BRANCHING PROCESSES AND APPLICATIONS

TAL ORENSHTEIN

ABSTRACT. These are notes of a talk given at the probability student seminar in the Weizmann institute of science on September 2011. After introducing Galton-Watson branching process, we consider conditions for which the process survives forever and for which it has a binary tree as a subtree with the same root. As an application, Mandelbrot's fractal percolation model is discussed.

CONTENTS

1. GALTON - WATSON PROCESS

The following is the first paragraph of the term *Galton-Watson process* in Wikipedia:

> The Galton–Watson process is a branching stochastic process arising from Francis Galton's statistical investigation of the extinction of family names. [...] There was concern amongst the Victorians that aristocratic surnames were becoming extinct. Galton originally posed the question regarding the probability of such an event in the Educational Times of 1873, and the Reverend Henry William Watson replied with a solution. Together, they then wrote an 1874 paper entitled *On the probability of extinction of families* ([9]). Galton and Watson appear to have derived their process independently of the earlier work by I. J. Bienaymé.

Date: September 14, 2011.

A gap in Galton and Watson's paper was found. Bienaymé failed to supply a valid derivation either. A solid proof was finally given by Stefengen in 1930 (see the footnote in page 173 of Grimmett and Stirzaker [6]).

Let X be an integer valued random variable. Assume that $\zeta_i^k,\ i,k=1,2,\ldots$ are i.i.d. copies of X. Define a sequence $(Z_n)_{n\geq 1}$ of random variables by:

$$Z_0=1,\ \ Z_{n+1}=\zeta_1^{n+1}+\zeta_2^{n+1}+\ldots+\zeta_{Z_n}^{n+1},\ n\geq 0.$$

The sequence $(Z_n)_{n\geq 1}$ is called **Galton-Watson process**, and the distribution of X, $p_k=\mathbb{P}(X=k),\ k=0,1,2,\ldots$, is called the **offspring distribution**. We will sometimes write X **process** in short. A way to think about this process is the following: Z_n is the number of individuals in the n-th generation. Each individual at the n-th generation gives birth to a number of children distributed by the offspring distribution, independently of all other individuals.

2. SURVIVAL

With a mathematical model in hands, Galton's question formulates to:

Problem. (Survival) When is there a positive probability for the process to survive forever?

In this section a full answer to the survival problem, in terms of the expected offspring distribution, will be presented. The solution is taken from chapter 5.4 of Grimmett and Stirzaker [6]. A few other recommended references are the first chapter of Athreta and Ney [1] and Section 5.3.4 of Durrett [5].

Remark. $\mathbb{P}(X=0)>0$ if and only if $\mathbb{P}(Z_n>0 \text{ for all } n)<1$.

In particular, unless the distribution of X is a.s. identically zero, one cannot expect an a.s. extinction.

Before dealing with the survival problem, we shall recall some properties of power series.

Let $G(s)=\sum_{i=0}^{\infty}a_is^i$, where $a_i\in\mathbb{R}$. The followings hold.

- **Convergence.** There is a **radius of convergence** $R\in[0,\infty]$ such that the series converges absolutely for $|s|<R$, and diverge if $|s|>R$. Moreover, it converges uniformly on $|s|<R'$ for each $R'<R$.
- **Differentiation.** $G(s)$ may be differentiated or integrated term by term on $|s|<R$, and moreover, this can be done **any number of times**.

- **Abel's theorem** may be used to understand the behavior at the radius of convergence: Assume that $a_i \geq 0$ for all i, and the sum is finite for all $|s| < 1$. Then exists $\lim_{s \nearrow 1} G(s) = \sum_{i=0}^{\infty} a_i \in [0, \infty]$.

Definition. The *probability generating function* (p.g.f.) G_Y of a random variable Y is given by $G_Y(s) = \mathbb{E}(s^Y)$. In particular, if Y has values $0, 1, 2, \ldots$, then $G_Y(s) = \sum_{i=0}^{\infty} \mathbb{P}(Y = i) s^i$.

Let Y be a random variable with values $0, 1, 2, \ldots$. Here are some **properties of** G_Y:

(1) $G_Y(0) = \mathbb{P}(Y = 0)$ and $G_Y(1) = 1$.
(2) G_Y is a power series with non-negative coefficients. Its radius of convergence $R \geq 1$.
(3) By 2 G_Y is an increasing function.
(4) By Abel's theorem, $G'_Y(1) = \sum_{i=1}^{\infty} i\mathbb{P}(Y = i) = \mathbb{E}X \in [0, \infty]$, where $G'(1)$ is a shorthand of $\lim_{s \nearrow 1} G'(s)$. Moreover, $E[X(X-1)\ldots(X-k+1)] = G^{(k)}(1)$, where $G^{(k)}(1)$ is a shorthand of $\lim_{s \nearrow 1} G^{(k)}(s)$.
(5) By 2 and 4, using differentiation term by term, G_Y is convex on $[0, 1]$ (i.e. $G'' \geq 0$).

