an infinite path from 0 becomes positive. But we're also interested in *quantitative* results in this direction — how *quickly* does this transition happen? The following theorem — which we'll discuss today — gives a statement along these lines.

**Theorem 1.5**

(1) For $p > p_c$ (the *supercritical* regime), we have

$$\mathbb{P}_p[0 \leftrightarrow \infty] \geq \frac{p - p_c}{p(1 - p_c)}.$$

(2) For each $p < p_c$ (the *subcritical* regime), there exists a constant $c$ (depending on $p$) such that for all $n$, we have

$$\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}] \leq e^{-cn}.$$

In words, (1) says that above the critical probability, we have a pretty good probability of having an infinite cluster. (We don't know the exact growth rate for what this probability should be, but (1) at least gives a reasonable lower bound.) Meanwhile, (2) says that just below the critical probability, the probability we escape a $2n \times 2n$ box is exponentially small in $n$. So not only does this probability tend to 0, but it does so exponentially quickly.

## §1.1 History

Theorem 1.5 was first proven simultaneously by Aizerman–Barsky (1987) and Menshikov (1986). Aizerman–Barsky used a strategy based on differential inequalities (we'll discuss later what this means). They defined $p_c'$ as the smallest probability for which the *expected* cluster size at 0 is infinite, and proved that $p_c = p_c'$. They did this by writing down some differential inequalities for both $p_c$ and $p_c'$, and analyzing these differential inequalities to show that the two are equal and satisfy the properties in Theorem 1.5; they used a pretty analytic approach. Menshikov used a slightly different strategy. The same differential inequalities show up, but he used a more geometric approach — using the structure of $\mathbb{Z}^2$ and how paths in this space look in order to analyze these inequalities.

These initial works were both specifically for $\mathbb{Z}^2$ (and maybe $\mathbb{Z}^d$ for general $d$). They were later generalized to the general setting (with arbitary locally finite vertex-transitive graphs); this required extra work in the initial arguments.

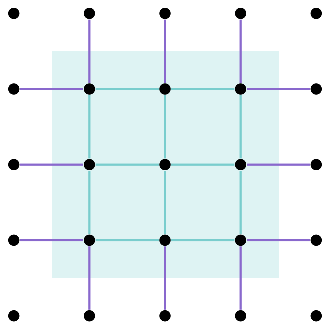
More recently, Duminil-Copin–Tassion (2016) reproved Theorem 1.5 in a nice, simple way that also extends to the more general setting (and this is the proof we'll discuss). They used a somewhat 'in-between' approach — they also defined differential inequalities for two quantities, but they analyzed these inequalities using a more geometric approach.

## §1.2 A new parameter

The main novelty of Duminil-Copin and Tassion is that they redefine the critical parameter in a new way (i.e., they define a new parameter $\widetilde{p_c}$ and show that it's the same as $p_c$). This definition may look somewhat strange at first, but after seeing the proof, it feels like a fairly natural parameter to define.

**Definition 1.6.** For a set of vertices $S$ (and vertices $x$ and $y$), we write $x \overset{S}{\leftrightarrow} y$ to mean that there is a path from $x$ to $y$ using only open edges inside (i.e., with both endpoints in) $S$.

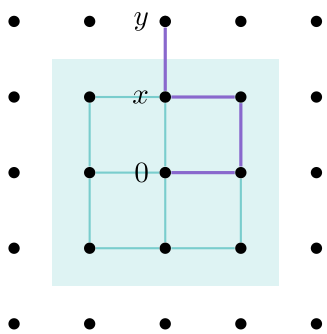
**Definition 1.7.** For a set of vertices $S$, we define $\partial S$ as the set of edges leaving $S$ — in other words, $xy \in \partial S$ means that $xy$ is an edge in the graph, $x \in S$, and $y \notin S$.



**Definition 1.8.** For a finite set of vertices $S$ containing the origin, we define

$$\varphi_p(S) = p \sum_{xy \in \partial S} \mathbb{P}_p[0 \overset{S}{\leftrightarrow} x].$$

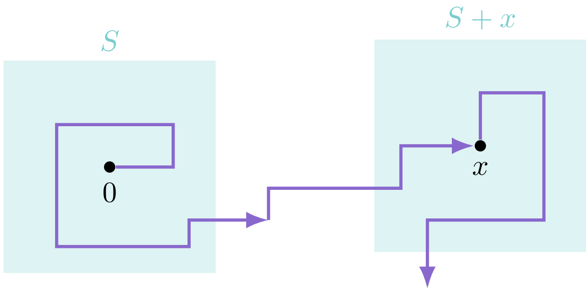
Intuitively, $\varphi_p(S)$ represents the expected number of ways to escape $S$ — we start off at the origin, and to escape $S$ we need to first move around within $S$ to reach some vertex $x$, and then go from $x$ to some vertex $y$ outside $S$ (the factor of $p$ corresponds to the probability that the edge $xy$ is open).




