Probability on Trees: An Introductory Climb

Yuval Peres

Lectures at the Saint Flour Summer School, July 1997
Notes prepared with Dimitris Gatzouras and David Levin

March 10, 1999

1 Preface

These notes are based on lectures delivered at the Saint Flour Summer School in July 1997. The first version of the notes was written and edited by Dimitris Gatzouras. The notes were then expanded and revised by David Levin and myself. I hope that they are useful to probabilists and graduate students as an introduction to the subject; a more complete account is in the forthcoming book co-authored with Russell Lyons.

The first 10 chapters are devoted to basic facts about percolation on trees, branching processes and electrical networks, with an emphasis on several key techniques: moment estimates, the use of percolation to determine dimension, and the "method of random paths" to construct flows of finite energy. These 10 chapters are the "introductory climb" alluded to in the title.

More advanced topics start in Chapter 11, where the method of random paths is refined in order to establish the Grimmett-Kesten-Zhang Theorem: *Simple random walk on the infinite percolation cluster in* $\mathbf{Z}^d$, $d \geq 3$ *is transient.*

Chapters 12 and 13 contain a regularity property of subperiodic trees, and its application to random walks on groups. In Chapter 14 we discuss capacity estimates for hitting probabilities; these are used in Chapter 15 to derive intersection-equivalence of fractal percolation and Brownian paths.

In Chapter 16 we analyze the phase transition in a broadcasting model considered by computer scientists: A random bit is propagated, with errors, from the root of a tree to its boundary, and the goal is to reconstruct the original bit from the boundary values. Remarkably, the same model arose independently in genetics, as a mutation model, and in mathematical physics, where it is equivalent to the Ising model on a tree. In Chapter 17, the Ising model on a tree is used to construct a nearest-neighbor process on $\mathbf{Z}$ that is "less predictable" than simple random walk.

In Chapters 18 and 19, we study the speed and recurrence properties of tree-indexed processes; in particular, we relate three natural notions of speed (cloud speed, burst speed, and sustainable speed) to three well-known dimension indices (Minkowski dimension, packing dimension, and Hausdorff dimension). In Chapter 20 we consider a dynamical variant of percolation, where edges open and close according to independent Poisson processes. At any fixed time, the random configuration is a sample of Bernoulli percolation, but we focus on exceptional random times when the number of infinite open clusters is atypical. There are striking parallels between the study of these exceptional times for dynamical percolation, and the study of multiple points for Brownian motion. We conclude in Chapter 21 by describing some results on stochastic domination between randomly labeled trees, and stating some open problems for other graphs.

I was first drawn to thinking about general trees in a lecture of I. Benjamini in 1989, when H. Furstenberg noted that certain trees that appeared in the lecture could be interpreted (via b-adic expansions) as Cantor sets with different Hausdorff and Minkowski dimensions. I. Benjamini and I proceeded to examine relations between properties of trees and properties of the corresponding compact sets; these connections had unexpected uses later (see Chapter 15). For example, consider a subset Λ of the unit square in the plane and the corresponding tree $T(\Lambda, b)$ in base b. Then Λ is hit by planar Brownian motion (i.e., it has positive logarithmic capacity) iff simple random walk on $T(\Lambda, b)$ is transient.

We then learned that a year earlier, R. Lyons (building on works of Furstenberg, Shepp, Kahane and Fan) had established some remarkably precise connections between random walks, percolation and capacity on trees. R. Lyons and R. Pemantle had already used these ideas to determine the *sustainable speed* of first-passage percolation on trees.

The point of view of these lectures was largely developed in the ensuing collaboration with Itai Benjamini, Russell Lyons and Robin Pemantle, whose influence pervades these notes. Other coauthors whose insights and ideas are represented here include Chris Bishop (see Chapter 15), Will Evans, Claire Kenyon, and Leonard Schulman (see Chapter 16), Olle Häggström and Jeff Steif (see Chapter 20).

In fact, probability on trees is a rich and fast-growing subject, so the account presented in these notes is necessarily incomplete. Natural complements are the two conference proceedings volumes: *Trees*, edited by B. Chauvin, S. Cohen and A. Rouault (Birkhäuser 1996) and *Classical and Modern Branching Processes*, edited by K. B. Athreya and P. Jagers (Springer 1996). Continuum random trees are fascinating objects studied in several papers by David Aldous; Tom Liggett is writing a detailed account of the contact process on trees. Superprocesses, which can be obtained as scaling limits of branching random walks, have been studied by numerous authors. I apologize to the many researchers whose results involving probability on trees are not described here.

Acknowledgements I am grateful to the participants in the St. Flour summer school for their comments and to the organizer, Pierre Bernard, for his warm hospitality.

I am greatly indebted to Dimitris Gatzouras and David Levin for their help in preparing these notes. I thank Itai Benjamini, Dayue Chen, Amir Dembo, Nina Gantert, David Grabiner, Olle Häggström, Davar Khoshnevisan, Elon Lindenstrauss, Elhanan Mossel, Oded Schramm and Balint Virag for their comments on the manuscript.

Yuval Peres
Jerusalem, December 1998

Contents

2 Basic Definitions and a Few Highlights

A **tree** is a connected graph containing no cycles. All trees considered in these notes are locally finite: the degree $\deg(v)$ is finite for each vertex v, although $\deg(v)$ may be unbounded as a function of v.

Why study general trees?

1. More can be done on trees than on general graphs. Percolation problems, for example, are easier to analyze on trees. The insight and techniques developed for trees can sometimes be extended to more general models later.

2. Trees occur naturally. Some examples are:

   (a) **Galton-Watson trees**. Let L be a non-negative integer-valued random variable and set $Z_0 \equiv 1$, $Z_1 = L$, and $Z_{n+1} = \sum_{i=1}^{Z_n} L_i^{(n+1)}$, where the $L_i^{(n)}$ are i.i.d. copies of L. Then Z_n is the number of individuals in generation n of a Galton-Watson branching process, a population which starts with one individual and in which each individual independently produces a random number of offspring with the same distribution as L. The collection of all individuals form the vertices of a tree, with edges connecting parents to their children.

   (b) **Random spanning trees in networks**. A *spanning tree* of a graph G is a tree which is a subgraph of G including all the vertices of G. There are several interesting algorithms for generating random spanning trees of finite graphs.

3. Trees describe well the complicated structure of certain compact sets in $\mathbf{R}^d$. Examples include Cantor sets on intervals and fractal percolation, a collection of nested random subsets of the unit cube described below.

Example 2.1 Fractal Percolation is a recursive construction generating random subsets $\{A_n\}$ of the unit cube $[0,1]^d$. Tile $A_0 = [0,1]^d$ by b^d similar subcubes with side-length b^{-1}. Generate A_1 by taking a union of some of these subcubes, including each independently with probability p. In general, A_n will be a union of b-adic cubes of order n (cubes with side-length b^{-n} and vertices with coordinates of the form kb^{-n}). A_{n+1} is obtained by tiling each such cube contained in A_n by b^d b-adic subcubes of order $n+1$, and taking a union which includes each subcube independently with probability p. The limit set of this construction $\bigcap_{n=0}^{\infty} A_n$ is denoted by $Q_d(p)$.

There is a tree associated with each realization of fractal percolation. The vertices at level n correspond to b-adic cubes of order n which are contained in A_n, and a vertex v at level n is the parent of a vertex w at level $n+1$ if the cube corresponding to v contains the cube corresponding to w. $\triangle$

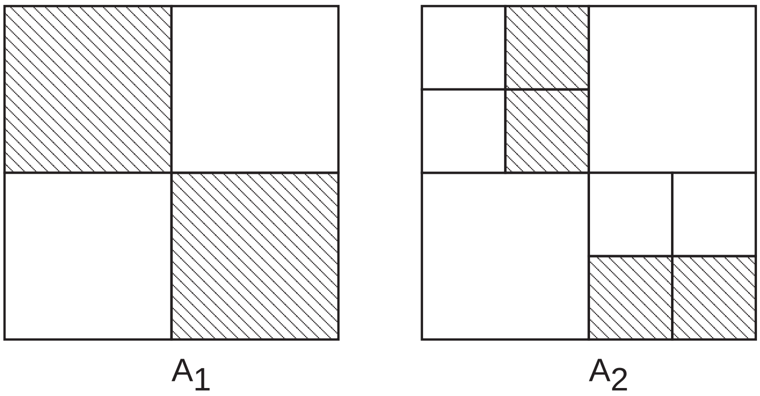


Figure 1: **A realization of A_1 and A_2 for $d = 2$, $b = 2$.**

Let $Q_3(\frac{1}{2}) \subset [0,1]^3$ denote the limit set of fractal percolation with $b = 2$, $d = 3$, and $p = \frac{1}{2}$. In Chapter 15, we will see that the random set $Q_3(\frac{1}{2})$ is **intersection-equivalent** in the cube to the Brownian motion path started uniformly in the cube. By this we mean the following: if $[B]$ denotes the range $\{B(t) : t \geq 0\}$ of a three-dimensional Brownian motion started uniformly in $[0,1]^3$, then for some constants $C_0, C_1 > 0$ and all closed sets $\Lambda \subset [0,1]^3$,

$$C_0 \, \mathbf{P}(Q_3(1/2) \cap \Lambda \neq \emptyset) \;\leq\; \mathbf{P}([B] \cap \Lambda \neq \emptyset) \;\leq\; C_1 \, \mathbf{P}(Q_3(1/2) \cap \Lambda \neq \emptyset) \, .$$

Consequently, hitting probabilities for Brownian motion can be related to hitting probabilities of $Q_3(\frac{1}{2})$. This gives a new perspective on the classical study of intersections and multiple points of Brownian paths.

For example, consider two independent copies $Q_3(\frac{1}{2})$ and $Q_3'(\frac{1}{2})$. Then the intersection $Q_3(\frac{1}{2}) \cap Q_3'(\frac{1}{2})$ has the same distribution as $Q_3(\frac{1}{4})$. Since the tree corresponding to $Q_3(\frac{1}{4})$ is a Galton-Watson tree with mean offspring 2, it survives with positive probability. Hence $Q_3(\frac{1}{4}) \neq \emptyset$ with positive probability, and intersection-equivalence shows that two independent Brownian paths in $\mathbf{R}^3$ intersect with positive probability, a result first proved in [21].

It also follows that three Brownian paths in space do not intersect (as first proved in [22]). By intersection-equivalence, it is enough to show that the intersection of the limit sets of three independent fractal percolations, which has the same distribution as $Q_3(\frac{1}{8})$, is empty a.s. But the tree corresponding to $Q_3(\frac{1}{8})$ is a critical Galton-Watson process and hence dies out, see Chapter 3.

Infinite family trees arising from supercritical Galton-Watson Branching processes, (Galton-Watson trees in short) play a prominent role in these notes.

Question 2.2 *In what ways are Galton-Watson trees like regular trees?*

First we establish a simple property of regular trees.

Example 2.3 Simple random walk $\{X_n\}_{n\geq 0}$ on a graph is a Markov chain on the vertices, with transition probabilities

$$\mathbf{P}(X_{n+1}=w|X_n=v)=\begin{cases}\frac{1}{\deg(v)} & \text{if } w\sim v\,,\\ 0 & \text{otherwise}\,.\end{cases}$$

The notation $u\sim v$ means that the vertices u and v are connected by an edge. Now suppose the graph is a tree, and let $|v|$ stand for the distance of a vertex v from the root ρ, *i.e.*, $|v|$ is the number of edges on the unique path from ρ to v. On the b-ary tree,

$$\mathbf{E}\Big[|X_{n+1}|-|X_n|\;\Big|\;X_n\Big]\geq\frac{b}{b+1}(+1)+\frac{1}{b+1}(-1)=\frac{b-1}{b+1}\,.$$

(We have an inequality here because X_n may be at the root.) Hence the distance of the random walk on the tree from the root stochastically dominates an upwardly biased random walk on $\mathbf{Z}$. It is therefore transient and will visit 0 only finitely many times. After the last visit of the random walk to the root,

$$\mathbf{E}\Big[|X_{n+1}|-|X_n|\;\Big|\;X_n\Big]=\frac{b-1}{b+1}\,,$$

and the strong law of large numbers for martingale differences implies that, almost surely, $n^{-1}|X_n|\to\frac{b-1}{b+1}$ $\triangle$

One specific case of Question 2.2 is

Question 2.4 *On a Galton-Watson (GW) tree with mean $m=\sum_k kp_k>1$, is simple random walk transient on survival of the GW process?*

We will see later that the answer is positive; this was first proved by Grimmett and Kesten (1984).

For a tree Γ, denote $\Gamma_n=\{v\colon |v|=n\}$. Define the **lower growth** and **upper growth** of Γ as $\underline{\mathrm{gr}}(T):=\liminf|\Gamma_n|^{1/n}$ and $\overline{\mathrm{gr}}(T)=\limsup|\Gamma_n|^{1/n}$ respectively. If $\underline{\mathrm{gr}}(\Gamma)=\overline{\mathrm{gr}}(\Gamma)$, we speak of the **growth** of the tree Γ and denote it by $\mathrm{gr}(\Gamma)$.

Question 2.5 *Is $\underline{\mathrm{gr}}(\Gamma)>1$ sufficient for transience of simple random walk on Γ? Is it necessary?*

The answer to both questions is negative. An analogous situation holds for Brownian motion on manifolds, where exponential volume growth is not sufficient and not necessary for transience.

Example 2.6 (3–1 tree) The 3-1 tree Γ has $\mathrm{gr}(\Gamma)=2$ (actually $|\Gamma_n|=2^n$), but simple random walk is recurrent on it. Γ can be embedded in the upper half-plane, with its root ρ at the origin. The root has two offspring, and for $n\geq 1$, each level Γ_n

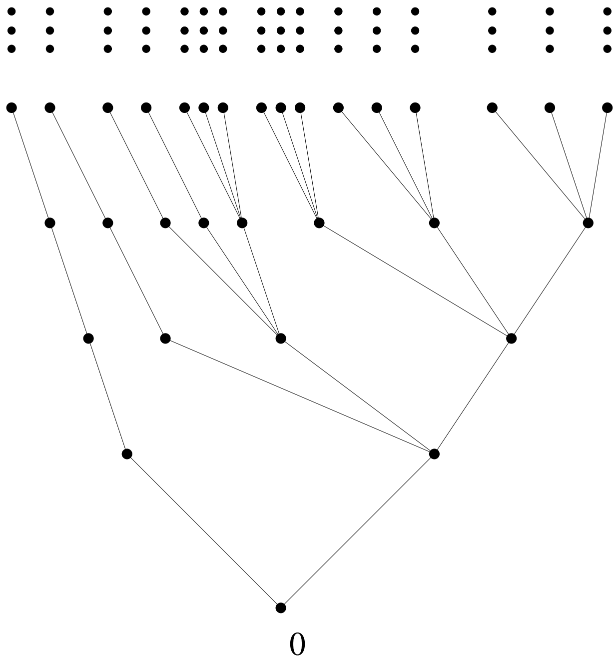


Figure 2: **The 3-1 Tree.**

has 2^n vertices which can be ordered from left to right as $v_1^n, \ldots, v_{2^n}^n$. For $k \leq 2^{n-1}$, each v_k^n has only one child, while for $2^{n-1} < k \leq 2^n$, each v_k^n has three children. Observe that for any vertex not on the right-most path to infinity, the subtree above it will eventually have no more branching (because "powers of 3 beat powers of 2"). The random walk on Γ will have excursions on left-hand branches, but must always return to the right-most branch (because of recurrence of simple random walk on the line). If these excursions are ignored, then we have a simple random walk on the right-most path, i.e., on $\mathbf{Z}^+$, which is recurrent. $\triangle$

It is even easier to construct transient trees of polynomial growth: *E.g.*, replace every edge at level k of the ternary tree by a path consisting of 2^k edges. Simple random walk on the resulting tree, considered just when it visits branch points, dominates an upward biased random walk on the integers, whence it is transient.