Remark. The probability generating function G_Y gained its name by the **formula for the coefficients:** $\mathbb{P}(Y = n) = \frac{1}{n!} G^{(n)}(0)$.

Let's go back to our problem. The following theorem is the heart of the solution.

Theorem 2.1. *Let $Z = (Z_n)_{n \geq 0}$ be a X - process. Let η be the probability for an ultimate extinction of the process (that is, η is the probability that there is some $n < \infty$ for which $Z_n = 0$). Then, η is the minimal fixed point of G in the interval $[0, 1]$, where $G(s) = \sum_{i=0}^{\infty} p_i s^i$ is the p.g.f. of X (i.e., $p_i = \mathbb{P}(X = i)$, $i = 0, 1, \ldots$).*

Let's see how to get an answer to our problem assuming theorem 2.1.

Corollary 2.2. $\eta = 1$ *if and only if* $\mathbb{E}X \leq 1$, *given that* $p_1 < 1$.

Proof. By theorem 2.1 we need to show that there is some $s < 1$ so that $G(s) = s$ if and only if $G'(1) > 1$. Indeed, assume first that $G'(1) > 1$. If $G(0) = \mathbb{P}(X = 0) = 0$ we are done, so assume that $G(0) > 0$. Define $f(s) = G(s) - s$ for $s \in [0, 1]$. Then $f'(1) > 0$ and by continuity of f' (the easy case that $G'(1) = \infty$ is excluded here), it means that there is an $\varepsilon > 0$ so that $f'(s) > 0$ for all $s \in [1-\varepsilon, 1]$. In particular $f(1-\varepsilon) < f(1) = 0$. On the other hand, $f(0) = G(0) = \mathbb{P}(X = 0) > 0$. By the mean value theorem for the function f we are done.

In the other direction, assume that $G'(1) \leq 1$. We note that as by assumption, $\mathbb{P}(X = 1) < 1$, then $G \mid_I \neq id \mid_I$ for all intervals or converging sequences $I \subset [0,1]$ (e.g. by the uniqueness theorem for analytic complex functions this would imply that $G = id$ on $[0,1]$, i.e. $p_1 = 1$). Therefore, if there is $0 \leq s < 1$ for which $G(s) = s$, then we have a *maximal* such, denote it by $s_0 < 1$. It means we have an interval $(s_0, 1)$ so that $f < 0$ on it. Let $s_m \in (s_0, 1)$ so that f is the minimal in $(s_0, 1)$ (exists such by continuity of f. Then for all $t \in (0,1)$ $f(t1 + (1-t)s_m) \geq f(s_m) \geq (1-t)f(s_m) = tf(1) + (1-t)f(s_m)$. By convexity of f (which is implied by the one of G), for all $t \in (0,1)$ $f(t1 + (1-t)s_m) = (1-t)f(s_m)$. Therefore $f'(1) = -\frac{1}{1-s_m} f(s_m) > 0$, or $G'(1) > 1$, a contradiction.

□

Remark. If $p_1 = 1$ *then* $\mathbb{E}X = 1$, but $\eta = 0$. Therefore the condition $p_1 < 1$ is necessary for deducing the direction $(\Rightarrow)$

A crucial property of the p.g.f. is the following.

Proposition 2.3. *Let $X_1, X_2, \ldots$ be independent random variables with common p.g.f G_X, and let N be a random variable, independent of all the rest, which takes values $0, 1, 2, \ldots$. Then $G_{X_1+X_2+\ldots+X_N}(s) = G_N(G_X(s))$, with the convention that $\sum_{i=1}^{0} x_i = 0$.*

Corollary 2.4. *Denote $G_n = G_{Z_n}$ and remember that $G = G_X$.*
Then, $G_n(s) = G(G_{n-1}(s)) = \ldots = G(G(\ldots(G(s)))) = G_{n-1}(G(s))$.

We are now ready to prove theorem 2.1.

Proof. (of theorem 2.1) Let $\eta_n = \mathbb{P}(Z_n = 0)$. By the continuity of the probability function, and the fact that the event that $Z_n = 0$ implies that also $Z_{n+1} = 0$, we have that $\eta_n \nearrow \eta$ as $n \to \infty$. By definition of G_n, $\eta_n = G_n(0)$ and by Corollary2.4, $\eta_n = G(G_{n-1}(0))$. By induction, $\eta_{n+1} = G(\eta_n)$ (note that η_0 =0). By continuity of G, $G(\eta) = \lim_{n\to\infty} G(\eta_n) = \lim_{n\to\infty} \eta_{n+1} = \eta$. To finish the proof, we need to show that η is the minimal non-negative fixed point (note that there is at most one fixed point < 1, by convexity, but as we wish to imitate this proof later on without a convexity assumption we will not show that). Let ρ be another non-negative fixed point of G. $\rho \geq 0 = \eta_0$. By monotonicity of G, $\rho = G(\rho) \geq G(\eta_0) = \eta_1$. By induction, using monotonicity, $\rho \geq \eta_n$ for all n and hence also for the limit $\rho \geq \eta$.