**Definition 1.9.** We define $\widetilde{p_c} = \sup\{p \mid \text{exists } S \text{ with } \varphi_p(S) < 1\}$.

So we're considering the threshold for there to be a set $S$ which is 'hard to escape' — meaning that on average there's less than one way to escape it.

First, here's a heuristic for why this definition makes sense (i.e., why we'd expect this to correspond to the actual critical parameter $p_c$). If $p < \widetilde{p_c}$, then there's a finite set $S$ that's hard to escape (for this argument, it's fine to visualize $S$ as being a box). Then the idea is that if our path from 0 escapes to $\infty$, it should have to escape $S$ (or more precisely, *translates* of $S$) infinitely many times. What does this mean? We start off at the origin, around which we have our set $S$, and in order to get to $\infty$ we need to first get out of $S$. Then we can imagine that we walk around a little bit, and now we place a translated copy of $S$ somewhere along this path. And then we need to get out of this translated copy of $S$ as well. And then we can walk around a bit more and then we'll need to escape another translated copy of $S$, and so on.




And the intuition is that each time, there's less than one way to escape $S$ on average, so it should be very unlikely that we're able to escape $S$ infinitely many times.

The main theorem of Duminil-Copin and Tassion is that Theorem 1.5 holds with $p_c$ replaced with $\widetilde{p_c}$.

**Theorem 1.10** (Duminil-Copin–Tassion 2016)

(1) For $p > \widetilde{p_c}$ (the *supercritical* regime), we have

$$\mathbb{P}_p[0 \leftrightarrow \infty] \geq \frac{p - \widetilde{p_c}}{p(1 - \widetilde{p_c})}.$$

(2) For each $p < \widetilde{p_c}$ (the *subcritical* regime), there exists a constant $c$ (depending on $p$) such that for all $n$, we have

$$\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}] \leq e^{-cn}.$$

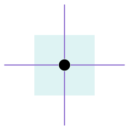
This automatically implies that $p_c = \widetilde{p_c}$ (since $p_c$ is defined as the smallest $p$ for which $\mathbb{P}[0 \leftrightarrow \infty]$ is positive, and it's positive above $\widetilde{p_c}$ and zero below $\widetilde{p_c}$), and therefore that the same statements hold for $p_c$ as well (as in Theorem 1.5).

**Remark 1.11.** For now, we can intuitively think of $S$ as a box — we want to think of $S$ as a set that's hard to escape, and a box is much harder to escape than e.g. a long path (since if $S$ were a long path, there'd be two ways to escape from each point). But when we're proving (1), we'll want to choose a specific $S$ that allows us to control the 'pivotal edges' of our percolation (we'll explain what this means later), which is why we need to allow general sets $S$ (rather than just boxes).

## §1.3 Some consequences

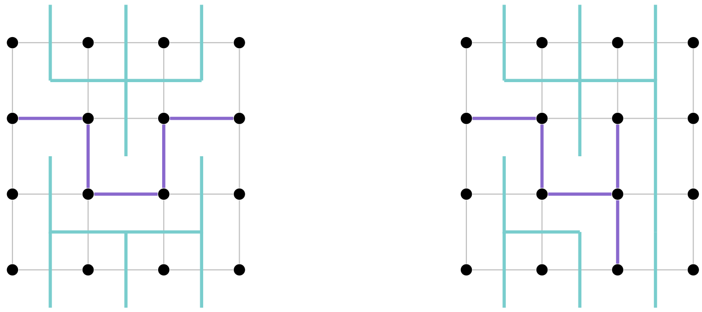
Before we prove Theorem 1.10, we'll discuss a few consequences. These consequences were already known or are standard results in percolation theory, but the fact that we can deduce them from Theorem 1.10 very quickly suggests that our definition of $\widetilde{p_c}$ really is a good thing to look at.

First, letting $S = \{0\}$, we have $\varphi_p(\{0\}) = 4p$ (since there are four edges out from $\{0\}$), so $\widetilde{p_c}$ (and therefore $p_c$) is at least $\frac{1}{4}$.



More generally, in $\mathbb{Z}^d$ this argument gives $p_c \geq 2^{-d}$; this is the same lower bound we'd get from a first moment argument where we count the number of paths.

A second consequence is that (2) implies $p_c \leq \frac{1}{2}$ (which is the harder direction of Kesten's theorem). The reason for this is that when $p = \frac{1}{2}$, by the self-duality of $\mathbb{Z}^2$, the probability there is a crossing from left to right in $\Lambda_n$ is roughly $\frac{1}{2}$. In more detail, there's a left-to-right crossing formed by edges if and only if there *isn't* an up-to-down crossing in the dual (where we draw edges between the centers of squares not separated by an edge); and the two are roughly symmetric (we may need to offset some of the dimensions by 1, but this isn't really important), so the probability of each is roughly $\frac{1}{2}$.



left-to-right crossing

no up-to-down crossing

no left-to-right crossing

up-to-down crossing

On the other hand, (2) implies that for $p < p_c$, the probability that there's a left-to-right crossing in $\Lambda_n$ is exponentially small (since it implies that for each point inside the box, the probability there's a left-to-right crossing through that point is exponentially small; and there's only polynomially many points). So we must have $p_c \leq \frac{1}{2}$, or else we could take $p = \frac{1}{2}$ and get a contradiction.