On the other hand, positive speed implies exponential growth:

Theorem 2.7 *Define the* **speed** *of a random walk as* $\lim_n n^{-1}|X_n|$, *when this limit exists. If the speed of simple random walk on a tree* Γ *exists and is positive, then* Γ *has exponential growth, i.e.,* $\underline{\mathrm{gr}}(\Gamma) > 1$.

This follows from Theorem 5.4 below.

Example 2.6 suggests that $\mathrm{gr}(\Gamma)$ does not give much information on the behavior of a random walk on Γ. The growth $\mathrm{gr}(\Gamma)$ barely takes into account the structure of Γ, and a more refined notion is required.

A **cutset** Π is a set of vertices such that any infinite self-avoiding path from emanating the root ρ must pass through some vertex in Π. The **branching number** of a tree Γ is defined as

$$\mathrm{br}(\Gamma) = \sup\Big\{ \lambda \geq 1 : \inf_{\Pi\,\mathrm{cutset}} \sum_{v\in\Pi} \lambda^{-|v|} > 0 \Big\}. \tag{1}$$

The function $\inf\Big\{ \sum_{v\in\Pi} \lambda^{-|v|} \ : \ \Pi \ \text{a cutset}\Big\}$ is decreasing in λ and positive at $\lambda = 1$.

The **boundary of a tree** Γ, denoted $\partial\Gamma$, is the set of all infinite self-avoiding paths (rays) emanating from the root ρ of Γ. A natural metric on the boundary $\partial\Gamma$ is $d(\xi, \eta) = e^{-n}$, where n is the number of edges shared by ξ and η. $\dim_H(\partial\Gamma)$ will denote the Hausdorff dimension of $\partial\Gamma$ with respect to this metric d. Because an open cover of $\partial\Gamma$ corresponds to a cutset of Γ, and vice-versa, the Hausdorff dimension of $\partial\Gamma$ is related to the branching number of Γ by

$$\log \mathrm{br}(\Gamma) \ = \ \dim_H(\partial\Gamma)\,.$$

Similarly, $\mathrm{gr}(\Gamma)$ is related to the Minkowski dimension $\dim_M(\partial\Gamma)$ by

$$\log \mathrm{gr}(\Gamma) \ = \ \dim_M(\partial\Gamma)\,.$$

Generally, $\mathrm{br}(\Gamma) \leq \underline{\mathrm{gr}}(\Gamma)$, since for $\lambda > \underline{\mathrm{gr}}(\Gamma)$ we must have

$$\inf_n |\Gamma_n| \lambda^{-n} = \inf_n \sum_{v\in\Gamma_n} \lambda^{-|v|} = 0;$$

using the fact that Γ_n is itself a cutset yields the inequality. If $\partial\Gamma$ is countable, then $\mathrm{br}(\Gamma) = 1$, because $\dim_H A = 0$ for countable sets A. For the 3-1 tree in Example 2.6, $\partial\Gamma$ is countable, and consequently $\mathrm{br}(\Gamma) = 1$.

As an indication that the branching number $\mathrm{br}(\Gamma)$ contains more information about the tree than the growth $\mathrm{gr}(\Gamma)$, we mention two results that we shall prove later, in Chapters 7 and 13.

Bernoulli(p) **percolation** on a tree Γ is the random subgraph of Γ obtained by independently including each original edge of Γ with probability p, and discarding each with probability $1 - p$. The retained edges are called **open**, and $\mathbf{P}_p$ is the probability corresponding to this process (see Chapter 4 for the formal definition of the probability space.) The first quantity of interest in percolation is

$$p_c(\Gamma) = \inf\{p \in [0,1]\colon \mathbf{P}_p(\rho \leftrightarrow \infty) > 0\}\,, \tag{2}$$

where $\{\rho \leftrightarrow \infty\}$ denotes the event that the root ρ is connected to ∞, i.e., that there is an infinite self-avoiding path emanating from ρ, that consists of open edges.

Theorem 2.8 (R. Lyons 1990) *For an infinite and locally finite tree* Γ*,*

$$p_c(\Gamma) = \frac{1}{\mathrm{br}(\Gamma)}. \tag{3}$$

Theorem 2.9 (R. Lyons 1990) *If* $\mathrm{br}(\Gamma) > 1$, *then simple random walk on* Γ *is transient.*

We close with an equivalent description of the branching number $\mathrm{br}(\Gamma)$ of a tree Γ. If u, v are vertices in Γ so that v is a child of u, denote by uv the edge connecting them. A **flow** θ on Γ from the root ρ to ∞ is an edge function obeying $\theta(uv) = \sum \theta(vw)$, where the sum is over all children w of v. This property is known as **Kirchhoff's node law**. Imagine the tree as a network of pipes through which water can flow entering at the root. However much water enters a pipe must leave through the other end, splitting up among the outgoing pipes (edges). Define $\theta(v)$, for a vertex $v \neq \rho$, to be the amount of flow that reaches v, i.e., $\theta(v) := \theta(uv)$ for u the parent of v. The **strength** of a flow θ, denoted $\|\theta\|$, is the amount flowing from the root, $\sum_{v:v\sim\rho} \theta(v)$. When $\|\theta\| = 1$, we call θ a **unit flow**.

Lemma 2.10 *For a tree* Γ,

$$\mathrm{br}(\Gamma) = \sup\{\lambda \geq 1 : \exists \text{ a nonzero flow } \theta \text{ from } \rho \text{ to } \infty \ : \forall v,\ \theta(v) \leq \lambda^{-|v|}\}. \tag{4}$$

Proof. This follows directly from the Min-cut/Max-flow Theorem, which in our setting says that

$$\sup\{\|\theta\| : \theta(v) \leq \lambda^{-|v|} \ \ \forall v\} = \inf_{\Pi \text{ cutset}} \sum_{v\in\Pi} \lambda^{-|v|}. \tag{5}$$

For details, see Lyons and Peres (1999). □

Remark: As mentioned above, $\mathrm{br}(\Gamma) \leq \underline{\mathrm{gr}}(\Gamma) = \liminf_n |\Gamma_n|^{1/n}$. In general, to get an upper bound for $\mathrm{br}(\Gamma)$ one can seek explicit 'good' cutsets. To get lower bounds use either

(i) Theorem 2.8, which in particular says that $\mathrm{br}(\Gamma) \geq 1/p_c(\Gamma)$, or

(ii) find a good flow θ on Γ such that $\theta(v) \leq \lambda^{-|v|}$ for all v; then $\mathrm{br}(\Gamma) \geq \lambda$. (Recall that $\theta(v)$ denotes the flow from the unique parent of v to v.)

A flow θ on Γ induces a measure μ on $\partial\Gamma$: for cylinder sets $[v] = \{\xi \in \partial\Gamma : \xi \text{ passes through } v\}$, define $\mu([v])$ as $\theta(v)$. If $[v_1], \ldots, [v_n]$ are disjoint cylinders (which means that no v_i is an ancestor of another), and $[v] = \bigcup_{i=1}^n [v_i]$ (i.e., the $\{v_i\}$ form a cutset for the subtree Γ^v rooted at v), then Kirchhoff's node law implies (by induction on n) that $\mu([v]) = \sum_{i=1}^n \mu([v_i])$. Countable additivity can be proven using the compactness of $\partial\Gamma$: Cylinders form a basis consisting of open sets and are also closed in the natural topology on $\partial\Gamma$. Thus countable additivity follows from finite additivity.

3 Galton-Watson Trees

Let L be a non-negative integer-valued random variable and let $p_k = \mathbf{P}(L = k)$ for $k = 0, 1, 2, \ldots$. To avoid trivial cases, we assume throughout that $p_1 < 1$. Let $\{L_i^{(n)}\}_{i,n\in\mathbf{N}}$ be independent and identically distributed copies of L, set $Z_0 = 1$, and define

$$Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} L_i^{(n+1)} & \text{if } Z_n > 0\,, \\ 0 & \text{if } Z_n = 0\,. \end{cases}$$

The variables Z_n are the population sizes of a Galton-Watson branching process. The tree associated with a realization of this process has Z_n vertices at level n, and for $i \le Z_n$, the i'th vertex in level n has $L_i^{(n+1)}$ children in level $n+1$.

Generating functions are an indispensable tool in the analysis of Galton-Watson processes. Set $f(s) = \mathbf{E}[s^L]$ and define inductively

$$f_0(s) = s\,, \qquad f_1(s) = f(s)\,, \qquad f_{n+1}(s) = f \circ f_n(s)\,, \qquad 0 \le s \le 1\,.$$

It can be verified by induction that $f_n(s) = \mathbf{E}[s^{Z_n}]$ for all n, that is, f_n is the generating function of Z_n. Note that $f(s) = \sum_{k=0}^{\infty} p_k s^k$ and $f'(1) = \mathbf{E}[L] = m$. We always have $f''(s) \ge 0$ for $s \ge 0$, so f is convex on $\mathbf{R}^+$.

Define q to be the smallest fixed point of f in $[0,1]$. Note that if $p_0 = 0$, then $q = 0$. Observe that $\lim_n \mathbf{P}(Z_n = 0) = \lim_n f_n(0) \le q$, and since $\lim_n f_n(0)$ must be a fixed point of f, it follows that $q = \lim_n \mathbf{P}(Z_n = 0)$. So

$$q = \mathbf{P}(Z_n \to 0) = \text{ probability of extinction.}$$

Since f is convex, if $1 \ge m = f'(1)$, then $q = 1$. If instead $1 < m = f'(1)$, then $q < 1$. Thus, a Galton-Watson process dies out a.s. if and only if $m \le 1$.

A property of trees A is **inherited** if all finite trees have property A, and all the immediate descendant subtrees $\Gamma^{(i)}$ of Γ have A when Γ has A. (The immediate descendant subtrees $\Gamma^{(i)}$ of Γ are the subtrees of Γ rooted at the children of the root ρ.)

Example 3.1 The following are all inherited properties:

1. $\{\Gamma : \sup_n |\Gamma_n| < \infty\}$.
2. $\{\Gamma : |\Gamma_n| \text{ grows polynomially in } n\}$.
3. $\{\Gamma : \Gamma \text{ finite or } \mathrm{br}(\Gamma) \le c\}$. $\triangle$

Proposition 3.2 (0-1 Law) *Let $\mathbf{P}$ be the probability measure on trees corresponding to a GW process with $m > 1$. If A is inherited, then*

$$\mathbf{P}(A \mid \textit{non-extinction}) \in \{0, 1\}.$$

Proof. We have

$$\mathbf{P}(\Gamma \in A | Z_1 = k) \le \mathbf{P}\left(\bigcap_{i=1}^{k} \{\Gamma^{(i)} \in A\} \mid Z_1 = k\right) = \mathbf{P}(\Gamma \in A)^k.$$

Thus,

$$\mathbf{P}(\Gamma \in A) = \sum_k p_k \mathbf{P}(\Gamma \in A | Z_1 = k) \le f(\mathbf{P}(\Gamma \in A)).$$

Convexity of f implies that the only numbers $x \in [0,1]$ satisfying $x \le f(x)$ are $x = 1$ and all $x \in [0, q]$. Since A holds for all finite trees, $\mathbf{P}(\Gamma \in A) \ge q$. So $\mathbf{P}(\Gamma \in A) \in \{q, 1\}$. $\square$

Observe that $m^{-n} Z_n$ is a non-negative martingale and hence converges to some finite random variable $W < \infty$. If $m \le 1$, then $Z_n = 0$ eventually, so a.s. $W = 0$. The case $m > 1$ is treated by the following theorem.

Theorem 3.3 (Kesten and Stigum (1966a)) *When $m > 1$,*

$$\mathbf{P}(W > 0 \mid \textit{non-extinction}) = 1 \quad \textit{if and only if} \quad \mathbf{E}[L \log^+ L] < \infty.$$

A conceptual proof of Theorem 3.3 appears in Lyons, Pemantle, and Peres (1995).

Hawkes (1981), under the assumption that $\mathbf{E}[L \log^2 L] < \infty$, proved that for Galton-Watson trees Γ,

$$\mathbf{P}(\dim_H(\partial\Gamma) = \log m \mid \text{non-extinction}) = 1.$$

This is equivalent to

$$\mathbf{P}(\text{br}(\Gamma) = m \mid \text{non-extinction}) = 1. \tag{6}$$

R. Lyons discovered a simpler proof without the assumption $\mathbf{E}[L \log^2 L] < \infty$, which is given below in Corollary 5.2. Because a.s. $m^{-n} Z_n \to W$, where $0 \le W < \infty$, it follows that a.s. $\overline{\text{gr}}(\Gamma) \le m$. This, together with the general inequality $\text{br}(\Gamma) \le \underline{\text{gr}}(\Gamma)$ and (6), implies that a.s. given non-extinction,

$$m = \text{br}(\Gamma) \le \underline{\text{gr}}(\Gamma) \le \overline{\text{gr}}(\Gamma) \le m.$$

4 General percolation on a connected graph

General (bond) percolation on a connected graph G is a random subgraph $G(\omega)$ of G such that, for any edge e in G, the event that e is an edge of $G(\omega)$ is measurable. **Independent $\{p_e\}$ percolation** is the percolation obtained when each edge e is retained (or declared open) with probability p_e, independently of other edges (and removed or declared closed otherwise). We already discussed in Chapter 2 the special case of **Bernoulli(p) percolation** where all probabilities p_e are the same, $p_e \equiv p$.

Formally, the sample space for a general bond percolation is $\Omega = \{0,1\}^E$, where E is the edge set of the graph G. The σ-field $\mathcal{F}$ on Ω is generated by the finite-dimensional cylinders, sets of the form $\{\omega \in \Omega : \omega(e_1) = x_1, \ldots, \omega(e_m) = x_m\}$ for $x_i \in \{0,1\}$. The probability measures $\mathbf{P}_{\{p_e\}}$ and $\mathbf{P}_p$, corresponding to independent $\{p_e\}$ percolation and Bernoulli(p) percolation respectively, are product measures on $(\Omega, \mathcal{F})$.

We write the event that vertex sets A and B are connected by a path in $G(\omega)$ by $\{A \leftrightarrow B\}$; when G is an infinite tree Γ, we write $\{\rho \leftrightarrow \partial\Gamma\}$ for the event that there is an infinite path emanating from ρ with all edges open.

The connected components of open edges in percolation are called **clusters**, and the cluster containing v is denoted by $\mathcal{C}(v)$. Define

$$\mathcal{C} := \{\exists v \in G \text{ with } |\mathcal{C}(v)| = \infty\};$$

$\mathcal{C}$ is the event that there is an infinite cluster somewhere in the percolation on G. We write $\mathcal{C}_G$ when there is a possibility of ambiguity.