□

Note that to prove $G^n(0) \to$ smallest fixed point of G, we only used two properties of G in the last proof: *monotonicity* and *continuity.*

Example 2.5. Use theorem 2.1 Galton-Watson process for some known offspring distributions:

- $\mathbb{E}(Poi(\lambda)) = \lambda$, and so $Poi(\lambda)$ process goes extinct a.s. iff $\lambda \leq 1$.
- $\mathbb{E}(Geom(p)) = \frac{1-p}{p}$. By theorem 2.1, and so $Geom(p)$ process goes extinct a.s. iff $p \geq \frac{1}{2}$ (remember that $X \sim Geom(p)$ if $\mathbb{P}(X = k) = (1-p)^k p$, $k = 0, 1, ...$).
- $\mathbb{E}(Binom(p,n)) = pn$, and so $Binom(p,n)$ process goes extinct a.s. iff $p \leq \frac{1}{n}$.

Example 2.6. If $p_0 = 1-p,\ p_2 = p$. Let $p_s = \sup\left\{p \in [0,1] : \mathbb{P}_p(\text{suvival}) > 0\right\}$. As the expectation is $2p$, we have a.s. extinction $\Leftrightarrow$ $p \geq \frac{1}{2}$. Therefore $p_s = \frac{1}{2}$. Note that $\mathbb{P}_{p_s}(\text{suvival}) = 0$. Compare later on with Example3.4.

3. Splitting

A natural realization of the process is discrete rooted trees. Hence, the process defines a probability measure on rooted trees. For example, the measure of a finite rooted tree of depth n, with number z_t of vertices in the tth level, $0 \leq t \leq n$ is $\mathbb{P}_{Z_0=1}(Z_t = z_t,\ 1 \leq t \leq n)$.

In this perspective we may translate questions of survival "in time" to questions of existence of infinite objects. For example Corollary 2.2 reformulated as

Corollary 3.1. $\mathbb{P}$(*Z contains an infinite tree with the same root*) $= 0$ *if and only if* $\mathbb{E}X \leq 1$, *given that* $p_1 < 1$.

With that formulation, the following question may be seemed more natural.

Problem. (Binary splitting) What are the conditions of having a **binary tree** as a sub-graph of Z with the same root?

The binary splitting question was raised and solved by Dekking [4]. As we shall see later on, a private case was discussed even earlier informally by Mandelbrot in his book "The fractal geometry of nature" from 1983 ([7], chapter 23), and formally by Chayes, Chayes and Durrett [3].

Definition 3.2. We say that the process $Z = (Z_n)$ has a *binary splitting* if it contains a **binary tree** as a sub-graph of Z with the same root.

Theorem 3.3. *(Dekking 1991) Let τ be the probability that the X - process Z has a binary splitting, then $1-\tau$ is the smallest fixed point of $f(s) = G(s) + (1-s)G'(s)$ in $[0,1]$.*

Proof. Let γ_n be the probability that Z does *not* contain a binary sub-tree of height n with the same root. Clearly, $\lim_{n\to\infty}\gamma_n \equiv \gamma = 1-\tau$. Given that at level one the process has k nodes, i.e. $Z_1 = k$, it does *not* contain a binary sub-tree of height $n+1$ if and only if $k=0$, $k=1$, all but (maybe) not more than one k sub-trees rooted at these nodes does *not* contain a binary sub-tree of height n. Therefore,

$$\begin{aligned}
\gamma_{n+1} &= p_0 + p_1 + \sum_{k\geq 2}(\gamma_n^k + k\gamma_n^{k-1}(1-\gamma_n))p_k \\
&= p_0 + p_1 + \sum_{k\geq 2}\gamma_n^k p_k + (1-\gamma_n)\sum_{k\geq 2} k\gamma_n^{k-1}p_k \\
&= p_0 + p_1 + \sum_{k\geq 0}\gamma_n^k p_k - p_0 - \gamma_n p_1 + (1-\gamma_n)\sum_{k\geq 2} k\gamma_n^{k-1}p_k \\
&= G(\gamma_n) + (1-\gamma_n)\sum_{k\geq 1} k\gamma_n^{k-1}p_k \\
&= G(\gamma_n) + (1-\gamma_n)G'(\gamma_n) \\
&= f(\gamma_n)
\end{aligned}$$

Note that (or define, if you wish) $\gamma_0 = 0$, and so $\gamma_n = f^n(0)$. Since $f'(s) = (1-s)G''(s)$ is non-negative f is continuously increasing on $[0,1]$. As in the note after the proof of theorem 2.1, it follows that $\gamma = \lim_{n\to\infty} f^n(0)$ is the smallest fixed point of f in $[0,1]$.