For a third consequence, first note that $\widetilde{p_c}$ is defined as the supremum of a union of open intervals (since the inequality $\varphi_p(S) < 1$ is strict), and such a supremum can't be attained by the set; this means we must have $\varphi_{p_c}(S) \geq 1$ for all $S$. Analyzing what happens at exactly the critical probability is generally one of the hardest things to do in these kinds of problems, but here this lets us deduce something — specifically, that the *expected* cluster size at the origin is infinite. This is because letting $\mathcal{C}(0)$ be the cluster at the origin, by linearity of expectation we have

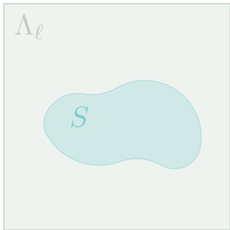
$$\mathbb{E}_{p_c} |\mathcal{C}(0)| = \sum_{x \in \mathbb{Z}^2} \mathbb{P}_{p_c}[0 \leftrightarrow x] \geq \sum_n \varphi_{p_c}(\Lambda_n)$$

(since $\varphi_p(\Lambda_n)$ is a sum over the vertices $x$ on the boundary of the box $\Lambda_n$, where we're essentially summing the probability that we can reach $x$ — i.e., that $0 \leftrightarrow x$ — and then escape from the box). And each of these terms $\varphi_{p_c}(\Lambda_n)$ is at least 1, so $\mathbb{E}_{p_c} |\mathcal{C}(0)| = \infty$.

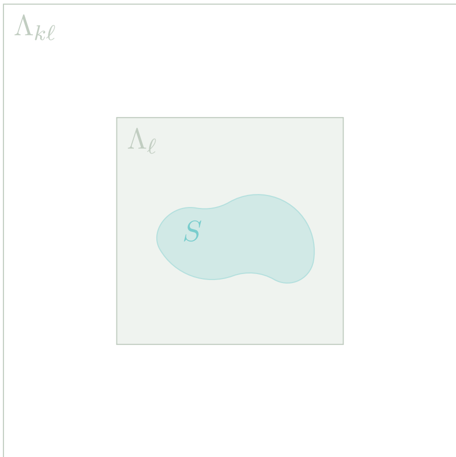
## §2 The subcritical case

We'll now talk about the proofs of (1) and (2); we'll start with (2), which is a bit more straightforward. The idea of the proof is essentially to formalize the heuristic mentioned earlier that to escape to $\infty$, we have to escape $S$ infinitely many times (and the same heuristic more quantitatively means that to escape $\Lambda_n$ we'll have to escape $S$ a number of times that's linear in $n$ — this is why we get exponential decay).

Let $p < \widetilde{p_c}$, and let $S$ be a finite set such that $\varphi_p(S) < 1$ (such a set exists by the definition of $\widetilde{p_c}$). Now fix some $\ell \in \mathbb{N}$ such that $S \subseteq \Lambda_\ell$ — so we're essentially encasing $S$ in some big box.

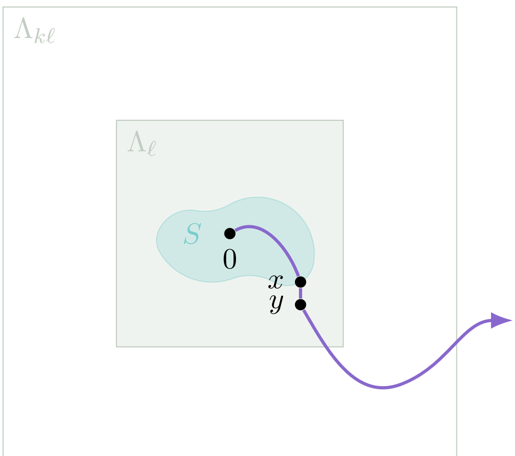



We're then going to inductively bound $\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_{k\ell}}]$, where $\ell$ is fixed and $k$ is the parameter we're inducting on — so we've fixed $S$ and enclosed it in a box $\Lambda_\ell$, and we're now trying to bound the probability we escape a much bigger box $\Lambda_{k\ell}$ (in terms of $\varphi_p(S)$).




**Claim 2.1** — We have $\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_{k\ell}}] \leq p \sum_{xy \in \partial S} \mathbb{P}_p[0 \overset{S}{\leftrightarrow} x] \mathbb{P}[y \leftrightarrow \overline{\Lambda_{k\ell}}]$.