For Bernoulli(p) percolation, at any fixed vertex v,

$$\mathbf{P}_p(|\mathcal{C}(v)| = \infty) > 0 \text{ if and only if } \mathbf{P}_p(\mathcal{C}) = 1. \tag{7}$$

One implication in (7) follows immediately from Kolmogorov's zero-one law: $\mathcal{C}$ does not depend on the status of any finite number of edges, hence $\mathbf{P}_p(\mathcal{C}) \in \{0,1\}$. To see the other implication, assume $\mathbf{P}_p(\mathcal{C}) = 1$ and take a ball $B_n(v)$ large enough so that

$$\mathbf{P}_p(\text{there exists an infinite path intersecting } B_n(v)) \; > \; 0.$$

Then clearly

$$\mathbf{P}_p(\partial B_n(v) \leftrightarrow \infty) \; > \; 0.$$

Because $B_n(v)$ is finite, the event that all edges in $B_n(v)$ are open has positive probability. By independence of disjoint edge sets,

$$\begin{aligned}
\mathbf{P}_p(|\mathcal{C}(v)| = \infty) &\geq \mathbf{P}_p(\text{all edges in } B_n(v) \text{ are open and } \partial B_n(v) \leftrightarrow \infty) \\
&= \mathbf{P}_p(\text{all edges in } B_n(v) \text{ are open})\mathbf{P}_p(\partial B_n(v) \leftrightarrow \infty) \\
&> 0\,.
\end{aligned}$$

Alternatively, one can use the FKG inequality for the events $A = \{\text{all edges in } B_n(v)$ are open$\}$ and $B = \{$there exists an infinite path connecting $B_n(v)$ to $\infty\}$, as both these events are increasing. See Grimmett (1989) for details.

For Bernoulli(p) percolation on an arbitrary graph G, the **critical probability** (already mentioned in the case of trees) is

$$p_c(G) = \inf\{\, p : \mathbf{P}_p(\mathcal{C}) = 1 \,\}\,.$$

For this definition to make sense, $p \mapsto \mathbf{P}_p(\mathcal{C})$ must be non-decreasing. This can be seen by by coupling the measures $\mathbf{P}_p$ for all p together, see Grimmett (1989).

5 The First-Moment Method

The *first moment method* is straightforward but useful. For general percolation on a tree Γ with root ρ, it asserts that

$$\mathbf{P}(\rho \leftrightarrow \infty) \leq \sum_{v \in \Pi} \mathbf{P}(\rho \leftrightarrow v) \tag{8}$$

for any cutset Π. For Bernoulli(p) percolation on the tree, the inequality becomes

$$\mathbf{P}_p(\rho \leftrightarrow \infty) \leq \sum_{v \in \Pi} p^{|v|}\,.$$

When $p < 1/\mathrm{br}(\Gamma)$, this can be made arbitrarily small for appropriate choice of cutset. This proves

Proposition 5.1 *For any locally finite* Γ*,*

$$p_c(\Gamma) \geq \frac{1}{\mathrm{br}(\Gamma)}\,. \tag{9}$$

In general there is equality here, as advertised previously in Theorem 2.8. The proof of equality is in §7.

Corollary 5.2 *Let T be a GW tree with mean $m > 1$. Almost surely on non-extinction,* $\mathrm{br}(T) = m$ *and* $p_c(T) = 1/m$.

Proof. Let $\mathbf{P}_{GW}$ be the distribution of T on the space of rooted trees $\mathcal{T}$, and let $Z_n = |T_n|$ be the size of level n of T. Given $t \in \mathcal{T}$, let $\mathbf{P}_{p,t}$ be Bernoulli(p) percolation on t.

Observe that

$$m \geq \overline{\mathrm{gr}}(T) \geq \underline{\mathrm{gr}}(T) \geq \mathrm{br}(T) \geq \frac{1}{p_c(T)}\,. \tag{10}$$

The first inequality follows since Z_n/m^n converges to a finite random variable, the middle inequalities hold in general, and the right-most is the content of Proposition 5.1. Thus it is enough to show that for $p > m^{-1}$,

$$\mathbf{P}_{GW}\left(t \ : \ \mathbf{P}_{p,t}(|\mathcal{C}(\rho)| = \infty) > 0 \ \middle|\ \text{non-extinction}\right) = 1. \tag{11}$$

Combine the measures $\mathbf{P}_{GW}$ and $\mathbf{P}_{p,t}$: Given the Galton-Watson tree T, perform Bernoulli(p) percolation on T and let T' be the component of ρ in the percolation. T' is itself a Galton-Watson tree, where the number of individuals in the first generation is $Z_1' = \sum_{i=1}^{Z_1} Y_i$, where $\{Y_i\}$ are i.i.d. Bernoulli(p) random variables. Because $\mathbf{E}[Z_1'] = mp > 1$, with positive probability T' is infinite:

$$\mathbf{P}(|T'| = \infty) = \int \mathbf{P}_{p,t}(|\mathcal{C}(\rho)| = \infty\})d\mathbf{P}_{GW}(t) > 0.$$

We conclude that the integrand must be positive with positive $\mathbf{P}_{GW}$-probability:

$$\mathbf{P}_{GW}\left(t \,:\, \mathbf{P}_{p,t}(|\mathcal{C}(\rho)| = \infty) > 0\right) > 0.$$

Since the set

$$\{t \,:\, \mathbf{P}_{p,t}(|\mathcal{C}(\rho)| = \infty) = 0\}$$

defines an inherited property, Proposition 3.2 implies that (11) holds. This proves that a.s. on survival, $p_c(T) = m^{-1}$, whence (10) yields that $\mathrm{br}(T) = m$. $\square$

Kahane and Peyrière (1976) calculated the dimension of the limit set of fractal percolation; their methods were different. The proof above is due to R. Lyons.

Question 5.3 (Häggström) *Suppose simple random walk $\{X_n\}_{n\geq 0}$ on Γ has positive lower speed, i.e., for some positive number s*

$$\mathbf{P}\left(\liminf_n \frac{|X_n|}{n} > s\right) > 0\,. \tag{12}$$

Is it necessarily true that $\mathrm{br}(\Gamma) > 1$*?*

The answer is positive, and the proof relies on the first-moment method again.

Theorem 5.4 *If (12) holds, then* $\mathrm{br}(\Gamma) \geq e^{I(s)/s}$*, where*

$$I(s) = \frac{1}{2}[(1+s)\log(1+s) + (1-s)\log(1-s)]\,.$$

Proof. By (12) above, there exists L such that

$$\mathbf{P}\left(|X_n| > ns \text{ for all } n \geq L\right) > 0\,.$$

Define a general percolation on Γ by

$$\Gamma(\omega) = \left\{\, v \in \Gamma :\, |v| \leq L \text{ or } X_n = v \text{ for some } n < |v|s^{-1}\right\}\,.$$

More precisely, if $e(v)$ denotes the edge from the parent of v to v, we retain $e(v)$ if $|v| \leq L$ or if $X_n = v$ for some $n < |v|s^{-1}$. By the definition of this percolation,

$$\mathbf{P}(\rho \leftrightarrow \infty) \geq \mathbf{P}\left(|X_n| > ns \text{ for all } n \geq L\right) > 0\,. \tag{13}$$

On the other hand, we claim that if S_n is simple symmetric random walk on $\mathbf{Z}$, then for $|v| > L$,

$$\mathbf{P}(\rho \leftrightarrow v) = \mathbf{P}(X_n = v \text{ for some } n < |v|s^{-1}) \leq \mathbf{P}\left(\max_{n<|v|s^{-1}} |S_n| \geq |v|\right)\,. \tag{14}$$

Consider a particle on Γ which moves with X when X moves along the unique path from ρ to v, but remains stationary during excursions (possibly infinite) of X from this path. This particle performs a simple random walk on the path with (possibly infinite) holding times between moves. The probability on the left in (14) is the chance that this particle reaches v before time $|v|s^{-1}$, which is at most the chance that simple random walk on $\mathbf{Z}$ travels distance $|v|$ from the origin in the same time. This proves (14).

By the reflection principle,

$$\mathbf{P}\left(\max_{n\le N}|S_n|\ge sN\right)\le 2\,\mathbf{P}\left(\max_{n\le N}S_n\ge sN\right)\le 4\,\mathbf{P}(S_N\ge sN)\le 4\,e^{-NI(s)}\,,$$

where $I(s)$ is the large deviations rate function for simple random walk on $\mathbf{Z}$ (see, e.g., Durrett 1996). Thus for $|v|>L$ we have

$$\mathbf{P}(\rho\leftrightarrow v)\le 4\exp\left(-|v|\frac{I(s)}{s}\right)\,.$$

Combine this with (13) and (8) to conclude that if $\lambda=e^{I(s)/s}$, then

$$0<\mathbf{P}(\rho\leftrightarrow\infty)\le\sum_{v\in\Pi}\mathbf{P}(\rho\leftrightarrow v)\le 4\sum_{v\in\Pi}\lambda^{-|v|}$$

for any cutset Π at distance more than L from the root. Hence $\mathrm{br}(\Gamma)\ge e^{I(s)/s}$. $\square$

Conjecture 1 *Under the assumptions of Question 5.3 above*

$$s\le\frac{\mathrm{br}(\Gamma)-1}{\mathrm{br}(\Gamma)+1}\,,\qquad \textit{i.e.,}\qquad \mathrm{br}(\Gamma)\ge\frac{1+s}{1-s}.$$

Remark. Very recently, this conjecture was proved by B. Virag (1998).

Recall that for simple random walk on the b-ary tree, the speed a.s. equals $\frac{b-1}{b+1}$.

Example 5.5 Take a binary tree and a ternary tree rooted together. The simple random walk on this tree does not have an a.s. constant speed. $\triangle$

The **Fibonacci tree** Γ_{fib} is a subtree of the binary tree. We label vertices as (L) and (R) (for "left" and "right"). The root is labeled (L). Every vertex labeled (L) has two offspring, one labeled (L) and one labeled (R). Every vertex labeled (R) has one offspring, which is labeled (L).

Exercise 5.6 *Justify the name Fibonacci tree. Also, show that*

$$\mathrm{br}(\Gamma_{\text{fib}})=\mathrm{gr}(\Gamma_{\text{fib}})=(1+\sqrt{5})/2.$$

Hint: Use a two state Markov chain to define a 'good' flow.

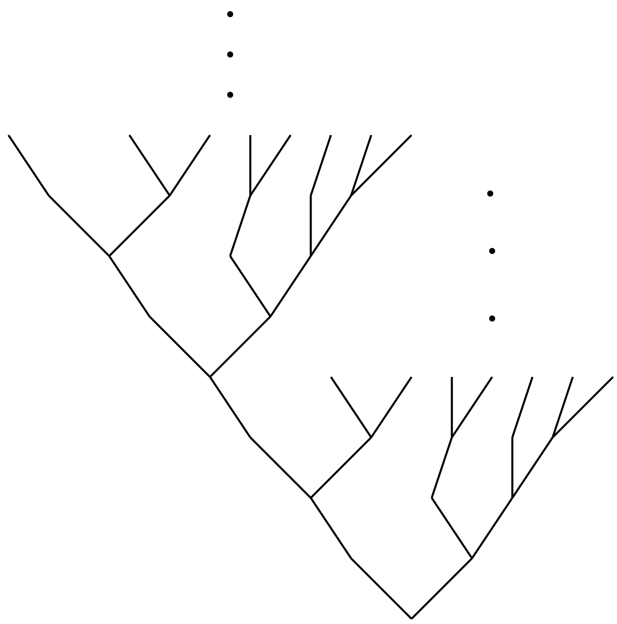


Figure 3: **The Fibonacci tree.**

6 Quasi-independent Percolation

Consider Bernoulli(p) percolation on a tree Γ. If v and w are vertices in Γ, then

$$\mathbf{P}(\rho \leftrightarrow u \ \text{ and } \ \rho \leftrightarrow w) = \frac{p^{|v|}p^{|w|}}{p^{|v\wedge w|}} = \frac{\mathbf{P}(\rho \leftrightarrow u)\mathbf{P}(\rho \leftrightarrow w)}{\mathbf{P}(\rho \leftrightarrow u \wedge w)},$$

where $v \wedge w$ is the vertex at which the paths from the root ρ to v and w separate. This turns out to be a key property of independent percolation, and we therefore make the following definition.

A **quasi-independent percolation** on a tree Γ is any general percolation so that for some $M < \infty$ and any vertices $u, v \in \Gamma$,

$$\mathbf{P}(\rho \leftrightarrow v \text{ and } \rho \leftrightarrow w) \leq M \frac{\mathbf{P}(\rho \leftrightarrow u)\mathbf{P}(\rho \leftrightarrow w)}{\mathbf{P}(\rho \leftrightarrow v \wedge w)}. \tag{15}$$

Example 6.1 *Percolation induced by i.i.d. labels.*

1. Let E be the edge set of a tree Γ, and let $\{X_e\}_{e\in E}$ be i.i.d. $\{-1,1\}$-valued random variables with $\mathbf{P}(X_e = 1) = 1/2$. Write $\mathtt{path}(v)$ for the unique path in Γ from the root to v. A **tree-indexed random walk** $\{S_v\}$ is defined for vertices v of Γ by
$$S_v = \sum_{e\in\mathtt{path}(v)} X_e \,.$$
Define $\Gamma(\omega) = \{v\colon S_v(\omega) \in [0,b)\}$. For $b = 2$, this is equivalent to Bernoulli(1/2) percolation: the only infinite paths in $\Gamma(\omega)$ are those for which each 1 is followed

by -1, and each -1 by 1 (with 1 in the first step). For $b > 2$, the corresponding percolation process is not independent, but it is quasi-independent.

2. Let $\{U_e\}$ be a collection of i.i.d. random variables, uniform on $[0,1)$, indexed by the edges of Γ. Define

$$\Gamma(\omega) = \left\{ v : \text{ for } \mathtt{path}(v) = e_1 e_2 \cdots e_{|v|},\ U_{e_1}(\omega) = \max_{k \le |v|} U_{e_k}(\omega) \right\}.$$

This is *not* quasi-independent.

For more on tree-indexed processes, see Chapter 18 and the survey article by Pemantle (1995). △

7 The Second Moment Method

For general percolation on a tree, the cutset sums (8) bound $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ from above. We get lower bounds by using the *second moment method*, which we describe next.