$\square$

Example 3.4. (discontinuity of τ at a critical point, by Dekking) For each p, let $\tau(p)$ be the binary splitting probability Galton-Watson process with off spring distribution given by $p_1 = 1-p$ and $p_3 = p$.

Then, $\tau(p) = \begin{cases} 0 & \text{for} \quad 0 \leq p < \frac{8}{9} \\ \frac{3}{4} + \frac{3}{4}\sqrt{1-\frac{8}{9p}} & \text{for} \quad \frac{8}{9} \leq p \leq 1 \end{cases}$

Note: Let $p_c = \inf\{0 \leq p \leq 1 : \ \tau(p) > 0\}$, then $p_c = \frac{8}{9}$ is the critical probability, and $\lim_{p\nearrow p_c}\tau(p) = 0 < \tau(p_c)$. Compare with Example 2.6.

Proof. The p.g.f. is $G(s) = (1-p)s + ps^3$, and so $1-\tau(p)$ is the smallest $s \in [0,1]$ for which

$1 - p + 3ps^2 - 2ps^3 = s$. Equivalently, $1 - s = p(2s^3 - 3s^2 + 1)$. Assume $s < 1$ and divide by $1-s$ to get $1 = -p(2s^2 - s - 1)$. Solving this we get

$s_{1,2} = \frac{3}{4} \pm \frac{1}{4p}\sqrt{p(9p-8)}$. Therefore, if $p < \frac{8}{9}$ then the smallest root is $s = 1$. For $p \geq \frac{8}{9}$ the smallest root is $\frac{3}{4} - \frac{1}{4p}\sqrt{p(9p-8)}$ so $\tau(p) = \frac{3}{4} + \frac{1}{4}\sqrt{(9-\frac{8}{p})}$.

$\square$

A generalization of theorem 3.3 to any N-ary tree was done by [8]. The proof is very similar to the binary case and it is a recommended exercise.

Theorem 3.5. *(Pakes and Dekking, 1991) Let $\tau(N)$ be the probability that the X - process Z has a N-ary splitting, then $1-\tau(N)$ is the smallest fixed point in $[0,1]$ of the function $f_N(s) = \sum_{j=0}^{N-1}(1-s)^j \frac{G^{(j)}(s)}{j!}$.*

4. MANDELBROT'S FRACTAL PERCOLATION

All results that will be presented in this chapter are from the paper [3] of Chayes, Chayes and Durrett from 1988 on Mandelbrot's fractal percolation process.

The model: Fix a natural number $N>1$ and some $p\in[0,1]$ as parameters. Let $A_0 = [0,1]^2$ and for $1\le i,j\le N$ let $S_{ij} = [\frac{i-1}{N},\frac{i}{N}]\times[\frac{j-1}{N},\frac{j}{N}]$. Let $\varepsilon_{i,j}\sim B(p)$ be i.i.d. Bernoulli with parameter p. If $\varepsilon_{ij}=1$ we say that S_{ij} is *occupied* and write $A_1 = \bigcup_{\varepsilon_{ij}=1} S_{ij}$. Assume A_{n-1} have been constructed, let $S^n_{ij} = [\frac{i-1}{N^n},\frac{i}{N^n}]\times[\frac{j-1}{N^n},\frac{j}{N^n}]$, $1\le i,j\le N^n$. Let $\varepsilon^n_{i,j}\sim B(p)$ be i.i.d. Bernoulli with parameter p. We define $A_n = A_{n-1}\cap(\bigcup_{\varepsilon^n_{ij}=1} S^n_{ij})$.

$A_0, A_1, \ldots$ is a decreasing sequence of compact sets. Call the limit $A_\infty = \bigcap_n A_n$. We are interested in properties of A_∞.

The probably first question to ask is "When is $A_\infty \ne \emptyset$?".

Theorem 4.1. $\mathbb{P}(A_\infty\ne\emptyset)>0$ *if and only if* $p>\frac{1}{N^2}$.

Proof. Let Z_n be the number of squares of the form B^n_{ij} which are contained in A_n. Note that Z_n is a Galton-Watson process with offspring distribution $Binom(p,N^2)$.

By Corollary 2.2, $\mathbb{P}(Z_n>0 \text{ for all } n)>0$ if and only if $pN^2>1$. Note that $Z_n > 0$ for all n if and only if $A_n\ne\emptyset$ for all n. Now, since A_n are decreasing compact set, their intersection A_∞ is not empty if and only if $A_n\ne\emptyset$ for all n. Hence $\mathbb{P}(A_\infty\ne\emptyset)=\mathbb{P}(A_n\ne\emptyset \text{ for all } n)=\mathbb{P}(Z_n>0 \text{ for all } n)>0$ if and only if $p>\frac{1}{N^2}$.