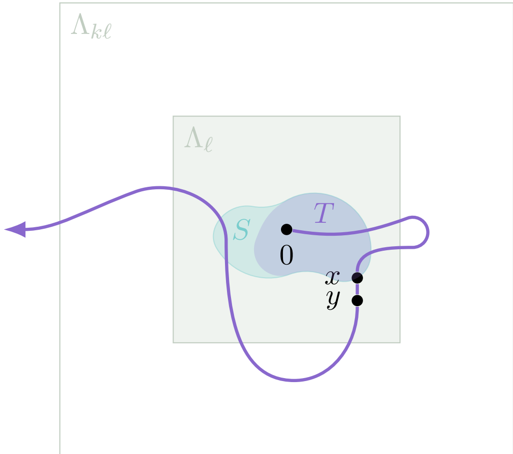
*Proof.* Intuitively, any path that escapes our big box $\Lambda_{k\ell}$ has to escape $S$ at some point, and we're partitioning the probability there exists such a path based on which edge $xy$ it escapes $S$ along — then $\mathbb{P}_p[0 \overset{S}{\leftrightarrow} x]$ corresponds to the probability we can get to $x$ (while staying within $S$), $p$ corresponds to the probability that the edge $xy$ is open, and $\mathbb{P}[y \leftrightarrow \overline{\Lambda_{k\ell}}]$ corresponds to the probability that we can escape the box from $y$.



However, we need to be a bit careful with how exactly we set this up to ensure that the three events depend on disjoint edges (because we want them to be independent).

So we let $T = \{x \mid 0 \overset{S}{\longleftrightarrow} x\}$ be the set of vertices reachable from the origin while staying within $S$, and we'll partition the probability $\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_{k\ell}}]$ by looking at the *last* vertex $x$ on our path such that $x \in T$. The path itself might go outside $S$ when going from 0 to $x$, but this doesn't matter — we only care that it's *possible* to get from 0 to $x$ while staying in $S$.

We then let $y$ be the vertex immediately after $x$ on this path — then we must have $y \notin S$ (or else $y$ would also be in $T$), so $xy \in \partial S$. And then our path must eventually escape the box from $y$, without ever returning to $T$ (since we chose $x$ to be the last vertex in $T$ on our path — it's possible that the path reenters $S$, but it can't reenter $T$).

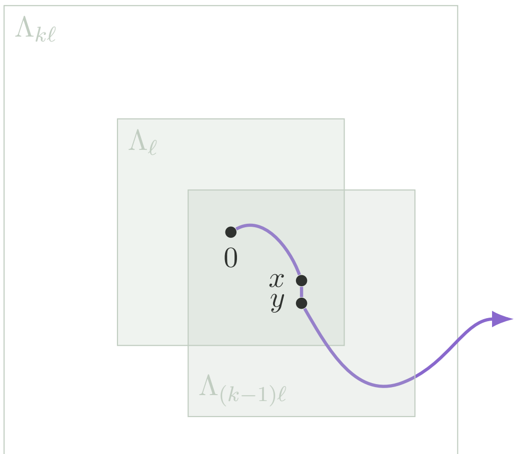



So then we must have $0 \overset{S}{\longleftrightarrow} x$, the edge $xy$ must be open, and we must have $y \overset{\overline{T}}{\longleftrightarrow} \overline{\Lambda_{k\ell}}$. These depend on disjoint sets of edges — the first event depends only on edges within $T$, and the third only depends on edges within $\overline{T}$ (and $xy$ goes between $T$ and $\overline{T}$). So they're independent, and we can upper-bound the probability this happens by

$$\mathbb{P}_p[0 \overset{S}{\longleftrightarrow} x]\mathbb{P}_p[xy \text{ open}]\mathbb{P}_p[y \leftrightarrow \overline{\Lambda_{k\ell}}] = p\mathbb{P}_p[0 \overset{S}{\longleftrightarrow} x]\mathbb{P}[y \leftrightarrow \overline{\Lambda_{k\ell}}],$$

as desired (and the final statement comes from summing over all possible edges $xy$). $\square$

But now for any $y$ with $xy \in \partial S$, we have $y \in \Lambda_\ell$ (since we chose $\ell$ such that $S \subseteq \Lambda_{\ell-1}$). So we can imagine drawing a slightly smaller big box $y + \Lambda_{(k-1)\ell}$ around $y$, and this box will have to live within our giant big box $\Lambda_{k\ell}$. In particular, if a path from $y$ escapes $\Lambda_{k\ell}$, then it also has to escape this slightly smaller box.



But by translational invariance, the probability that a path from $y$ escapes $y + \Lambda_{(k-1)\ell}$ is the same as the probability a path from 0 escapes $\Lambda_{(k-1)\ell}$. So we get

$$\begin{aligned}\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_{k\ell}}] &\leq p \sum_{xy \in \partial S} \mathbb{P}_p[0 \overset{S}{\leftrightarrow} x]\mathbb{P}[y \leftrightarrow \overline{\Lambda_{k\ell}}] \\ &\leq p \sum_{xy \in \partial S} \mathbb{P}_p[0 \overset{S}{\leftrightarrow} x]\mathbb{P}[0 \leftrightarrow \overline{\Lambda_{(k-1)\ell}}] \\ &= \varphi_p(S) \cdot \mathbb{P}[0 \leftrightarrow \overline{\Lambda_{(k-1)\ell}}],\end{aligned}$$

and so by induction, for all $k$ we have

$$\mathbb{P}[0 \leftrightarrow \overline{\Lambda_{k\ell}}] \leq \varphi_p(S)^k.$$

And we chose $S$ such that $\varphi_p(S) < 1$, which means this probability is exponentially small in $k$ (and therefore in $n = k\ell$, as $\ell$ is fixed).