By our standing assumption about local finiteness of trees,

$$\{\rho \leftrightarrow \partial\Gamma\} = \bigcap_n \{\rho \leftrightarrow \Gamma_n\}.$$

We extend the definition of the boundary $\partial\Gamma$ to finite trees by

$$\partial\Gamma = \begin{cases} \text{leaves of } \Gamma\text{, i.e., vertices with no offspring} & \text{if } \Gamma \text{ is finite,} \\ \text{infinite paths starting at } \rho & \text{if } \Gamma \text{ is infinite.} \end{cases}$$

Consider the case Γ finite first. Let μ be a probability measure on $\partial\Gamma$ and set

$$Y \;=\; \sum_{x \in \partial\Gamma} \mu(x) \mathbf{1}_{\{\rho \leftrightarrow x\}} \frac{1}{\mathbf{P}(\rho \leftrightarrow x)}\,.$$

Then $\mathbf{E}[Y] \;=\; \sum_{x \in \partial\Gamma} \mu(x) = 1$, and

$$\begin{aligned} \mathbf{E}[Y^2] \;&=\; \mathbf{E}\left[\sum_{x \in \partial\Gamma} \sum_{y \in \partial\Gamma} \mu(x)\mu(y) \, \frac{\mathbf{1}_{\{\rho \leftrightarrow x\} \cap \{\rho \leftrightarrow y\}}}{\mathbf{P}(\rho \leftrightarrow x)\mathbf{P}(\rho \leftrightarrow y)} \right] \\ &=\; \sum_{x \in \partial\Gamma} \sum_{y \in \partial\Gamma} \mu(x)\mu(y) \frac{\mathbf{P}(\rho \leftrightarrow x \text{ and } \rho \leftrightarrow y)}{\mathbf{P}(\rho \leftrightarrow x)\mathbf{P}(\rho \leftrightarrow y)}\,. \end{aligned} \tag{16}$$

Thus, in the case of quasi-independent percolation,

$$\mathbf{E}[Y^2] \;\le\; M \sum_{x,y \in \partial\Gamma} \mu(x)\mu(y) \, \frac{1}{\mathbf{P}(\rho \leftrightarrow x \wedge y)}\,. \tag{17}$$

In the case of independent percolation, there is an equality with $M = 1$ in (17).

Define the **energy** of the measure μ in the kernel K as

$$\mathcal{E}_K(\mu) = \sum_{x,y \in \partial\Gamma} K(x,y)\mu(x)\mu(y) = \int_{\partial\Gamma}\int_{\partial\Gamma} K(x,y)\mu(dx)\mu(dy).$$

When the kernel is

$$K(x,y) \;=\; \frac{1}{\mathbf{P}(\rho \leftrightarrow x \wedge y)} \qquad \text{for } x, y \in \partial\Gamma\,,$$

(17) can be rewritten as

$$\mathbf{E}[Y^2] \leq M\mathcal{E}_K(\mu)\,.$$

By the Cauchy-Schwarz inequality,

$$(\mathbf{E}[Y])^2 = (\mathbf{E}[Y\mathbf{1}_{\{Y>0\}}])^2 \leq \mathbf{E}[Y^2]\mathbf{P}(Y>0)\,,$$

and consequently

$$\mathbf{P}(Y>0) \geq \frac{(\mathbf{E}[Y])^2}{\mathbf{E}[Y^2]} \geq \frac{1}{M}\frac{1}{\mathcal{E}_K(\mu)}\,.$$

Since $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \geq \mathbf{P}(Y > 0)$,

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \geq \frac{1}{M}\frac{1}{\mathcal{E}_K(\mu)}.$$

The left-hand side does not depend on μ, so optimizing the right-hand side with respect to μ yields

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \geq \frac{1}{M} \sup_{\mu:\mu(\partial\Gamma)=1} \frac{1}{\mathcal{E}_K(\mu)} = \frac{1}{M}\mathrm{Cap}_K(\partial\Gamma)\,, \tag{18}$$

where we define the **capacity** of $\partial\Gamma$ in the kernel K to be

$$\mathrm{Cap}_K(\partial\Gamma) = \sup_{\mu:\mu(\partial\Gamma)=1} \frac{1}{\mathcal{E}_K(\mu)}\,.$$

For Γ infinite, let μ be any probability measure on $\partial\Gamma$. μ induces a probability measure on Γ_n : for a vertex $x \in \Gamma_n$, set

$$\mu(x) = \mu(\text{infinite paths through } x)\,.$$

By the finite case considered above,

$$\mathbf{P}(\rho \leftrightarrow \Gamma_n) \geq \frac{1}{M}\frac{1}{\sum_{x,y\in\Gamma_n} K(x,y)\mu(x)\mu(y)}\,.$$

Each path ξ from the root ρ to ∞ must pass through some vertex x in Γ_n ; write $x \in \xi$ if the path ξ goes through vertex x. If $x \in \xi$ and $y \in \eta$, then $\xi \wedge \eta$ is a descendant of $x \wedge y$. This implies that $K(x,y) \leq K(\xi,\eta)$ for $x \in \xi$ and $y \in \eta$. Therefore,

$$\begin{aligned}\int_{\partial\Gamma}\int_{\partial\Gamma} K(\xi,\eta)d\mu(\xi)d\mu(\eta) &= \sum_{x,y\in\Gamma_n}\int_{x\in\xi}\int_{y\in\eta} K(\xi,\eta)d\mu(\xi)d\mu(\eta) \\ &\geq \sum_{x,y\in\Gamma_n} K(x,y)\mu(x)\mu(y) \\ &\geq \frac{1}{M}\frac{1}{\mathbf{P}(\rho \leftrightarrow \Gamma_n)}\,.\end{aligned}$$

Hence

$$\mathbf{P}(\rho \leftrightarrow \Gamma_n) \geq \frac{1}{M}\,\frac{1}{\mathcal{E}_K(\mu)}$$

for any probability measure μ on $\partial\Gamma$. Optimizing over μ and passing to the limit as $n \to \infty$, we get

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \geq \frac{1}{M}\mathrm{Cap}_K(\partial\Gamma)\,. \tag{19}$$

To summarize, we have established the following proposition.

Proposition 7.1 *Let Γ be finite or infinite, $\mathbf{P}$ the probability measure corresponding to a quasi-independent percolation on Γ, and K the kernel on $\partial\Gamma$ defined by $K(x,y) = \mathbf{P}(\rho \leftrightarrow x \wedge y)^{-1}$. Then*

$$\mathbf{P}(\rho \leftrightarrow \partial\Gamma) \geq \frac{1}{M}\mathrm{Cap}_K(\partial\Gamma), \tag{20}$$

where $M = 1$ in the case of independent percolation.

For Bernoulli percolation, we have already proven that $p_c(\Gamma) \geq 1/\mathrm{br}(\Gamma)$ in Proposition 5.1, using the first-moment method. We will now prove the reverse inequality, thus showing equality. For convenience, we restate the result.

Theorem 2.8 (R. Lyons 1990) *For Bernoulli(p) percolation on a tree Γ,*

$$p_c(\Gamma) = 1/\mathrm{br}(\Gamma)\,.$$

Proof. Take $p > 1/\mathrm{br}(\Gamma)$ and $1/p < \lambda < \mathrm{br}(\Gamma)$. By Lemma 2.10, there exists a unit flow μ from ρ to the boundary satisfying $\mu(v) \leq C\lambda^{-|v|}$ for each vertex $v \in \Gamma$. We may identify μ with a probability measure on $\partial\Gamma$ (see the discussion following Lemma 2.10).

Consider the kernel

$$K(\xi,\eta) \;=\; \frac{1}{\mathbf{P}(\rho \leftrightarrow \xi \wedge \eta)} = p^{-|\xi\wedge\eta|}\,.$$

The energy $\mathcal{E}_K(\mu)$ of μ in the kernel K is given by

$$\int_{\partial\Gamma}\int_{\partial\Gamma} p^{-|\xi\wedge\eta|}d\mu(\xi)d\mu(\eta) = \sum_v p^{-|v|}\int\int_{\xi\wedge\eta=v} d\mu(\xi)d\mu(\eta).$$

Since the set of pairs (ξ,η) with $\xi\wedge\eta = v$ is contained in the set of pairs (ξ,η) with $v\in\xi, v\in\eta$, the right-hand side above is not larger than

$$\begin{aligned}\sum_v p^{-|v|}[\mu(v)]^2 &= \sum_{n=0}^{\infty} p^{-n}\sum_{|v|=n}[\mu(v)]^2 \\ &\le \sum_{n=0}^{\infty} p^{-n}\sum_{|v|=n} C\lambda^{-|v|}\mu(v) \\ &= C\sum_{n=0}^{\infty}(p\lambda)^{-n}\mu(\Gamma_n).\end{aligned}$$

The last sum is finite since $\lambda p > 1$. Applying Proposition 7.1 yields

$$\mathbf{P}_p(\rho \leftrightarrow \partial\Gamma) \ge C^{-1}(1 - 1/\lambda p) > 0\,.$$

□

8 Electrical Networks

The basic reference for the material in this chapter is Doyle and Snell (1984). Here we will not restrict ourselves to trees, but will discuss general graphs.

While electrical networks are only a different language for reversible Markov chains, the electrical point of view is useful because of the insight gained from the familiar physical laws of electrical networks.

A **network** is a finite connected graph G, endowed with non-negative numbers $\{c_e\}$, called **conductances**, that are associated to the edges of G. The reciprocal $r_e = 1/c_e$ is the **resistance** of the edge e. A network will be denoted by the pair $\langle G, \{c_e\}\rangle$. Vertices of G are often called **nodes**. A real-valued function h defined on the vertices of G is **harmonic** at a vertex x of G if

$$\sum_{y\sim x}\frac{c_{xy}}{\pi_x}h(y) = h(x)\,, \qquad \text{where } \ \pi_x = \sum_{y\sim x} c_{xy}. \tag{21}$$

(Recall that the notation $y \sim x$ means y is a neighbor of x.)

We distinguish two nodes, $\{a, z\}$, which are called the *source* and the *sink* of the network. A function V which is harmonic on $G \setminus \{a, z\}$ will be called a **voltage**. A voltage is completely determined by its boundary values, V_a, V_z. In particular, the following result is derived from the maximum principle.

Proposition 8.1 *Let h be a function on a network G which is harmonic on $G\setminus\{a, z\}$ and such that $h(a) = h(z) = 0$. Then h must vanish everywhere on G.*

Proof. We will first show that $h \le 0$. Suppose this is not the case. Then $h(x_*) := \max_G h > 0$. By harmonicity on $G \setminus \{a,z\}$, if $x \notin \{a,z\}$ belongs to the set $A = \{x : h(x) = \max_G h\}$ and $y \sim x$, then $y \in A$ also. By connectedness, $a, z \in A$, hence $h(a) = h(z) = \max_G h > 0$, contradicting our assumption. Thus $h \le 0$, and an application of this result to $-h$ also yields $h \ge 0$. ☐

This proves that given boundary conditions $h(a) = x$ and $h(z) = y$, if there is a function harmonic on $G \setminus \{a,z\}$ with these boundary conditions, it is unique. To prove that a harmonic function with given boundary values exists, observe that the conditions (21) in the definition of harmonic functions form a system of linear equations with the same number of equations as unknowns, namely (number of nodes in G) $- 2$; for such a system, uniqueness of solutions implies existence.

A more informative way to prove existence is via the probabilistic interpretation of harmonic functions and voltages. Consider the Markov chain on the nodes of G with transition probabilities

$$p_{xy} = \mathbf{P}(X_{n+1} = y \mid X_n = x) = \frac{c_{xy}}{\pi_x}\,.$$

This process is called the **weighted random walk** on G with edge weights $\{c_e\}$, or the Markov chain associated to the network $\langle G, \{c_e\}\rangle$. This Markov chain is reversible with respect to the measure π:

$$\pi_x p_{xy} = c_{xy} = \pi_y p_{yx}\,.$$

A special case is the **simple random walk** on G, which has transition probabilities

$$p_{xy} = \frac{1}{\deg(x)} \qquad \text{for } y \sim x$$

and corresponds to the weighted walk with conductances $c_{xy} = 1$ for $y \sim x$.

To get a voltage with boundary values 0 and 1 at z and a respectively, set

$$V_x^* = \mathbf{P}_x(\{X_n\} \text{ hits } a \text{ before } z)\,,$$

where $\mathbf{P}_x$ is the probability for the walk started at node x. For arbitrary boundary values V_a and V_z, define

$$V_x = V_z + V_x^*(V_a - V_z)\,.$$

Until now, we have focused on *undirected graphs.* Now we need to consider also **directed graphs**. An edge in a directed graph is an *ordered* pair of nodes (x,y), which we denote by $\vec{e} = \vec{xy}$.

A **flow** θ from a to z, previously discussed when the underlying graph is a tree, is a function on oriented edges which is antisymmetric, $\theta(\vec{xy}) = -\theta(\vec{yx})$, and which obeys **Kirchhoff's node law** $\sum_{w \sim v} \theta(\vec{vw}) = 0$ at all $v \notin \{a,z\}$. This is just the requirement "flow in equals flow out" for any node $\neq a, z$. Despite notational differences, it is easily seen that these definitions generalize the ones given earlier for trees.

Observe that it is only flows that are defined on oriented edges. Conductance and resistance are defined for unoriented edges; we may of course define them on oriented edges by $c_{\vec{xy}} = c_{\vec{yx}} = c_{xy}$ and $r_{\vec{xy}} = r_{\vec{yx}} = r_{xy}$.

Given a voltage V on the network, the **current flow** associated with V is defined on oriented edges by

$$I(\vec{e}) = \frac{V_y - V_x}{r_e}, \qquad \text{where } \vec{e} = \vec{xy}\,.$$

Notice that I is antisymmetric and satisfies the node law at every $x \notin \{a, z\}$:

$$\sum_{y \sim x} I(\vec{xy}) = \sum_{y \sim x} c_{xy}(V_y - V_x) = 0.$$

Thus the node law for the current is equivalent to the harmonicity of the voltage.

The current flow also satisfies the **cycle law:** if the edges $\vec{e}_1, \ldots, \vec{e}_m$ form a cycle, i.e., $\vec{e}_i = \overrightarrow{x_{i-1}x_i}$ and $x_n = x_0$, then

$$\sum_{i=1}^{m} r_{e_i} I(\vec{e}_i) = 0\,.$$

Finally, by definition, a current flow also satisfies **Ohm's law:** if $\vec{e} = \vec{xy}$,

$$r_e I(\vec{e}) = V_y - V_x\,.$$

The particular values of a voltage function V are less important than the voltage differences, so fix a voltage function V on the network normalized to have $V_z = 0$.

By definition, if θ is an arbitrary flow on oriented edges satisfying Ohm's law $r_{xy}\theta(\vec{xy}) = V_y - V_x$ (with respect to the voltage V), then θ equals the current flow I associated with V.

Define the **strength** of an arbitrary flow θ as

$$\|\,\theta\,\| = \sum_{x \sim a} \theta(\vec{ax})\,.$$

Proposition 8.2 (Node law/cycle law/strength) *If θ is a flow from a to z satisfying the cycle law*

$$\sum_{i=1}^{m} r_{e_i}\theta(\vec{e}_i) = 0$$

for any cycle $\vec{e}_1 \ldots, \vec{e}_m$, and if $\|\,\theta\,\| = \|\,I\,\|$, then $\theta = I$.

Proof. The function $J = \theta - I$ satisfies the node-law at all nodes and the cycle law. Define

$$h(x) = \sum_{i=1}^{m} J(\vec{e}_i) r_{e_i}\,,$$

where $\vec{e}_i, \ldots, \vec{e}_m$ is an arbitrary path from a to x. By the cycle law, J is well defined. By the node law, it is harmonic everywhere, except possibly at a and z. Now $\| \theta \| = \| I \|$ implies that J is also harmonic at a and z. By the maximum principle, h must be constant. This implies that $J = 0$. $\square$

Given a network, the ratio $(V_a - V_z)/ \| I \|$, where I is the current flow corresponding to the voltage V, is independent of the voltage V applied to the network. Define the **effective resistance** between vertices a and z as

$$\mathcal{R}(a \leftrightarrow z) := \frac{V_a - V_z}{\| I \|} .$$

We think of effective resistance as follows: replace the whole network by a single edge joining a to z and require that the two networks be equivalent, in the sense that the amount of current flowing from a to z in the new network is the same as in the original network if we apply the same voltage to both.