□

What about connectivity properties? We first understand when does A_∞ avoid "inner lines":

Theorem 4.2. *Assume that x is* not *of the form $\frac{m}{N^n}$.* $\mathbb{P}(A_\infty\cap(\{x\}\times[0,1])=\emptyset)=1$ *if and only if $p\le\frac{1}{N}$.*

Proof. The number Y_n of interval of the form $\{x\}\times[\frac{(j-1)}{N^n},\frac{j}{N^n}]$ contained in A_n is a Galton-Watson branching process with expected offspring Np. The rest of the proof is as in the proof of the last theorem.

□

Corollary 4.3. *If $p \leq \frac{1}{N}$ then largest connected component of A_∞ is a point.*

Proof. Let $(x,y) \in A_\infty$. If either x or y are not in the form $\frac{m}{N^n}$ we are done by theorem 4.2. Otherwise, if there is a different another $(x',y') \in A_\infty$ in the same component of (x,y), then we can find $(x'',y'') \in A_\infty$ in the same component (perhaps $(x',y') = (x'',y'')$) s.t. either x or y are not of the form form $\frac{m}{N^n}$, contradicting theorem 4.2. □

Actually, even more is achieved by changing the value of x.

Theorem 4.4. *If $p \leq \frac{1}{\sqrt{N}}$ then the largest connected component of A_∞ is a point.*

Proof. Call a segment $[\frac{j-1}{N^n}, \frac{j}{N^n}] \times \left\{\frac{1}{N}\right\}$ *vacant* if either either one of the two adjacent squares in the n-th subdivision is. It is called *occupied* if it is not vacant. Let Y_n be the number of occupied segments of the form $[\frac{j-1}{N^n}, \frac{j}{N^n}] \times \left\{\frac{1}{N}\right\}$. Y_n is a branching process with expected offspring p^2N. Hence Y_n dies out a.s. if $p \leq \frac{1}{\sqrt{N}}$, and and so there will be a.s. a path *arbitrarily closed* to $[0,1] \times \left\{\frac{1}{N}\right\}$ in $A^c_\infty \cup (\textit{squares corners})$. Fix n and j, the same argument shows that there will be a.s. a path arbitrarily closed to $[0,1] \times \left\{\frac{j}{N^n}\right\}$ in $A^c_\infty \cup (\textit{squares corners})$ and by symmetry also for the lines $\left\{\frac{j}{N^n}\right\} \times [0,1]$. After waiting a finite time, all the relevant cornered squares will become vacant. Therefore, as in the last corollary, each path between two points necessarily intersects such a path, contradiction. Hence a.s. the largest component is a point. □

We have seen that for small p the set is totally disconnected. Is it possible that there is some $p < 1$ for which it is not the case? Here are some extra notation and definitions. Let $B_n = \{x \in A_n : x \text{ can be connected to both } \{0\} \times [0,1] \text{and } \{1\} \times [0,1]\}$, and let $B_\infty = \bigcap_n B_n$. For $x \in B_\infty$ let $C_n(x)$ be the component of B_n containing x. Let $\Gamma(x) = \bigcap_n C_n(x)$ then it has an intersection with $\{0\} \times [0,1]$ and $\{1\} \times [0,1]$ as all $C_n(x)$ do. It is also connected as an intersection of a decreasing sequence of connected compact Hausdorff spaces. If $[B_\infty \neq \emptyset]$ occur we say that there is a *left-to-right crossing* of $[0,1]^2$. Let $p_c(N) = \inf\{p \geq 0 : \mathbb{P}(B_\infty \neq \emptyset) > 0\}$. Note that for all $p < 1$ the expected Lebesgue measure $|A_n|$ decreases to zero exponentially fast, so it could hint that one *might not* expect to find $p < 1$ for which there is a positive probability for such a "long" path. But, well, we have the following.

Theorem 4.5. *$p_c(N) < 1$ for all $N > 1$.*

Proof. Assume that $N \geq 3$ (the case $N = 2$ will follow by a modification of $N = 4$). The basic observation, due to Mandelbrot, is that if $N^2 - 1$ of the N^2 are occupied, then any two adjacent squares in the nth level, must have adjacent occupied boundary children squares, which implies $B_n \neq \emptyset$. As before, if we let Z_n to be the number of squares in the n-th stage, then it is a Galton-Watson process with offspring distribution $\sim Binom(N^2, p)$. Hence for achieving left to right crossing it suffices if Z has (N^2-1)-ary splitting.

In other words, we wish to show that $\tau(N^2-1)$, the probability that Z has (N^2-1)-ary splitting, is positive for some $p < 1$, then by theorem 3.5 on the general case, $1-\tau(N^2-1)$ is the smallest fixed point in $[0,1]$ of $f_{N^2-1}(s) = \sum_{j=0}^{N^2-2}(1-s)^j \frac{G^{(j)}(s)}{j!}$, where $G(s) = (1-p+ps)^{N^2}$ is the generating offspring function.