**Remark 2.2.** Intuitively, what's happening in this proof is that we start off with a single set $S$ that we need to escape, and we use induction to formalize the idea that we have to escape an infinite string of sets $S$ (or more precisely, to escape $\Lambda_{k\ell}$ we need to escape a string of $k$ translates of $S$).

# §3 The supercritical case

We'll now prove (1), which gives a lower bound on $\mathbb{P}_p[0 \leftrightarrow \infty]$ for $p > \widetilde{p_c}$.

## §3.1 Russo's formula

The proof of (1) aligns with a general theory of Boolean functions; in particular, a key ingredient is Russo's formula, which we'll now state.

**Definition 3.1.** We let $\mu_p$ denote the $p$-biased measure on the hypercube $\{0,1\}^n$ — i.e., the probability measure where each coordinate is set to 1 independently with probability $p$.

**Definition 3.2.** We say $\mathcal{A} \subseteq \{0,1\}^n$ is an increasing event if for every $x \in \mathcal{A}$, if we obtain $y$ from $x$ by changing some 0's to 1's, then $y \in \mathcal{A}$ as well.

**Definition 3.3.** Let $\mathcal{A}$ be an increasing event. For each index $i \in [n]$, we define the influence of $i$ on $\mathcal{A}$, denoted $\mathrm{Inf}_i(\mathcal{A})$, as the probability that the coordinate $i$ is 'pivotal' to $\mathcal{A}$ — i.e.,

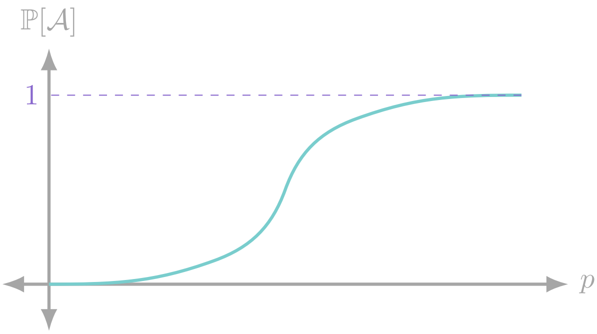
$$\mathrm{Inf}_i(\mathcal{A}) = \mu_p(\{\omega \in \{0,1\}^{[n]\setminus\{i\}} \mid \omega^{i=1} \in \mathcal{A},\, \omega^{i=0} \notin \mathcal{A}\}).$$

In words, we imagine choosing all the coordinates of our point except the $i$th one, and we consider the probability that the $i$th coordinate is pivotal for whether we end up in $\mathcal{A}$ — meaning that setting the $i$th coordinate to 0 places us outside $\mathcal{A}$, while setting it to 1 places us inside $\mathcal{A}$. (We use $\omega^{i=1}$ to denote $\omega$ with the $i$th coordinate set to 1.)

**Theorem 3.4** (Russo's formula)
For any increasing event $\mathcal{A}$, we have

$$\frac{d}{dp}\mathbb{P}[\mathcal{A}] = \sum_{i=1}^{n} \mathrm{Inf}_i(\mathcal{A}).$$

The reason this is useful is that if we're interested in the probability of some increasing event $\mathcal{A}$, then we can imagine graphing this probability as a function of $p$. If $\mathcal{A}$ is nontrivial, this probability will go from 0 to 1 (as $p$ goes from 0 to 1). We're often interested in how quickly this transition happens, and that depends on the derivative of this probability; and Russo's formula says that we can compute this derivative by looking at how much each coordinate influences our event.




For example, if we've seen the Friedgut–Kalai result on monotone graph properties having a sharp threshold, its proof works by invoking Russo's formula and then using some other results to bound this influence.

Percolation also often uses some version of this formula, and there we often bind the influence using geometric ideas. In this case, we'll see that the influence is strongly related to the function $\varphi_p$. If there's time, we'll later talk about some other problems where the influence has a nice geometric interpretation.

## §3.2 A proof of (1)

We can put percolation into the framework of Boolean functions by considering $\{0,1\}^{e(G)}$, where an entry of 1 represents an open edge and 0 represents a closed edge (and $G = \mathbb{Z}^2$ is our graph). (Technically Russo's formula only applies to *finite* hypercubes, so we really only consider the edges in a large box.) Then Russo's formula gives

$$\frac{d}{dp}\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}] = \sum_{e \in \Lambda_n} \mathbb{P}[e \text{ is pivotal}],$$

where we say an edge $e$ is *pivotal* if toggling whether it's open or closed would change whether $0 \leftrightarrow \overline{\Lambda_n}$. (Russo's formula gives a sum over *all* edges $e$, but edges outside of $\Lambda_n$ are certainly not pivotal — they don't affect whether we can escape $\Lambda_n$.) It'll be more convenient to rewrite this as