Next, we discuss the probabilistic interpretation of effective resistance. Denote

$$\mathbf{P}(a \to z) := \mathbf{P}_a(\text{hit } z \text{ before returning to } a) .$$

For any vertex x

$$\mathbf{P}_x(\text{hit } z \text{ before } a) = \frac{V_a - V_x}{V_a - V_z} .$$

If $p_{xy} = c_{xy}\pi_x^{-1}$ are the transition probabilities of the Markov chain, then

$$\begin{aligned} \mathbf{P}(a \to z) &= \sum_x p_{ax}\mathbf{P}_x(\text{hit } z \text{ before } a) \\ &= \sum_{x \sim a} \frac{c_{ax}}{\pi_a} \frac{V_a - V_x}{V_a - V_z} \\ &= \frac{1}{\pi_a(V_a - V_z)} \sum_{x \sim a} I(\vec{ax}) \\ &= \frac{1}{\pi_a} \frac{\| I \|}{V_a - V_z} \\ &= \frac{1}{\pi_a \mathcal{R}(a \leftrightarrow z)} . \end{aligned}$$

Call $[\mathcal{R}(a \leftrightarrow z)]^{-1}$ the **effective conductance**, written as $\mathcal{C}(a \leftrightarrow z)$. Then

$$\mathbf{P}(a \to z) = \frac{1}{\pi_a}\mathcal{C}(a \leftrightarrow z) . \tag{22}$$

The **Green function** for the random walk stopped at z, is defined by

$$G(a, x) = \mathbf{E}_a[\# \text{ visits to } x \text{ before hitting } z] .$$

(The subscript in $\mathbf{E}_a$ indicates the initial state.) Then $G(a, a) = \pi_a \mathcal{R}(a \leftrightarrow z)$, since the number of visits to a before visiting z has a geometric distribution with parameter $\mathbf{P}(a \to z)$. It is often possible to replace a network by a simplified one without changing quantities of interest, for example the effective resistance between a pair of nodes. The following laws are very useful.

Parallel Law. Conductances in parallel add: Suppose edges e_1 and e_2, with conductances c_1 and c_2 respectively, share vertices v_1 and v_2 as endpoints. Then both edges can be replaced with a single edge of conductance $c_1 + c_2$ without affecting the rest of the network. All voltages and currents in $G \setminus \{e_1, e_2\}$ are unchanged and the current $I(\vec{e})$ equals $I(\vec{e}_1) + I(\vec{e}_2)$. For a proof, check Ohm's and Kirchhoff's laws with $I(\vec{e}) := I(\vec{e}_1) + I(\vec{e}_2)$.

Series Law. Resistances in series add: If $v \in G \setminus \{a, z\}$ is a node of degree 2 with neighbors v_1 and v_2, the edges (v_1, v) and (v, v_2) can be replaced by a single edge (v_1, v_2) of resistance $r_{v_1 v} + r_{v v_2}$. All potentials and currents in $G \setminus \{v\}$ remain the same and the current that flows from v_1 to v_2 equals $I(\overrightarrow{v_1 v}) = I(\overrightarrow{v v_2})$. For a proof, check again Ohm's and Kirchhoff's laws, with $I(\overrightarrow{v_1 v_2}) := I(\overrightarrow{v_1 v}) = I(\overrightarrow{v v_2})$.

Glue. Another convenient operation is to identify vertices having the same voltage, while keeping all existing edges. Because current never flows between vertices with the same voltage, potentials and currents are unchanged.

Example 8.3 Consider a spherically symmetric tree Γ, a tree in which all vertices of Γ_n have the same number of children for all $n \geq 0$. Suppose that all edges at the same distance from the root have the same resistance, that is, $r_e = r_i$ if $|e| = i$, $i \geq 1$. Glue all the vertices in each level; This will not affect effective resistances, so we infer that

$$\mathcal{R}(\rho \leftrightarrow \Gamma_N) = \sum_{i=1}^{N} \frac{r_i}{|\Gamma_i|}$$

and

$$\mathbf{P}(\rho \to \Gamma_N) = \frac{r_1/|\Gamma_1|}{\sum\limits_{i=1}^{N} r_i/|\Gamma_i|} \,.$$

Therefore the corresponding random walk on Γ is transient iff $\sum\limits_{i=1}^{\infty} r_i/|\Gamma_i| < \infty$. $\triangle$

Theorem 8.4 (Thomson's Principle) *For any finite connected graph,*

$$\mathcal{R}(a \leftrightarrow z) = \inf \{ \mathcal{E}(\theta) \colon \theta \text{ a unit flow from } a \text{ to } z \} ,$$

where $\mathcal{E}(\theta) := \sum_e [\theta(e)]^2 r_e$. *The unique minimizer in the* inf *above is the unit current flow.*

Note: The sum in $\mathcal{E}(\theta)$ is over unoriented edges, so each edge $\{x, y\}$ is only considered once in the definition of energy. Although θ is defined on oriented edges, it is antisymmetric and hence $\theta(e)^2$ is unambiguous.

Proof. By compactness, there exists flows minimizing $\mathcal{E}(\theta)$ subject to $\|\theta\| = 1$. By Proposition 8.2, to prove that the unit current flow is the unique minimizer, it is enough to verify that any unit flow θ of minimal energy satisfies the cycle law.

Let the edges $\vec{e}_1, \dots \vec{e}_n$ form a cycle. Set $\gamma(\vec{e}_i) = 1$ for all $1 \le i \le n$ and set γ equal to zero on all other edges. Note that γ satisfies the node law, so it is a flow, but $\sum \gamma(\vec{e}_i) = n \neq 0$. For any $\epsilon \in \mathbf{R}$, we have that

$$0 \le \mathcal{E}(\theta + \epsilon\gamma) - \mathcal{E}(\theta) = \frac{1}{2}\sum_{i=1}^{n} [(\theta(\vec{e}_i) + \epsilon)^2 - \theta(\vec{e}_i)^2] r_{e_i} = \epsilon \sum_{i=1}^{n} r_{e_i}\theta(\vec{e}_i) + O(\epsilon^2)\,.$$

By taking $\epsilon \to 0$ from above and from below, we see that $\sum\limits_{i=1}^{n} r_{e_i}\theta(\vec{e}_i) = 0$, thus verifying that θ satisfies the cycle law.

To complete the proof, we show that the unit current flow I has $\mathcal{E}(I) = \mathcal{R}(a \leftrightarrow z)$:

$$\begin{aligned}
\sum_e r_e I(e)^2 &= \frac{1}{2}\sum_x \sum_y r_{xy}\left(\frac{V_y - V_x}{r_{xy}}\right)^2 \\
&= \frac{1}{2}\sum_x \sum_y c_{xy}(V_y - V_x)^2 \\
&= \frac{1}{2}\sum_x \sum_y (V_y - V_x) I(\vec{xy}).
\end{aligned}$$

Since I is antisymmetric,

$$\frac{1}{2}\sum_x \sum_y (V_y - V_x) I(\vec{xy}) = -\sum_x V_x \sum_y I(\vec{xy})\,. \tag{23}$$

Applying the node law and recalling that $\|I\| = 1$, we conclude that the right-hand side of (23) is equal to

$$\frac{V_z - V_a}{\|I\|} = \mathcal{R}(a \leftrightarrow z)\,.$$

$\square$

Let a, z be vertices in a network, and suppose that we add to the network an edge which is not incident to a. How does this affect the escape probability from a to z? Probabilistically the answer is not obvious. In the language of electrical networks, this question is answered by:

Theorem 8.5 (Rayleigh's Monotonicity Law) *If $\{r_e\}$ and $\{r'_e\}$ are sets of resistances on the edges of the same graph G, and if $r_e \le r_{e'}$ for all e, then*

$$\mathcal{R}(a \leftrightarrow z; r) \;\le\; \mathcal{R}(a \leftrightarrow z; r')\,.$$

Proof. Note that $\inf\limits_\theta \sum\limits_e r_e\theta(e)^2 \le \inf\limits_\theta \sum\limits_e r'_e\theta(e)^2$ and apply Thomson's Principle (Theorem 8.4). $\square$

Corollary 8.6 *Adding an edge weakly decreases the effective resistance $\mathcal{R}(a \leftrightarrow z)$. If the added edge is not incident to a, the addition weakly increases the escape probability* $\mathbf{P}(a \to z) = [\,\pi_a \mathcal{R}(a \leftrightarrow z)\,]^{-1}$.

Proof. Before we add an edge to a network we can think of it as existing already with $c = 0$ or $r = \infty$. By adding the edge we reduce its resistance to a finite number. □

Thus, combining the relationship (22) and Corollary 8.6 shows that the addition of an edge not incident to a (which we regard as changing a conductance from 0 to 1) cannot decrease the escape probability $\mathbf{P}(a \to z)$.

Exercise 8.7 *Show that $\mathcal{R}(a \leftrightarrow z)$ is a concave function of $\{r_e\}$.*

Corollary 8.8 *The operation of gluing vertices cannot increase effective resistance.*

Proof. When we glue vertices together, we take an infimum over a larger class of flows. □

Moreover, if we glue together vertices with different potentials, then effective resistance will strictly decrease.

9 Infinite Networks

For an infinite graph G containing vertex a, let $\{G_n\}$ be a collection of finite connected subgraphs containing a and satisfying $\cup_n G_n = G$. If all the vertices in $G \setminus G_n$ are replaced by a single vertex z_n, then

$$\mathcal{R}(a \leftrightarrow \infty) := \lim_{n\to\infty} \mathcal{R}(a \leftrightarrow z_n \text{ in } G_n \cup \{z_n\}) .$$

Now

$$\mathbf{P}(a \to \infty) = \frac{\mathcal{C}(a \leftrightarrow \infty)}{\pi_a} .$$

A flow on G from a to infinity is an antisymmetric edge function obeying the node law at all vertices except a. Thomson's Principle remains valid for infinite networks:

$$\mathcal{R}(a \leftrightarrow \infty) = \inf\{\,\mathcal{E}(\theta) : \theta \text{ a unit flow from } a \text{ to } \infty\,\} . \tag{24}$$

Let us summarize the facts in the following proposition.

Proposition 9.1 *Let $\langle G, \{c_e\}\rangle$ be a network. The following are equivalent.*

1. *The weighted random walk on the network is transient.*
2. *There is some node a with $\mathcal{C}(a \leftrightarrow \infty) > 0$ (equivalently, $\mathcal{R}(a \leftrightarrow \infty) < \infty$).*
3. *There is a flow θ from some node a to infinity with $\|\,\theta\,\| > 0$ and $\mathcal{E}(\theta) < \infty$.*

In particular, any subgraph of a recurrent graph must be recurrent.

Recall that an edge-cutset Π separating a from z is a set of edges so that any path from a to z must include some edge in Π.

Corollary 9.2 (Nash-Williams (1959)) *If $\{\Pi_n\}$ are disjoint edge-cutsets which separate a from z, then*

$$\mathcal{R}(a \leftrightarrow z) \geq \sum_n \left(\sum_{e \in \Pi_n} c_e \right)^{-1} . \tag{25}$$

In an infinite network $\langle G, \{c_e\} \rangle$, the analogous statement with z replaced by ∞ is also valid; in particular, if there exist disjoint edge-cutsets $\{\Pi_n\}$ that separate a from ∞ and satisfy

$$\sum_n \left(\sum_{e \in \Pi_n} c_e \right)^{-1} = \infty ,$$

then the weighted random walk on $\langle G, \{c_e\} \rangle$ is recurrent.

Proof. Let θ be a unit flow from a to z. For any n

$$\sum_{e \in \Pi_n} c_e \cdot \sum_{e \in \Pi_n} r_e \theta(e)^2 \geq \left(\sum_{e \in \Pi_n} \sqrt{c_e}\sqrt{r_e} |\theta(e)| \right)^2 = \left(\sum_{e \in \Pi_n} |\theta(e)| \right)^2 \geq \| \theta \|^2 = 1 ,$$

because Π_n is a cutset and $\| \theta \| = 1$. Therefore

$$\sum_e r_e \theta(e)^2 \geq \sum_n \sum_{e \in \Pi_n} r_e \theta(e)^2 \geq \sum_n \left(\sum_{e \in \Pi_n} c_e \right)^{-1} .$$

$\square$

Example 9.3 ($\mathbf{Z}^2$ is recurrent) Take $r_e = 1$ on $G = \mathbf{Z}^2$ and consider the cutsets consisting of edges joining vertices in $\partial \Box_n$ to vertices in $\partial \Box_{n+1}$, where $\Box_n = [-n, n]^2$. Then by Nash-Williams (25),

$$\mathcal{R}(a \leftrightarrow \infty) \geq \sum_n \frac{1}{4(2n+1)} = \infty .$$

Thus simple random walk on $\mathbf{Z}^2$ is recurrent. Moreover, we obtain a lower bound for the resistance from the center of a square $\Box_n = [-n, n]^2$ to its boundary:

$$\mathcal{R}(0 \leftrightarrow \partial \Box_n) \geq c \log n .$$

In the next chapter, we will obtain an upper bound of the same type. $\triangle$

The Nash-Williams inequality (25) is useful, but in general is not sharp. For example, for the 3-1 tree in Example 2.6, the effective resistance from the root to ∞ is infinite because the random walk is recurrent, yet the right-hand side of (25) is at most 1 for any sequence of disjoint cutsets (prove this, or see Lyons and Peres 1999).

Example 9.4 ($\mathbf{Z}^3$ is transient) To each directed edge $\vec{e}$ in the lattice $\mathbf{Z}^3$, attach an orthogonal unit square $\Box_e$ intersecting $\vec{e}$ at its midpoint m_e. Define $\theta(\vec{e})$ to be the area of the radial projection of $\Box_e$ onto the sphere $\partial B(0, \frac{1}{4})$, taken with a positive sign if $\vec{e}$ points in the same direction as the radial vector from 0 to m_e, and with a negative sign otherwise. By considering a unit cube centered at each lattice point and projecting it to $\partial B(0, \frac{1}{4})$, we can easily verify that θ satisfies the node law at all vertices except the origin. Hence θ is a flow from 0 to ∞ in $\mathbf{Z}^3$. It is easy to bound its energy:

$$\mathcal{E}(\theta) \leq \sum_n C_1 n^2 \left(\frac{C_2}{n^2}\right)^2 < \infty .$$

By Proposition 9.1, $\mathbf{Z}^3$ is transient. This works for any $\mathbf{Z}^d$, $d \geq 3$. An analytic description of the same flow was given by T. Lyons (1983).

△

Exercise 9.5 *Fix $k > 1$. Define the k-**fuzz** of an undirected graph $G = (V, E)$ as the graph $G_k = (V, E_k)$ where for any two distinct vertices $v, w \in V$, the edge $\{v, w\}$ is in E_k iff there is a path of at most k edges in E connecting v to w. Show that for G with bounded degrees, G is transient iff G_k is transient.*

A solution can be found in Doyle and Snell (1984, §8.4).