As $G^{(j)}(s) = \frac{N^2!}{(N^2-j)!}p^j(1-p+ps)^{N^2-j}$, we have

$$\begin{aligned}
f(s) &= f_{N^2-1,p}(s) \\
&= \sum_{j=0}^{N^2-2}(1-s)^j \frac{N^2!}{(N^2-j)!j!}p^j(1-p+ps)^{N^2-j} \\
&= \sum_{j=0}^{N^2-2}\binom{N^2}{j}(p(1-s))^j(1-p+ps)^{N^2-j} \\
&= 1-(p(1-s))^{N^2} - N^2(p(1-s))^{N^2-1}(1-(p(1-s)) \\
&= \mathbb{P}[Bin(N^2,(1-s)p) < N^2-1]
\end{aligned}$$

Therefore,

$$1-f(s) \quad = (p(1-s))^{N^2} + N^2(p(1-s))^{N^2-1}(1-(p(1-s))$$

Since $f(0) > 0$ for all p, also $1-f(0) < 1-0$ for all p. So, in order to find ($p < 1$ for which there is) $s < 1$ with $1-f(s) = 1-s$ it is enough to find $s < 1$ for which $1-f(s) > 1-s$.

Let $M > 3$. Using Tylor expansion around 0, for $0 < x$ we may write $(1-x)^M = 1 - Mx + \frac{M(M-1)}{2}x^2 - \frac{M(M-1)(M-2)}{2\cdot 3}x^3 + \frac{M(M-1)(M-2)(M-3)}{2\cdot 3\cdot 4}(1-c)^{M-4}x^4$, where $0 < c < x$ is some intermediate point. The last summand is positive for $0 < x < 1$, and therefore, if $\frac{M(M-1)}{2}x^2 - \frac{M(M-1)(M-2)}{2\cdot 3}x^3 > 0$, or equivalently, if $\frac{3}{M-2} > x$, then

$$(1-x)^M \geq 1-Mx. \tag{4.1}$$

On the other hand $(1-x)^M = 1-Mx+\frac{M(M-1)}{2}x^2 - \frac{M(M-1)(M-2)}{2\cdot 3}(1-c)^3x^3$, for another $0 < c < x$. The last summand is negative for $0 < x < 1$, therefore

$$(1-x)^M \leq 1-Mx+\frac{M(M-1)}{2}x^2. \tag{4.2}$$

Using equation (4.1) and equation (4.2), we have for $s < \frac{3}{N^2-2}$

$$\begin{aligned}
1-f(s) &= (p(1-s))^{N^2} + N^2(p(1-s))^{N^2-1}(1-(p(1-s)) \\
&= p^{N^2}(1-s)^{N^2} + N^2 p^{N^2-1}(1-s)^{N^2-1} - N^2 p^{N^2}(1-s)^{N^2} \\
&= N^2 p^{N^2-1}(1-s)^{N^2-1} - (N^2-1)p^{N^2}(1-s)^{N^2} \\
&\geq N^2 p^{N^2-1}(1-(N^2-1)s) - (N^2-1)p^{N^2}(1-N^2 s + \frac{N^2(N^2-1)}{2}s^2) \\
&= [p^{N^2-1}(N^2(1-p)+p)] - [N^2(N^2-1)p^{N^2-1}(1-p)]s - [\frac{N^2(N^2-1)^2}{2}p^{N^2}]s^2 \\
&\equiv a - bs - cs^2.
\end{aligned}$$

Note that as $p \to 1$ then $a \to 1$, $b \to 0$ and $c \to \frac{N^2(N^2-1)^2}{2}$. Therefore, for $p = 1$, for all $s \in (0, \frac{1}{N^2(N^2-1)^2})$, $1 - \frac{N^2(N^2-1)^2}{2}s^2 > 1-s$. By continuity of f with respect to p and s there are $\delta, \delta' > 0$ small enough so that if $1-\delta < p < 1$ then for all $s \in (\delta', \frac{2}{N^2(N^2-1)^2} - \delta') \neq \emptyset$ $1-f(s) > 1-s$. In particular there is $p < 1$, for which there is some $s < 1$ s.t. $1-f(s) > 1-s$, and we are done.

□

Let $p_d(N) = \sup\{p : \mathbb{P}_p(\text{largest connected component of } A_\infty \text{is a point}) = 1\}$ and remember $p_c = \inf\{p : \mathbb{P}_p(A_\infty \text{ has a left-to-right crossing}) > 0\}$. Clearly, $P_d(N) \leq P_c(N)$. What we have seen thus far is $0 < \frac{1}{\sqrt{N}} \leq P_d(N) \leq P_c(N) < 1$. In [3] it is shown that actually $P_d(N) = P_c(N)$ and $\mathbb{P}_{p_c(N)}(A_\infty$ has a left-to-right crossing$) > 0$ unlike the "usual" percolation on $\mathbb{Z}^2$. Moreover, they show that if we put an independent copy of A_∞ on each square of $\mathbb{Z}^2$, then $p_c(N)$ is actually the critical value of p for which there is an unbounded connected component. This means that there is a remarkable phase transition for this model: there is a point for which for all p smaller the maximal connected component is a point, for all p larger there is a unique unbounded component with positive probability, and this probability as a function of p has discontinuity at the critical point.