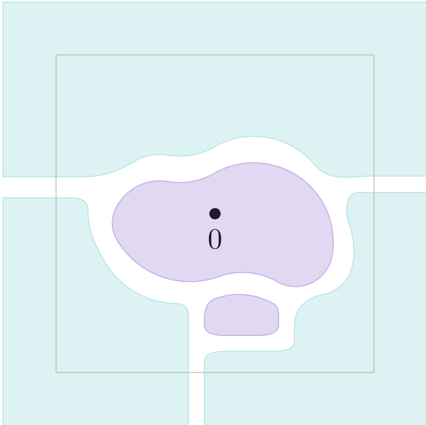
$$\frac{d}{dp}\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}] = \frac{1}{1-p}\sum_{e \in \Lambda_n} \mathbb{P}[e \text{ is pivotal and } 0 \not\leftrightarrow \overline{\Lambda_n}] \tag{1}$$

(this is because $e$ being pivotal tells us that if we set $e$ to be closed — which occurs with probability $1-p$ — then we have $0 \not\leftrightarrow \overline{\Lambda_n}$).

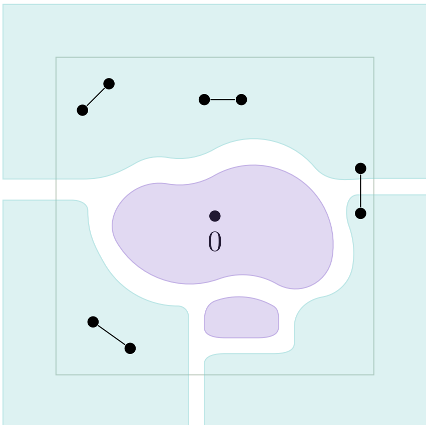
Now suppose that $0 \not\leftrightarrow \overline{\Lambda_n}$, and let

$$I = \{x \in \Lambda_n \mid x \not\leftrightarrow \overline{\Lambda_n}\}$$

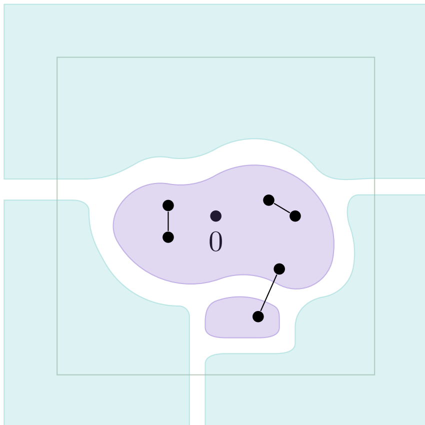
(so that $0 \in I$) — i.e., $I$ is the set of vertices *not* connected to the outside of our box. So we've got a blob $I$ (or possibly a collection of several blobs) not connected to the outside of our box, while everything else *is* connected to the outside. (In the below figure, $I$ consists of the two purple blobs.)




This definition of $I$ lets us pinpoint which edges are pivotal — an edge is pivotal if flipping it from closed to open means that suddenly there *is* a path from 0 to outside our box. Flipping an edge outside $I$ doesn't do anything to help us, since we can't have reached that edge from 0 in the first place. Flipping an edge *inside* $I$ doesn't help us either — neither endpoint of that edge can reach the outside of the box, so even if we can now reach one of those endpoints from 0, this doesn't help us reach the outside.

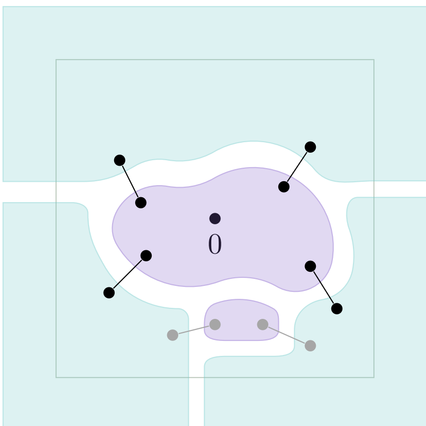



edges outside $I$




edges inside $I$

So the only edges that could possibly be pivotal are the ones with one endpoint in $I$ and one endpoint not in $I$ — i.e., edges $xy \in \partial I$. And such an edge $xy$ is pivotal if and only if $x$ (its endpoint in $I$) is in the island containing 0 — this is because when we open the edge $xy$, now $x$ becomes connected to the outside of our box, and so 0 becomes connected to the outside if and only if 0 was already connected to $x$.




edges in $\partial I$

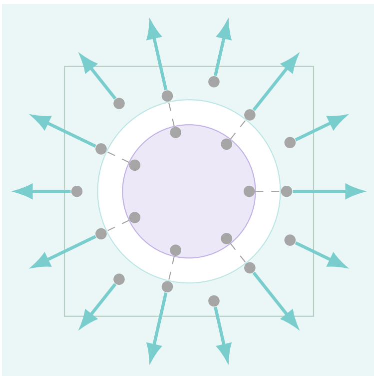
Now we'll break up the right-hand side of (1) based on what $I$ is — we can rewrite it as

$$\frac{1}{1-p}\sum_{S}\sum_{xy\in\partial S}\mathbb{P}[xy \text{ pivotal and } I = S]$$