10 The Method of Random Paths

A **self-avoiding path** from a to z is a sequence of vertices $v_0, \ldots, v_n$ such that $v_0 = a$ and $v_n = z$, adjacent vertices v_{i-1} and v_i are connected by an edge, and $v_i \neq v_j$ for $i \neq j$. If φ and ψ are two self-avoiding paths from a to z, define

$$|\varphi \cap \psi| = \text{number of edges in the intersection of } \varphi \text{ and } \psi .$$

If $\vec{e}$ is the oriented edge pointing from vertex v to w, let $\overleftarrow{e}$ be the reversed edge pointing from w to v. If μ is a measure on the set of self-avoiding paths from a to z, define

$$\mu(e) = \mu(\varphi : \varphi \ni e) = \mu(\varphi : \varphi \ni \vec{e} \text{ or } \varphi \ni \overleftarrow{e}).$$

The Nash-Williams inequality yields lower bounds for effective resistance. For upper bounds the following result is useful. Assume that $r_e = 1$ for all e; the result can be extended easily to arbitrary resistances.

Theorem 10.1 (Method of random paths)

$$\mathcal{R}(a \leftrightarrow z) = \inf_\mu \sum_e [\mu(e)]^2 = \inf_\mu \mathbf{E}_{\mu \times \mu}[\,|\varphi \cap \psi|\,] ,$$

where the infimum is over all probability measures μ on the set of self-avoiding paths from a to z, and φ and ψ are independent paths with distribution μ. Similarly, if there is a measure μ on infinite self-avoiding paths in a graph G with $\mathbf{E}_{\mu \times \mu}[\,|\varphi \cap \psi|\,] < \infty$, then simple random walk on G is transient.

Remark. The useful direction here is $\mathcal{R}(a \leftrightarrow z) \le \sum \mu(e)^2$ for all μ.

Proof. The second equality is trivial: write $|\varphi \cap \psi|$ as $\sum_e \mathbf{1}_{\{\varphi \ni e, \psi \ni e\}}$.

Given a probability measure μ on the set of self-avoiding paths from a to z, define

$$\begin{aligned}\theta(\vec{e}) &:= \mu(\varphi : \varphi \ni \vec{e}) - \mu(\varphi : \varphi \ni \overleftarrow{e}) \\ &= \mathbf{E}_\mu \left[\mathbf{1}\{\varphi \ni \vec{e}\} - \mathbf{1}\{\varphi \ni \overleftarrow{e}\}\right].\end{aligned}$$

By definition, θ is antisymmetric. To see that θ obeys the node law, observe that

$$\sum_{w : w \sim v} \theta(\vec{vw}) = \mathbf{E}_\mu \left[\sum_{w : w \sim v} \mathbf{1}\{\varphi \ni \vec{vw}\} - \mathbf{1}\{\varphi \ni \overleftarrow{vw}\}\right].$$

Assume $v \notin \{a, z\}$. If, for a sample path φ, a term in the sum is nonzero, then φ must use either an edge directed *to* v or an edge directed *from* v. But because φ is a self-avoiding walk which terminates at z, it must also use exactly one other edge incident to v, in the first case directed *away* from v and in the second case directed *to* v. Hence the net contribution of φ to the sum is zero. We conclude that θ is a flow.

Clearly, θ is a unit flow, *i.e.*.

$$\|\,\theta\,\| = \sum_{x \sim a} \theta(\vec{ax}) = 1\,.$$

so we can apply Thomson's principle:

$$\mathcal{R}(a \leftrightarrow z) \le \sum_e [\theta(e)]^2 \le \sum_e [\mu(e)]^2.$$

The other inequality $\mathcal{R}(a \leftrightarrow z) \ge \inf_\mu \sum \mu(e)^2$ will not be used in these notes, so we only sketch a proof. Let I denote a unit current flow. Then

$$\mathcal{R}(a \leftrightarrow z) = \sum_e I(e)^2.$$

Notice that a unit current flow is acyclic. Define a Markov chain by making transitions according to the flow I normalized. This chain then defines a measure on paths and $\mu(\vec{e}) = I(\vec{e})$, because I is acyclic. For details, see Lyons and Peres (1999). $\square$

Example 10.2 In $\mathbf{Z}^2$, consider the boundary $\partial\Box_n = \{x \in \mathbf{Z}^2 \colon \|\,x\,\|_1 = n\}$ of the square $\Box_n = [-n, n]^2$. Using Nash-Williams we have seen that

$$\mathcal{R}(0 \leftrightarrow \Box_n) \ge c \log n\,.$$

Now define a measure μ on self-avoiding paths in $\Box_n$ as follows: Pick a ray $\vec{\ell}$ emanating from the origin in a random uniformly distributed direction, and let μ be the distribution of the lattice path that best approximates ℓ. By considering edges e according to their distance from the origin, we also get

$$\sum_e [\mu(e)]^2 \le \sum_{k=1}^{n} c_1 k \left(\frac{c_2}{k}\right)^2 \le C \log n\,.$$

So in $\mathbf{Z}^2$ we have

$$c \log n \leq \mathcal{R}(0 \leftrightarrow \partial \Box_n) \leq C \log n\,.$$

$\triangle$

Example 10.3 In $\mathbf{Z}^3$, define μ analogously, but this time on the whole infinite lattice. Now

$$\mathcal{R}(0 \leftrightarrow \infty) \leq \sum_k c_1 k^2 \left(\frac{c_2}{k^2}\right)^2 < \infty.$$

$\triangle$

Example 10.4 (Wedges in $\mathbf{Z}^3$) Given a non-negative and non-decreasing function f, consider the wedge

$$W_f = \{(x, y, z) \colon 0 \leq y \leq x, \quad 0 \leq z \leq f(x)\}\,.$$

By Nash-Williams, the resistance from the origin to ∞ in W_f satisfies

$$\mathcal{R}(0 \leftrightarrow \infty) \geq C \sum_k \frac{1}{k f(k)}\,.$$

In particular, if this sum diverges, then W_f is recurrent. The converse also holds: $\triangle$

Theorem 10.5 (T. Lyons 1983) *If $\sum_k [k f(k)]^{-1} < \infty$, then the wedge W_f is transient.*

Proof Idea. Choose a random point (U_1, U_2) according to the uniform distribution on $[0, 1]^2$ and find the lattice path closest to $\{(k, U_1 k, U_2 f(k))\}_{k=0}^{\infty}$. The completion of this proof is left as an exercise. $\Box$

11 Transience of Percolation Clusters

The graph $\mathbf{Z}^3$ supports a flow of finite energy, described in Example 9.4, and hence simple random walk in three dimensions is transient. Equivalently, if each edge of $\mathbf{Z}^3$ is assigned unit conductance, then the effective conductance from any vertex to infinity is positive. If a finite number of edges are removed, then the random walk on the infinite component of the modified graph is also transient, because the effective conductance remains nonzero.

A much deeper result, first proved by Grimmett, Kesten, and Zhang (1993), is that if $d \geq 3$ and $p > p_c(\mathbf{Z}^d)$, then simple random walk on $\mathcal{C}_\infty(\mathbf{Z}^d, p)$ is transient, where $\mathcal{C}_\infty(\mathbf{Z}^d, p)$ is the unique infinite cluster of Bernoulli(p) percolation on $\mathbf{Z}^d$. Benjamini, Pemantle and Peres (1998) (hereafter referred to as BPP (1998)) gave an alternative proof of this result and extended it to high-density oriented percolation. Their argument uses certain "unpredictable" random paths that have *exponential intersection tails* to construct random flows of finite energy on $\mathcal{C}_\infty(\mathbf{Z}^d, p)$.

Let $G = (V_G, E_G)$ be an infinite graph with all vertices of finite degree and let $v_0 \in V_G$. Denote by $\Upsilon = \Upsilon(G, v_0)$ the collection of infinite oriented paths in G which emanate from v_0. Let $\Upsilon_1 = \Upsilon_1(G, v_0) \subset \Upsilon$ be the set of **paths with unit speed**, *i.e.*, those paths for which the n^{th} vertex is at distance n from v_0.

Let $0 < \zeta < 1$. A Borel probability measure μ on $\Upsilon(G, v_0)$ has **exponential intersection tails** with parameter ζ (in short, EIT(ζ)) if there exists C such that

$$\mu \times \mu\Big\{(\varphi, \psi) : |\varphi \cap \psi| \geq n\Big\} \leq C\zeta^n \tag{26}$$

for all n, where $|\varphi \cap \psi|$ is the number of edges in the intersection of φ and ψ. If such a measure μ exists for some basepoint v_0 and some $\zeta < 1$, then we say that G *admits random paths with* EIT(ζ). By the previous chapter, such a graph G must be transient.

Theorem 11.1 (Cox-Durrett 1983, BPP 1998) *For every $d \geq 3$, there exists $\zeta < 1$ such that the lattice $\mathbf{Z}^d$ admits random paths with EIT(ζ).*

Proof: For $d \geq 4$, we will show (following Cox and Durrett 1983, who attribute the idea to Kesten) that the "uniform distribution" on $\Upsilon_1(\mathbf{Z}^d, 0)$ has the required EIT property; for $d = 3$ such a simple choice cannot work, and we will delay the proof to Chapter 17. Let $d \geq 4$, and define μ to be the distribution of the random walk with i.i.d. increments uniformly distributed on the d standard basis vectors $(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$. Let $\{X_n\}$ and $\{Y_m\}$ be two independent random walks with distribution μ. It suffices to show that the number of vertex intersections of these two walks has an exponential tail. Since $\| X_n \|_1 = n$ for all n, we can have $X_n = Y_m$ only if $n = m$. The process $\{X_n - Y_n\}$ is a mean 0 random walk in the $d-1$ dimensional sublattice of $\mathbf{Z}^d$ consisting of vectors orthogonal to $(1, 1, \ldots, 1)$, and its increments generate this sublattice. Since $d - 1 \geq 3$, the random walk $\{X_n - Y_n\}$ is transient, and (26) holds with

$$\zeta := \mathbf{P}[\exists n \geq 1 \;\; X_n - Y_n = 0], \quad \text{and } \; C = 1.$$

$\square$

Proposition 11.2 (BPP 1998) *Suppose that the directed graph G admits random paths with EIT(ζ), and consider Bernoulli(p) percolation on G. If $p > \zeta$ then with probability 1 there is a vertex v in G such that the open cluster $\mathcal{C}(v)$ is transient.*

Proof. The hypothesis means that there is some vertex v_0 and a probability measure μ on $\Upsilon = \Upsilon(G, v_0)$ satisfying (26). We will assume here that μ is supported on Υ_1; the general case is treated in BPP (1998).

For $N \geq 1$ and any infinite path $\varphi \in \Upsilon_1(G, v_0)$, denote by φ_N the finite path consisting of the first N edges of φ. Consider the random variable

$$Z_N = \int_{\Upsilon_1} p^{-N} \mathbf{1}_{\{\varphi_N \text{ is open}\}}\, d\mu(\varphi)\,. \tag{27}$$

Except for the normalization factor p^{-N}, this is the μ-measure of the paths that stay in the open cluster of v_0 for N steps.

Since each edge is open with probability p (independently of other edges), $\mathbf{E}(Z_N) = 1$, but we can say more. Let $\mathcal{B}_N$ be the σ-field generated by the status (open or closed) of all edges on paths φ_N with $\varphi \in \Upsilon_1$. It is easy to check that for each $\varphi \in \Upsilon_1$, the sequence $\{p^{-N}\mathbf{1}_{\{\varphi_N \text{ is open}\}}\}$ is a martingale adapted to the filtration $\{\mathcal{B}_N\}_{N\geq 1}$. Consequently, $\{Z_N\}_{N\geq 1}$ is also a non-negative martingale. By the Martingale Convergence Theorem, $\{Z_N\}$ converges a.s. to a random variable Z_∞. In fact, we now show that $\{Z_N\}$ is bounded in L^2, and hence converges in L^2. Since each edge is open with probability p (independently of other edges), $\mathbf{E}(Z_N) = 1$. The second moment of Z_N satisfies

$$
\begin{aligned}
\mathbf{E}(Z_N^2) &= \mathbf{E}\int_{\Upsilon_1}\int_{\Upsilon_1} p^{-2N}\mathbf{1}_{\{\varphi_N \text{ and } \psi_N \text{ are open}\}}\, d\mu(\varphi)\, d\mu(\psi) \\
&\leq \int_{\Upsilon_1}\int_{\Upsilon_1} p^{-|\varphi\cap\psi|}\, d\mu(\varphi)\, d\mu(\psi) \\
&= \sum_{k=1}^{\infty} p^{-k}\mu\times\mu\{(\varphi,\psi) : |\varphi\cap\psi| = k\}\,. \qquad (28)
\end{aligned}
$$

By (26), the sum on the right-hand side of (28) is bounded by $\sum_{k=1}^{\infty} C\left(\frac{\zeta}{p}\right)^k$, which does not depend on N and is finite since $\zeta < p$.

On the event $\{Z_\infty > 0\}$, the cluster $\mathcal{C}(v_0)$ is infinite, and by Cauchy-Schwarz,

$$\mathbf{P}(|\mathcal{C}(v_0)| = \infty) \geq \mathbf{P}(Z_\infty > 0) \geq \frac{(\mathbf{E}Z_\infty)^2}{\mathbf{E}Z_\infty^2}\,.$$

Since $\mathbf{E}Z_N^2$ is bounded, by Fatou's Lemma the right-hand side is positive. Thus with positive probability $\mathcal{C}(v_0)$ is infinite, and it remains to prove that $\mathcal{C}(v_0)$ is a.s. transient on this event.

We will construct a flow of finite energy on $\mathcal{C}(v_0)$. For each $N \geq 1$, and every edge $\vec{e}$ directed away from v_0, define

$$f_N(\vec{e}) = \int_{\Upsilon} p^{-N}\mathbf{1}_{\{\varphi_N \text{ is open}\}}\mathbf{1}_{\{\vec{e}\in\varphi_N\}}\, d\mu(\varphi)\,. \qquad (29)$$

If $\vec{e}$ is directed towards v_0, let $f(\vec{e}) = -f(\overleftarrow{e})$, where $\overleftarrow{e}$ is the reversal of $\vec{e}$. Let $B(v_0, N)$ denote the set of all vertices within distance N of v_0. Then f_N is a flow on $\mathcal{C}(v_0)\cap B(v_0, N+1)$ from v_0 to the complement of $B(v_0, N)$, *i.e.,* for any vertex $v \in B(v_0, N)$ except v_0, the incoming flow to v equals the outgoing flow from v. The strength of f_N (the total outflow from v_0) is exactly Z_N.