APPENDIX A. UNIQUE PHASE TRANSITION

Remember that the event $[B_\infty \neq \emptyset]$ means that there is a left-to-right crossing in A_∞.

Theorem A.1. *There is $\varepsilon_0 > 0$ so that if $\mathbb{P}(B_n \neq \emptyset) \leq \varepsilon_0$ then $\mathbb{P}(B_\infty \neq \emptyset) = 0$ and furthermore, the largest connected component is a point.*

Denote $\Omega_{n,K}$ the event that there is a left to right crossing of $[0,1] \times [0,K]$ when independent copies of A_n are place in each of the squares $[0,1] \times [k-1,k]$, $1 \leq k \leq K$.

Here are two lemmas that will help us to prove Theorem A.1. The proof of these lemmas uses a verity of ideas. Among them are Harris/FKG inequality and Russo-Seymour-Welsh theory. The proof is essentially a careful fine tuning where needed, of the estimations of rectangle's left-to-right crossings probabilities for site percolation process on $\mathbb{Z}^2$. Proof and references can be found in Chayes, Chayes and Durrett [3].

Lemma A.2. *If* $\mathbb{P}(B_n \neq \emptyset) \leq \varepsilon$ *then* $\mathbb{P}(\Omega_{n,K}) \leq f_K(\varepsilon)$*, for some* $f_K(\varepsilon)$ *with* $f_K(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Lemma A.3. *If* $\mathbb{P}(\Omega_{n,2}) \leq 0.01$ *then* $\mathbb{P}(\Omega_{n+m,2}) \leq \frac{1}{25} e^{-N^{m-1}}$.

Proof. (of Theorem A.1): Let ε_0 for which $f_2(\varepsilon) \leq 0.01$ for all $\varepsilon \leq \varepsilon_0$. Assume that n is such that $\mathbb{P}(B_n \neq \emptyset) \leq \varepsilon_0$. Note that $[B_n \neq \emptyset] = \mathbb{P}(\Omega_{n,1})$ and that $\mathbb{P}(\Omega_{m,K})$ increases with K and decreases with m. By lemma A.2, $\mathbb{P}(\Omega_{n,2}) \leq f_2(\varepsilon) \leq 0.01$, and so by lemma A.3, $\mathbb{P}(\Omega_{n+m,2}) \leq \frac{1}{25} e^{-N^{m-1}}$. This implies that $\mathbb{P}(B_m \neq \emptyset) = \mathbb{P}(\Omega_{m,1}) \to 0$ as $m \to \infty$. Indeed, let $\delta > 0$. Choose m large enough so that $\frac{1}{25} e^{-N^{m-1}} < \delta$. Then for all $n' = n + m' \geq n + m$, $\mathbb{P}(B_{n'} \neq \emptyset) = \mathbb{P}(\Omega_{n+m',1}) \leq \mathbb{P}(\Omega_{n+m',2}) \leq \frac{1}{25} e^{-N^{m'-1}} \leq \frac{1}{25} e^{-N^{m-1}} < \delta$.

Using lemma A.2 again, we have that for all $K = 1, 2, \ldots$ $\mathbb{P}(\Omega_{n,K}) \to 0$ as $n \to \infty$. Therefore with probability one, for all m and $j = 1, 2, \ldots, N^m$ there is a "crack" in $[\frac{j-1}{N^m}, \frac{j}{N^m}] \times [0,1]$, i.e. a a curve from bottom to top in $A_\infty \cap [\frac{j-1}{N^m}, \frac{j}{N^m}] \times [0,1]$. As we saw before this implies that a.s. the maximal connected component of A_∞ is a point (and in particular $\mathbb{P}(B_\infty \neq \emptyset) = 0$) . ☐

The following are two consequences of Theorem A.1.

Corollary A.4. $p_d(N) = p_c(N)$

Proof. If $p < p_c(N)$, then $\mathbb{P}_p(B_\infty \neq \emptyset) = 0 < \varepsilon_0$. Therefore, there exists n so that $\mathbb{P}_p(B_n \neq \emptyset) < \varepsilon_0$. By Theorem A.1, the largest connected component of A_∞ is a point $\mathbb{P}_p$-a.s., implying that $p \leq p_d(N)$. Therefore, $p_c(N) \leq p_d(N)$ and we are done. ☐

Fact A.5. *Let* $f_n : [0,1] \to \mathbb{R}$ *be a point-wise decreasing sequence of continuous increasing functions. Assume* $f : [0,1] \to \mathbb{R}$ *is a point-wise limit function of* f_n*. Then* f *is a right continuous increasing function.*

Proof. First we will show that f is increasing. Indeed, let $x < y$ and let $\eta > 0$. We will show that $f(x) - f(y) < \eta$. Since η is arbitrarily small we may deduce that $f(x) - f(y) \leq 0$ as wanted. Take n so large that both $|f(x) - f_n(x)|, |f(y) - f_n(y)| \leq \frac{\eta}{3}$, then $f(x) - f(y) = f(x) - f_n(x) + f_n(x) - f_n(y) + f_n(y) - f(y) \leq \frac{\eta}{3} + 0 + \frac{\eta}{3} < \eta$.