(where we're summing $S$ over all possibilities for $I$ for which $0 \not\leftrightarrow \overline{\Lambda_n}$ — i.e., all sets with $0 \in S$ and $S \subseteq \Lambda_n$). And if $I = S$, then as seen above, an edge $xy \in \partial S$ is pivotal if and only if 0 is connected to $x$; so we can rewrite this as

$$\frac{1}{1-p}\sum_{S}\sum_{xy\in\partial S}\mathbb{P}[0 \overset{S}{\leftrightarrow} x,\, I = S].$$

And now the key point is that the event $0 \overset{S}{\leftrightarrow} x$ only depends on the edges inside $S$, while the event $I = S$ only depends on edges *not* inside $S$ (i.e., edges either outside $S$ or on its boundary). This is because the event $I = S$ requires just that all edges leaving $S$ are closed, and that all vertices outside $S$ are connected to the outside of our box (through paths that stay outside $S$) — this doesn't at all reference the edges inside $S$. (Opening or closing edges inside $S$ can affect how it splits into islands, but won't affect whether we can get from $S$ to the outside.)



And so the two events depend on disjoint sets of edges, which means they're independent, and we can split this probability as a product; then our sum becomes

$$\frac{1}{1-p}\sum_{S}\sum_{xy\in\partial S}\mathbb{P}[0 \overset{S}{\leftrightarrow} x]\mathbb{P}[I = S].$$

And now we can see where $\varphi_p$ comes in — we have

$$p\sum_{xy\in\partial S}\mathbb{P}[0 \overset{S}{\leftrightarrow} x] = \varphi_p(S) \geq 1$$

for all $S$ (as $p > \widetilde{p_c}$). So we can use this to get rid of the inner sum, and we get

$$\frac{1}{1-p}\sum_{S}\sum_{xy\in\partial S}\mathbb{P}[0 \overset{S}{\leftrightarrow} x]\mathbb{P}[I = S] \geq \frac{1}{p(1-p)}\sum_{S}\mathbb{P}[I = S] = \frac{1}{p(1-p)}\mathbb{P}[0 \not\leftrightarrow \overline{\Lambda_n}].$$

(Here we're collapsing our choice of $S$ back together — we're summing $S$ over all possibilities for $I$ for which we have $0 \not\leftrightarrow \overline{\Lambda_n}$, so $\sum_S \mathbb{P}[I = S] = \mathbb{P}[0 \not\leftrightarrow \overline{\Lambda_n}]$.) Plugging this into (1) gives

$$\frac{d}{dp}\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}] \geq \frac{1}{p(1-p)}\mathbb{P}[0 \not\leftrightarrow \overline{\Lambda_n}]. \tag{2}$$

So we started out with Russo's formula for the derivative of $\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}]$ and bounded the expression it gave by looking carefully at what pivotal edges look like; and we ended up with an expression involving the same probability (or rather, its complement). This is a *differential inequality* — it says that the derivative of a certain function of $p$ (namely, $\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}]$) is lower-bounded by the same function of $p$.

And this differential inequality (2) is true for all $p > \widetilde{p_c}$, so now we can integrate from $\widetilde{p_c}$ to our value of $p$ to get a lower bound for $\mathbb{P}_p[0 \leftrightarrow \overline{\Lambda_n}]$; and what comes out of this is exactly the bound in (1).

**Remark 3.5.** This proof works with $\mathbb{Z}^2$ replaced with more general graphs as well (though the value of $p_c$ will be different — the fact that $p_c = \frac{1}{2}$ for $\mathbb{Z}^2$ relies on the self-duality of $\mathbb{Z}^2$).

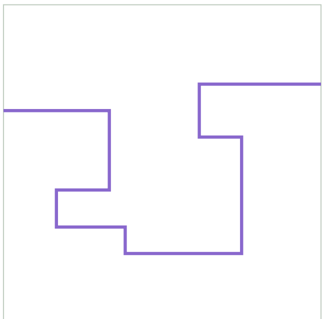
# §4 A related problem — noise sensitivity of crossings

The proof of Theorem 1.10(1) we saw gives a flavor of how many arguments in this area work — we produce a differential inequality using Russo's formula, often by using something about our picture to control influences. We'll now briefly discuss another result whose proof fits into the same idea.

## §4.1 The setup

Imagine we sample a percolation $\omega$ with probability $p = \frac{1}{2}$. We're interested in whether there's a left-to-right crossing in a large square box $\Lambda_n$.

**Definition 4.1.** We define $\mathrm{LR}(\omega)$ as the event that there is a left-to-right crossing of $\Lambda_n$ (in $\omega$).