Next, we estimate the expected energy of f_N by summing over edges directed away from v_0:

$$\mathbf{E}\sum_{\vec{e}} f_N(\vec{e})^2 = \mathbf{E}\int_{\Upsilon_1}\int_{\Upsilon_1} p^{-2N}\mathbf{1}_{\{\varphi_N,\psi_N \text{ are open}\}}\sum_{\vec{e}}\mathbf{1}_{\{\vec{e}\in\varphi_N\}}\mathbf{1}_{\{\vec{e}\in\psi_N\}}\, d\mu(\varphi)\, d\mu(\psi)$$

$$\begin{aligned}
&\leq \int_{\Upsilon_1}\int_{\Upsilon_1} |\varphi \cap \psi|\, p^{-|\varphi\cap\psi|}\, d\mu(\varphi)\, d\mu(\psi) \\
&= \sum_{k=1}^{\infty} kp^{-k}\mu \times \mu\{(\varphi,\psi) : |\varphi \cap \psi| = k\}. \qquad (30)
\end{aligned}$$

Again using (26) and $p > \zeta$, from (30) we conclude that

$$\mathbf{E}\sum_{\vec{e}} f_N(\vec{e})^2 \leq \sum_{k=1}^{\infty} k \left(\frac{\zeta}{p}\right)^k = C < \infty, \qquad (31)$$

where C does not depend on N.

For each directed edge $\vec{e}$ of G, the sequence $\{f_N(\vec{e})\}$ is a $\{\mathcal{B}_N\}$-martingale which converges a.s. and in L^2 to a nonnegative random variable $f(\vec{e})$. The edge function f is a flow with strength Z_∞ on $\mathcal{C}(v_0)$, and has finite expected energy by (31) and Fatou's Lemma.

Thus

$$\mathbf{P}[\mathcal{C}(v_0) \text{ is transient}] \geq \mathbf{P}[Z_\infty > 0] > 0\,,$$

so the tail event $\{\exists v : \mathcal{C}(v) \text{ is transient}\}$ must have probability 1 by Kolmogorov's zero-one law. $\square$

Theorem 11.3 (Grimmett, Kesten and Zhang 1993) *Consider Bernoulli(p) percolation on $\mathbf{Z}^d$, where $d \geq 3$. For all $p > p_c$, the unique infinite cluster is a.s. transient.*

Proof. It follows from Theorem 11.1 and Proposition 11.2 that the infinite cluster is transient if p is close enough to 1.

Recall that a set of graphs $\mathbf{B}$ is called **increasing** if for any graph G that contains a subgraph in $\mathbf{B}$, necessarily G must also be in $\mathbf{B}$.

Consider now percolation with *any* parameter $p > p_c$ in $\mathbf{Z}^d$. Following Pisztora (1996), call an open cluster $\mathcal{C}$ contained in some cube Q a *crossing cluster* for Q if for all d directions there is an open path contained in $\mathcal{C}$ joining the left face of Q to the right face. For each v in the lattice $N\mathbf{Z}^d$, denote by $\square_N(v)$ the cube of side-length $5N/4$ in $\mathbf{Z}^d$, centered at v. Let $A_p(N)$ be the set of $v \in N\mathbf{Z}^d$ with the following property: *The cube $\square_N(v)$ contains a crossing cluster $\mathcal{C}$ such that any open cluster in $\square_N(v)$ of diameter greater than $N/10$ is connected to $\mathcal{C}$ by an open path in $\square_N(v)$.*

Proposition 2.1 in Antal and Pisztora (1996), which relies on the work of Grimmett and Marstrand (1990), implies that $A_p(N)$ stochastically dominates site percolation with parameter $p_*(N)$ on the stretched lattice $N\mathbf{Z}^d$, where $p_*(N) \to 1$ as $N \to \infty$. By Liggett, Schonmann and Stacey (1996), it follows that $A_p(N)$ stochastically dominates bond percolation with parameter $p^*(N)$ on $N\mathbf{Z}^d$, where $p^*(N) \to 1$ as $N \to \infty$. This domination means that for any increasing Borel set of graphs $\mathbf{B}$, the probability that the subgraph of open sites under independent bond percolation with parameter $p^*(N)$ lies in $\mathbf{B}$, is at most $\mathbf{P}[A_p(N) \in \mathbf{B}]$. If N is sufficiently large, then the infinite cluster determined by bond percolation with parameter $p^*(N)$ on the lattice $N\mathbf{Z}^d$, is a.s. transient. The set of subgraphs of $N\mathbf{Z}^d$ that contain a transient subgraph is increasing, so $A_p(N)$ contains a transient subgraph $\widehat{A}_p(N)$ with probability 1. Observe that $\widehat{A}_p(N)$ is isomorphic to a subgraph of the "$3N^d$-fuzz" of the infinite cluster $\mathcal{C}_p$ in the original lattice, so by Rayleigh's monotonicity principle, we conclude that $\mathcal{C}_p$ is also transient a.s. (See Ex. 9.5, or §8.4 in Doyle and Snell (1984) for the definition and properties of the k-fuzz of a graph.) Alternatively, it can be verified that $\widehat{A}_p(N)$ is "roughly isometric" to a subgraph of $\mathcal{C}_p$, and therefore $\mathcal{C}_p$ is transient a.s. (see Soardi 1994). □

Remark. Hiemer (1998) proved a renormalization theorem for oriented percolation, that allowed him to extend the result of [6] on transience of oriented percolation clusters in $\mathbf{Z}^d$ for $d \geq 3$, from the case of high p to the whole supercritical phase for oriented percolation.

Recall that a collection of edges Π is a **cutset separating** v_0 **from** ∞ if any infinite self-avoiding path emanating from v_0 must intersect Π. Nash-Williams proved that if $\{\Pi_n\}_{n=1}^{\infty}$ is a sequence of disjoint cutsets separating v_0 from infinity in a connected transient graph, then $\sum_n |\Pi_n|^{-1} < \infty$.

The following extension of Theorem 11.3 provides finer information about the permissible growth rates of cutsets on supercritical infinite percolation clusters.

Exercise 11.4 *Show that for $d \geq 2$,*

$$\inf\{q : \exists \text{ a flow } f \neq 0 \text{ from } 0 \text{ to } \infty \text{ on } \mathbf{Z}^d \text{ with } \sum |f(e)|^q < \infty\} = \frac{d}{d-1}\,.$$

Theorem 11.5 (Levin and Peres 1998) *Let $\mathcal{C}_\infty(\mathbf{Z}^d, p)$ be the infinite cluster of Bernoulli(p) percolation on $\mathbf{Z}^d$. Then for $d \geq 3$ and $p > p_c(\mathbf{Z}^d)$, a.s.,*

$$\inf\{q : \exists \text{ a flow } f \neq 0 \text{ from } 0 \text{ to } \infty \text{ on } \mathcal{C}_\infty(\mathbf{Z}^d, p) \text{ with } \sum_e |f(e)|^q < \infty\} = \frac{d}{d-1}\,.$$

Corollary 11.6 *Let $d \geq 3$ and $p > p_c(\mathbf{Z}^d)$. With probability one, if $\{\Pi_n\}$ is a sequence of disjoint cutsets in the infinite cluster $\mathcal{C}_\infty(\mathbf{Z}^d, p)$ that separate a fixed vertex v_0 from ∞, then $\sum_n |\Pi_n|^{-\beta} < \infty$ for all $\beta > \frac{1}{d-1}$.*

Proof. Pick $\beta > \frac{1}{d-1}$, and let f be a unit flow on $\mathcal{C}_\infty(\mathbf{Z}^d, p)$ with $\sum |f(e)|^{1+\beta} < \infty$, which exists by Theorem 11.5. Observe first that

$$\mathcal{E}_{1+\beta}(f) = \sum_{e \in E_G} |f(e)|^{1+\beta} \geq \sum_n \sum_{e \in \Pi_n} |f(e)|^{1+\beta},$$

since the $\{\Pi_n\}$ are disjoint. By Jensen's inequality, for all $n \geq 1$,

$$\frac{1}{|\Pi_n|} \sum_{e \in \Pi_n} |f(e)|^{1+\beta} \geq \Big(\frac{1}{|\Pi_n|} \sum_{e \in \Pi_n} |f(e)|\Big)^{1+\beta} = |\Pi_n|^{-1-\beta}\,.$$

Multiplying by $|\Pi_n|$ and summing over n establishes the Corollary. □

Remark. Theorem 11.5 was recently sharpened by Hoffman and Mossel.

12 Subperiodic Trees

For a tree Γ, let Γ^v denote the subtree of Γ rooted at vertex v that contains all descendants of v. Γ is **N-subperiodic** if for any vertex $v \in \Gamma$ there exists a 1-1 adjacency preserving map $f\colon \Gamma^v \to \Gamma^{f(v)}$ with $|f(v)| \le N$.

Example 12.1 *Examples of subperiodic trees.*

- b-ary trees for any integer $b \ge 2$.
- The Fibonacci tree Γ_{fib} described in Exercise 5.6.
- The tree of all self-avoiding walks in $\mathbf{Z}^d$.
- Directed covers of finite connected directed graphs: to every directed path of length n in the graph corresponds $v \in \Gamma$ with $|v| = n$; extensions of the path correspond to descendants of v.
- Universal covers of undirected graphs: to every non-backtracking path of length n in the graph corresponds $v \in \Gamma$ with $|v| = n$; extensions correspond to descendants, as above.

$\triangle$

Suppose that $b \ge 2$ is an integer. For a closed nonempty set $\Lambda \subseteq [0,1]$, define a tree $\Gamma(\Lambda, b)$ as follows. Consider the system of b-adic subintervals of $[0,1]$; those which have a non-empty intersection with Λ form the vertices of the tree. Two vertices are connected by an edge if one of the corresponding intervals is contained in the other and their orders differ by one (i.e., the ratio of lengths is b). The root of this tree is $[0,1]$. Clearly, $\Gamma([0,1], b)$ is the usual b-ary tree. If $b\Lambda (\text{mod } 1) \subset \Lambda$, i.e., Λ is invariant under the transformation $x \mapsto bx (\text{mod } 1)$, then $\Gamma(\Lambda, b)$ is 0-subperiodic.

Theorem 12.2 (Furstenberg 1967) *For Γ which is subperiodic, $\text{gr}(\Gamma)$ exists and $\text{gr}(\Gamma) = \text{br}(\Gamma)$. Furthermore,*

$$\inf_{\Pi} S(\text{br}(\Gamma), \Pi) > 0\,,$$

where $S(\lambda, \Pi) = \sum_{v \in \Pi} \lambda^{-|v|}$ for a cutset Π.

Corollary 12.3 (Furstenberg's formulation) *Let $\Lambda \subseteq [0,1]$ be a compact set. If $b\Lambda (\text{mod } 1) \subseteq \Lambda$, then*

$$\dim_H(\Lambda) = \dim_M(\Lambda) = \beta$$

for some β, and moreover, $\mathcal{H}^\beta(\Lambda) > 0$, where $\mathcal{H}^\beta$ denotes β-dimensional Hausdorff measure.

Proof of Theorem 12.2. We will give the proof for Γ 0-subperiodic. The N-subperiodic case can be reduced to the 0-subperiodic case; this reduction is left as an exercise. Assume first that Γ has no leaves.

Suppose that for some finite cutset Π,

$$S(\lambda, \Pi) < 1 \,. \tag{32}$$

Denote $d = \max_{v \in \Pi} |v|$. By 0-subperiodicity, for any $v \in \Pi$, there exists a cutset $\Pi(v)$ of Γ^v such that

$$\sum_{w \in \Pi(v)} \lambda^{-(|w|-|v|)} < 1 \,.$$

In other words,

$$\sum_{w \in \Pi(v)} \lambda^{-|w|} < \lambda^{-|v|} \,.$$

Replace v in Π by the vertices in $\Pi(v)$ to obtain a new cutset $\tilde{\Pi}$ in Γ with $S(\lambda, \tilde{\Pi}) < 1$. Given n, repeat this kind of replacement for every vertex v in the current cutset with $|v| \le n$ to get a cutset Π^* such that all vertices $u \in \Pi^*$ satisfy $n \le |u| \le n + d$. Then

$$|\Gamma_n| \lambda^{-n-d} \le S(\lambda, \Pi^*) < 1 \,.$$

This inequality depends on the assumption of no leaves. Thus $|\Gamma_n| < \lambda^{n+d}$ for all n, whence $\overline{\mathrm{gr}}(\Gamma) \le \lambda$. Since (32) holds for any $\lambda > \mathrm{br}(\Gamma)$, we infer that $\overline{\mathrm{gr}}(\Gamma) < \lambda$. Therefore

$$\overline{\mathrm{gr}}(\Gamma) \le \mathrm{br}(\Gamma) \le \underline{\mathrm{gr}}(\Gamma) \,.$$

Finally, consider $\lambda_1 = \mathrm{br}(\Gamma)$. If $S(\lambda_1, \Pi) < 1$ for some finite cutset Π, then we could find $\lambda < \lambda_1$ such that $S(\lambda, \Pi) < 1$, and the preceding argument would yield that $\overline{\mathrm{gr}}(\Gamma) \le \lambda < \lambda_1$, a contradiction. Thus for all cutsets Π,

$$S(\mathrm{br}(\Gamma), \Pi) \ge 1.$$

If Γ has leaves, create a modified tree Γ' by attaching to each leaf an infinite path. Γ' is periodic as well, and so the theorem can be applied to it, yielding $\mathrm{br}(\Gamma') = \mathrm{gr}(\Gamma')$. But since $\mathrm{br}(\Gamma) = \mathrm{br}(\Gamma')$ and $\overline{\mathrm{gr}}(\Gamma) \le \overline{\mathrm{gr}}(\Gamma')$, we have

$$\mathrm{br}(\Gamma) \le \underline{\mathrm{gr}}(\Gamma) \le \overline{\mathrm{gr}}(\Gamma) \le \overline{\mathrm{gr}}(\Gamma') = \mathrm{br}(\Gamma') = \mathrm{br}(\Gamma),$$

and hence $\mathrm{gr}(\Gamma) = \mathrm{br}(\Gamma)$. $\square$

Exercise 12.4 *Construct a subperiodic tree with superlinear polynomial growth (more precisely, construct a subperiodic tree T such that $|T_n \| to\infty$ as $n \to \infty$, but $|T_n| = O(n^d)$ for some $d < \infty$.*

(Hint: build a subtree of the binary tree where all finite paths are labeled by words in the Morse sequence 0110100110010110.... This sequence is obtained by iterating the substitution $0 \mapsto 01$, $1 \mapsto 10$. Alternatively, use a lexicographic spanning tree in $\mathbf{Z}^d$, as described in the next chapter.)

Exercise 12.5 *Does every subperiodic tree with exponential growth have a subtree without leaves that has bounded pipes?*

(Hint: Consider the subtree T of the binary tree T_2, containing all self-avoiding paths from the root in T_2 with the property that for every $n > 100$, any n^2 consecutive edges on the path contain a run of n consecutive left turns.)

13 The Random Walks RW_λ

For a graph G, fix an origin o, and define $|e|$ as the length of a shortest path from o to an end-vertex of e. We will define a family of processes RW_λ as weighted random walks on G. Specifically, each edge e is assigned conductance $\lambda^{-|e|}$. We will mostly consider the case where Γ is a tree and o is the root ρ, although we will also consider these processes defined on Cayley graphs of groups. By fine tuning λ, we obtain random walks that explore the graph better than the simple random walk. The following result is stronger than Theorem 2.9 mentioned in Chapter 2.

Theorem 13.1 (R. Lyons 1990) RW_λ *is transient on a tree* Γ *if* $\lambda < \mathrm{br}(\Gamma)$, *and recurrent if* $\lambda > \mathrm{br}(\Gamma)$.