Fix $x_0 \in [0,1)$ and $\varepsilon > 0$. We wish to find $\delta > 0$ so that if $x \in (x_0, x_0 + \delta)$ then $f(x_0) - f(x) < \varepsilon$. Let n_0 large enough so that for all $n \geq n_0$, $(0 \leq)$ $f_n(x_0) - f(x_0) < \frac{\varepsilon}{3}$. Let $\delta > 0$ such that $0 \leq f_{n_0}(x_0) - f_{n_0}(x) \leq \frac{\varepsilon}{3}$ for all $x \in (x_0, x_0 + \delta)$ (can take it even for all $x \in (x_0 - \delta, x_0 + \delta)$, but it's more than needed (the right inequality is since f_{n_0} is increasing in x and $x > x_0$). We claim that $\delta$is as wanted. Indeed, fix $x \in (x_0, x_0 + \delta)$. Let $n \geq n_0$ so large that $|f_n(x) - f(x)| < \frac{\varepsilon}{3}$. $0 \leq f(x) - f(x_0) = (f(x) - f_n(x)) + (f_n(x) - f_n(x_0)) + (f_n(x_0) - f(x_0)) \leq \frac{\varepsilon}{3} + 0 + \frac{\varepsilon}{3} < \varepsilon$. ☐

Note that we did *not* use left continuity of the f_n's

Corollary A.6. $\mathbb{P}_{p_c(N)}(B_\infty \neq \emptyset) > 0$. *In particular,* $p_c(N)$ is a discontinuity point of $p \mapsto \mathbb{P}_p(B_\infty \neq \emptyset)$.

Proof. Since for all p, $\mathbb{P}_p(B_n \neq \emptyset) \searrow \mathbb{P}_p(B_\infty \neq \emptyset)$, if $\mathbb{P}_{p_c(N)}(B_\infty \neq \emptyset)$ would be zero, then let n large enough so that $\mathbb{P}_{p_c(N)}(B_n \neq \emptyset) < \varepsilon_0$. Now using continuity of $p \mapsto \mathbb{P}_p(B_n \neq \emptyset)$ (it is a polynomial), let $p > p_c(N)$ so that still $\mathbb{P}_p(B_n \neq \emptyset) < \varepsilon_0$, then by Theorem A.1, $\mathbb{P}_p(B_\infty \neq \emptyset) = 0$, contradicting the definition of $p_c(N)$. Note that as a point-wise limit of a decreasing sequence of continuous increasing functions, $p \mapsto \mathbb{P}_p(B_\infty \neq \emptyset)$ is in fact right continuous, by Fact A.5. Hence, $\mathbb{P}_{p_c(N)}(B_\infty \neq \emptyset) = \inf\left\{\mathbb{P}_p(B_\infty \neq \emptyset) : p > p_c(N)\right\} \geq \varepsilon_0$. But by the last corollary, $\mathbb{P}_p(B_\infty \neq \emptyset) = 0$ for all $p < p_c(N)$. □

We end with a fact without a proof.

Theorem A.7. *(Chayes-Chayes 1989 [2])* $p_c(N) \to p_c(\mathbb{Z}^2)$ *as* $N \to \infty$, *where* $p_c(\mathbb{Z}^2)$ *is the critical probability for* **site** *percolation on* $\mathbb{Z}^2$.

References

[1] K.B. Athreya and P.E. Ney. *Branching Processes (Die Grundlehren der Mathematischen Wissenschaften 196)*. Springer, Berlin, 1972.

[2] JT Chayes and L. Chayes. The large-n limit of the threshold values in mandelbrot's fractal percolation process. *Journal of Physics A: Mathematical and General*, 22:L501, 1989.

[3] JT Chayes, L. Chayes, and R. Durrett. Connectivity properties of mandelbrot's percolation process. *Probability theory and related fields*, 77(3):307–324, 1988.

[4] FM Dekking. Branching processes that grow faster than binary splitting. *The American mathematical monthly*, 98(8):728–731, 1991.

[5] R. Durrett. *Probability: theory and examples*. Cambridge Univ Pr, 2010.

[6] G. Grimmett and D. Stirzaker. *Probability and random processes*. Oxford University Press, USA, 2001.

[7] B.B. Mandelbrot. *The fractal geometry of nature*. Wh Freeman, 1983.

[8] A.G. Pakes and FM Dekking. On family trees and subtrees of simple branching processes. *Journal of Theoretical Probability*, 4(2):353–369, 1991.

[9] H.W. Watson and F. Galton. On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4:138–144, 1875.