By the self-duality of $\mathbb{Z}^2$, we know that the probability there is such a crossing is $\frac{1}{2}$. But what we're *actually* interested in is how sensitive the existence of such a crossing is to 'noise.' So once we've got $\omega$, we define a new percolation $\omega_t$ by resampling each edge with probability $t$. (This means for each edge, with probability $1-t$ we keep the edge as it is in $\omega$, and with probability $t$ we flip a new coin to resample whether that edge is open or closed.)

**Question 4.2.** What is $\mathrm{Cov}[\mathrm{LR}(\omega), \mathrm{LR}(\omega_t)]$?

In other words, we're interested in how this resampling affects whether there's a crossing — i.e., how sensitive the property of having a crossing is to some amount of noise (where the amount of noise corresponds to $t$).

**Theorem 4.3**

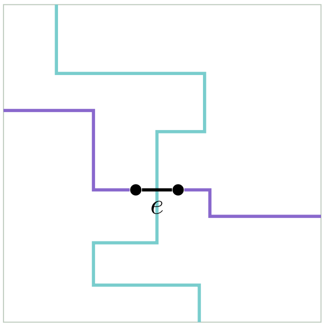
There is a constant $c$ such that

$$\mathrm{Cov}[\mathrm{LR}(\omega), \mathrm{LR}(\omega_t)] = \begin{cases} 0 & \text{if } t \gg n^{-c} \\ \frac{1}{4} & \text{if } t \ll n^{-c}. \end{cases}$$

(We might not know the explicit value of $c$ for $\mathbb{Z}^2$.)

The idea is to use Russo's formula to take the derivative with respect to $t$ of this covariance; then we get an expression of the form $\sum_e \mathbb{P}[e \text{ is pivotal}]$, where we say $e$ is pivotal if opening $e$ means there is a left-to-right crossing and closing $e$ means there isn't.

So for $e$ to be pivotal, we know that when $e$ is open, there's a left-to-right crossing that goes through $e$. Meanwhile, when $e$ is closed, we use duality — the fact that there's no left-to-right crossing means there's instead an up-to-down crossing in the dual (where we draw edges in the dual connecting the centers of squares for which the edge between them is closed).

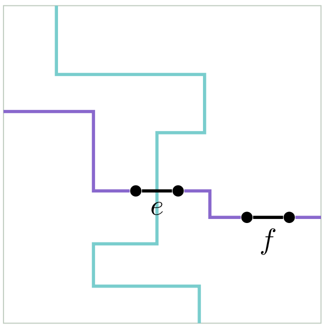



So $e$ being pivotal corresponds to the event that we've got these four arms coming off of $e$ (with the left and right arms consisting of open edges, and the up and down arms consisting of the duals of closed edges); we call this a 4-*arm event*. Then Russo's formula tells us that the derivative of $\mathrm{Cov}[\mathrm{LR}(\omega), \mathrm{LR}(\omega_t)]$ is controlled by the probabilities of these 4-arm events.

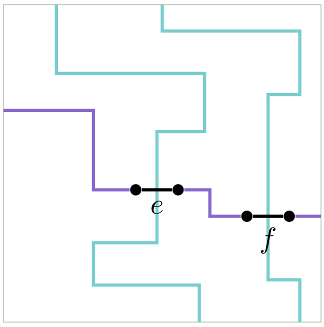
But now we're kind of stuck — in our proof of Theorem 1.10(1), when we took a derivative of our function, we got back a bound in terms of the *same* function, which let us integrate. But here we *don't* have the same function; instead we have some weird-looking event.

But the point is that we can try using Russo's formula again to compute the probabilities of these 4-arm events. To do so, we want to figure out the pivotal events for these 4-arm events — what does it mean for an edge $f$ to be pivotal? This means we originally have a crossing left-to-right (consisting of open edges) and up-to-down (consisting of the duals of closed edges), and flipping $f$ kills one of these crossings. This means $f$ has to lie on one of our arms (more precisely, either $f$ is on the left-to-right crossing or its dual is on the up-to-down crossing) — otherwise flipping $f$ wouldn't affect either crossing.

Let's suppose for concreteness that $f$ is on the right arm of our left-to-right crossing.




Then $f$ is originally open, and closing it needs to destroy *all* left-to-right crossings through $e$ — and we can still reach the left boundary from $e$ (since the left arm of our left-to-right crossing is still intact), so closing $f$ has to cut off $e$ from the right boundary. And this means in the dual there's an up-to-down crossing that goes through $f$ and separates $e$ from the right boundary.




And now it's possible to essentially split this picture into three separate 4-arm events — roughly speaking, we draw a box around our original pivotal edge $e$ (where we have one 4-arm event) and a box around our new pivotal edge $f$ (where we have another 4-arm event), and on the outside there's a bigger 4-arm event. (Actually figuring out how to design this such that the events are independent is quite intricate — this is the case for lots of arguments of this form — and we won't go into the details.)

So we can decompose our pivotal event for a 4-arm event into three separate 4-arm events. And then Russo's formula tells us that the derivative of the probability of a 4-arm event is controlled by the probabilities of other 4-arm events. Then we can do calculus to get precise estimates for these probabilities (since we have 4-arm events on both sides); and then we can integrate to get estimates for our original covariance.