Proof. If $\lambda > \mathrm{br}(\Gamma)$, then for any ϵ there exists a cutset Π such that $\sum_{v\in\Pi}\lambda^{-|v|} < \epsilon$. By Nash-Williams (for just one cutset)

$$\mathcal{R}(\rho \leftrightarrow \infty) \geq \frac{1}{\sum\limits_{v\in\Pi}\lambda^{-|v|}} > \frac{1}{\epsilon}\,.$$

Letting $\epsilon \downarrow 0$ shows that $\mathcal{R}(\rho \leftrightarrow \infty)$ is infinite, and hence the walk is recurrent.

If $\lambda < \mathrm{br}(\Gamma)$ choose $\lambda < \lambda_* < \mathrm{br}(\Gamma)$ so that there exists a unit flow θ from ρ to ∞ with $\theta(e) \leq C\lambda_*^{-|e|}$. Then

$$\mathcal{E}(\theta) = \sum_e r_e[\theta(e)]^2 \leq \sum_n \lambda^n \sum_{|e|=n}\theta(e)C\lambda_*^{-|e|} = C\sum_n \left(\frac{\lambda}{\lambda_*}\right)^n \sum_{|e|=n}\theta(e) < \infty\,,$$

since θ is a unit flow. $\square$

Let G be a countable group with a finite set of generators $S = \langle g_1, \ldots g_m\rangle$. With every generator we include its inverse, so $S = S^{-1}$. The **Cayley graph** of G has as vertices the elements of the group, and contains an (unoriented) edge between u and v if $u = g_i v$ for some $g_i \in S$. Each element $g \in G$ can be represented as a word in the generators, $g = g_{i(1)}\cdots g_{i(m)}$; let $|g|$ be the minimal length of words which represent g, and let $G_n = \{g \in G : |g| = n\}$. The growth $\mathrm{gr}(G) := \lim_n |G_n|^{1/n}$ exists for such groups, and the group is of exponential growth if $\mathrm{gr}(G) > 1$.

Corollary 13.2 (R. Lyons 1995) RW_λ *on the Cayley graph of a group G of exponential growth is transient for* $\lambda < \mathrm{gr}(G)$ *and recurrent for* $\lambda > \mathrm{gr}(G)$.

Proof. The second statement follows from the Nash-Williams inequality. For the first, we will show that random walk on a subgraph is transient; by Rayleigh's Monotonicity Principle, this is enough. We will use the lexicographic spanning tree Γ in G. Assign g its lexicographically minimal representation $g = g_{i(1)} \cdots g_{i(m)}$ where $m = |g|$ and if $g = g_{j(1)} \cdots g_{j(m)}$ is another representation of g, then at the smallest k such that $i(k) \neq j(k)$ we have $i(k) < j(k)$. The edge gh is in Γ if $|\,|g| - |h|\,| = 1$ and either g is an initial segment of h or h is an initial segment of g. Let the identity be the root. Since there is a unique path from the root to any element in Γ, and Γ contains all elements of G, it is indeed a spanning tree. One can check that it is 0–subperiodic.

Observe that $|\Gamma_n| = |G_n|$, so $\mathrm{gr}(\Gamma) = \mathrm{gr}(G)$. Since Γ is subperiodic, Theorem 12.2 implies that $\mathrm{br}(\Gamma) = \mathrm{gr}(G)$. By Theorem 13.1, for $\lambda < \mathrm{gr}(G)$ the biased walk RW_λ is transient on Γ, hence also on G. $\square$

Open Problem 1 *For* $1 < \lambda < \mathrm{gr}(G)$*, is it true that*

$$\mathrm{speed}(\mathrm{RW}_\lambda) := \lim_{n\to\infty} \frac{|X_n|}{n} > 0\,, \qquad a.s.\ ?$$

Here $|v|$ *denotes the distance of* v *from the identity.*

We remark that there exist groups of exponential growth where the speed of simple random walk is 0 a.s. An example is the simple random walk on the lamplighter group; see Lyons, Pemantle and Peres (1996).

14 Capacity

In Chapter 6 we considered capacity on the boundary of a tree. We now generalize the definition to any metric space X equipped with the Borel σ-field $\mathcal{B}$. A **kernel** F is a measurable function $F : X \times X \to [0, \infty]$. For a measure μ on $(X, \mathcal{B})$, the **energy** of μ in the kernel F is defined as

$$\mathcal{E}_F(\mu) = \int_X \int_X F(x, y)\, d\mu(x) d\mu(y)\,.$$

We will mostly consider F of the form $F(x, y) = f(|x - y|)$ for f non-negative and non-increasing; we write $\mathcal{E}_f$ for $\mathcal{E}_F$ in this case. Define the **capacity** of a set Λ in the kernel F as

$$\mathrm{Cap}_F(\Lambda) \;=\; \left[\inf_{\mu:\mu(\Lambda)=1} \mathcal{E}_F(\mu)\right]^{-1}.$$

The first occurrence of capacity in probability theory was the following result.

Theorem 14.1 (Kakutani 1944a, 1944b) *If* $\Lambda \subset \mathbf{R}^d$ *is compact with* $0 \notin \Lambda$ *and* B *is a Brownian motion, then*

$$\mathbf{P}_0(B \textit{ hits } \Lambda) > 0 \;\; \textit{if and only if} \;\; \mathrm{Cap}_G(\Lambda) > 0\,,$$

where G is the Green kernel

$$G(x,y) \;=\; \begin{cases} |x-y|^{2-d} & d \geq 3\,, \\ \log^+\left(|x-y|^{-1}\right) & d = 2\,. \end{cases}$$

R. Lyons discovered connections between capacity and percolation on trees, already discussed in Chapter 6. Let $\{p_e\}$ be a set of probabilities indexed by the edges of a tree Γ. Let $\mathtt{path}(v)$ denote the unique path from the root to v, and let F be the kernel

$$F(x,y) \;=\; \prod_{e \in \mathtt{path}(x \wedge y)} p_e^{-1}\,. \tag{33}$$

If $p_e \equiv p$, then $F(x,y) = p^{-|x \wedge y|}$. More generally, if $\mathbf{P}$ is the probability measure corresponding to independent $\{p_e\}$ percolation, then $F(x,y) = [\mathbf{P}(\rho \leftrightarrow x \wedge y)]^{-1}$. A. H. Fan proved that on an infinite tree of bounded degree, $\mathbf{P}(\rho \leftrightarrow \partial\Gamma) > 0$ iff $\mathrm{Cap}_F(\partial\Gamma) > 0$. This was sharpened by R. Lyons to a quantitative estimate.

Theorem 14.2 (R. Lyons 1992) *Let $\mathbf{P}$ be the probability measure corresponding to independent $\{p_e\}$ percolation on a tree Γ and F the kernel defined in (33). Then*

$$\mathrm{Cap}_F(\partial\Gamma) \;\leq\; \mathbf{P}(\rho \leftrightarrow \partial\Gamma) \;\leq\; 2\mathrm{Cap}_F(\partial\Gamma)\,. \tag{34}$$

Consider Brownian motion in dimension $d \geq 3$. One obstacle to obtaining quantitative estimates for Brownian hitting probabilities with capacity in Green's kernel is translation invariance of that kernel: If B is a Brownian motion started at the origin, then $\mathbf{P}(B \,\text{hits}\, \Lambda + x)$ becomes small as $x \to \infty$. If we had a scale invariant kernel instead, we would have more hope, as $\mathbf{P}(B \text{ hits } c\Lambda) = \mathbf{P}(B \text{ hits } \Lambda)$ for any $c > 0$. Hence we use capacity in the **Martin kernel**

$$K(x,y) \;=\; \frac{G(x,y)}{G(0,y)} \;=\; \left(\frac{|y|}{|x-y|}\right)^{d-2} \tag{35}$$

for $d \geq 3$.

Theorem 14.3 (Benjamini, Pemantle, and Peres 1995) *Let B be a Brownian motion in $\mathbf{R}^d$ for $d \geq 3$, started at the origin. Let K be the Martin kernel defined in (35). Then for any closed set Λ in $\mathbf{R}^d$,*

$$\frac{1}{2}\mathrm{Cap}_K(\Lambda) \leq \mathbf{P}_0(B \textit{ hits } \Lambda) \leq \mathrm{Cap}_K(\Lambda)\,.$$

Remark. An analogous statement holds for planar Brownian motion, provided it is killed at an appropriate finite stopping time (e.g., an independent exponential time, or the first exit from a bounded domain) and the corresponding Green function $G(x,y)$ is used to define the Martin Kernel.

Theorem 14.4 (BPP 1995) *Let $\{X_n\}$ be a transient Markov chain on a countable state space S with initial state $\rho \in S$, and set*

$$G(x,y) \;=\; \mathbf{E}_x\left[\sum_{n=0}^{\infty} \mathbf{1}_{\{y\}}(X_n)\right] \text{ and } K(x,y) \;=\; \frac{G(x,y)}{G(\rho,y)}\,.$$

Then for any initial state ρ and any subset Λ of S,

$$\frac{1}{2}\mathrm{Cap}_K(\Lambda) \le \mathbf{P}_\rho(\{X_n\} \text{ hits } \Lambda) \le \mathrm{Cap}_K(\Lambda)\,.$$

Exercise 14.5 *Verify the analogous result for the stable-$\frac{1}{2}$ subordinator and the kernel*

$$G(s,t) := \begin{cases} (t-s)^{-1/2} & 0 < s \le t\,, \\ 0 & s > t > 0\,. \end{cases}$$

Problem: Find the class of Markov processes for which the above estimate (for suitable kernel G and resulting K) holds.

Proof of Theorem 14.4. To prove the right-hand inequality, we may assume that the hitting probability is positive. Let $\tau = \inf\{n : X_n \in \Lambda\}$ and let ν be the measure $\nu(A) = \mathbf{P}_\rho(\tau < \infty \text{ and } X_\tau \in A)$. In general, ν is a sub-probability measure, as τ may be infinite. By the Markov property, for $y \in \Lambda$,

$$\int_\Lambda G(x,y)d\nu(x) = \sum_{x\in\Lambda} \mathbf{P}_\rho(X_\tau = x)G(x,y) = G(\rho,y)\,,$$

whence $\int_\Lambda K(x,y)d\nu(x) = 1$. Therefore $\mathcal{E}_K(\nu) = \nu(\Lambda)$, $\mathcal{E}_K(\nu/\nu(\Lambda)) = [\nu(\Lambda)]^{-1}$; consequently, since $\nu/\nu(\Lambda)$ is a probability measure,

$$\mathrm{Cap}_K(\Lambda) \;\ge\; \nu(\Lambda) \;=\; \mathbf{P}_\rho(\{X_n\} \text{ hits } \Lambda)\,.$$

This yields one inequality. Note that the Markov property was used here.

For the reverse inequality, we use the second moment method. Given a probability measure μ on Λ, set

$$Z = \int_\Lambda \sum_{n=0}^{\infty} \mathbf{1}_{\{y\}}(X_n)\frac{d\mu(y)}{G(\rho,y)}\,.$$

$\mathbf{E}_\rho[Z] = 1$, and the second moment satisfies

$$\begin{aligned} \mathbf{E}_\rho[Z^2] \;&=\; \mathbf{E}_\rho \int_\Lambda\int_\Lambda \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \mathbf{1}_{\{x\}}(X_m)\mathbf{1}_{\{y\}}(X_n)\frac{d\mu(x)d\mu(y)}{G(\rho,x)G(\rho,y)} \\ &\le\; 2\mathbf{E}_\rho \int_\Lambda\int_\Lambda \sum_{m\le n} \mathbf{1}_{\{x\}}(X_m)\mathbf{1}_{\{y\}}(X_n)\frac{d\mu(x)d\mu(y)}{G(\rho,x)G(\rho,y)}\,. \end{aligned}$$

Observe that

$$\sum_{m=0}^{\infty} \mathbf{E}_\rho \sum_{n=m}^{\infty} \mathbf{1}_{\{x\}}(X_m)\mathbf{1}_{\{y\}}(X_n) = \sum_{m=0}^{\infty} \mathbf{P}_\rho(X_m = x)G(x,y) = G(\rho,x)G(x,y)\,.$$

Hence

$$\mathbf{E}_\rho[Z^2] \le 2\int_\Lambda\int_\Lambda \frac{G(x,y)}{G(\rho,y)}d\mu(x)d\mu(y) = 2\mathcal{E}_K(\mu)\,,$$

and therefore

$$\mathbf{P}_\rho(\{X_n\} \text{ hits } \Lambda) \ge \mathbf{P}_\rho(Z>0) \ge \frac{(\mathbf{E}_\rho[Z])^2}{\mathbf{E}_\rho[Z^2]} \ge \frac{1}{2\mathcal{E}_K(\mu)}\,.$$

We conclude that $\mathbf{P}_\rho(\{X_n\} \text{ hits } \Lambda) \ge \frac{1}{2}\mathrm{Cap}_K(\Lambda)$. □

The upper bound on $\mathbf{P}(\rho \leftrightarrow \partial\Gamma)$ obtained by the first moment method (8) is not sharp enough to prove Theorem 14.2. For example, take the binary tree with Bernoulli(p) percolation for $p=\frac{1}{2}$; if $\Gamma_n = \{v\colon |v| \le n\}$, then the first-moment method yields an upper bound of 1 for any n, while $\mathrm{Cap}_F(\partial\Gamma_n) = 2(n+2)^{-1}$. However, we can use Theorem 14.4 to give a short proof of Theorem 14.2.

Proof of Theorem 14.2. The first inequality was already proven in Proposition 7.1.

It remains to prove the right-hand inequality in (34). Assume first that Γ is finite. There is a Markov chain $\{V_k\}$ hiding here: Embed Γ in the lower half-plane, with the root at the origin. The random set of $r \ge 0$ leaves that survive the percolation may be enumerated from left to right as $V_1, V_2, \ldots, V_r$. The key observation is that the random sequence $\rho, V_1, V_2, \ldots, V_r, \Delta, \Delta, \ldots$ is a Markov chain on the state space $\partial\Gamma \cup \{\rho, \Delta\}$, where ρ is the root and Δ is a formal absorbing cemetery.

Indeed, given that $V_k = x$, all the edges on the unique path from ρ to x are retained, so that survival of leaves to the right of x is determined by the edges strictly to the right of the path from ρ to x, and is thus conditionally independent of $V_1, \ldots, V_{k-1}$. This verifies the Markov property, so Theorem 14.4 may be applied.

The transition probabilities for the Markov chain above are complicated, but it is easy to write down the Green kernel. Clearly, $G(\rho, y)$ equals the probability that y survives percolation, so

$$G(\rho,y) = \prod_{e\in \mathtt{path}(y)} p_e\,.$$

If x is to the left of y, then $G(x,y)$ is equal to the probability that the range of the Markov chain contains y given that it contains x, which is just the probability of y surviving given that x survives. Therefore,

$$G(x,y) = \prod_{e\in \mathtt{path}(y)\setminus\mathtt{path}(x)} p_e\,,$$

and hence

$$K(x,y) = \frac{G(x,y)}{G(\rho,y)} = \prod_{e\in\mathtt{path}(x\wedge y)} p_e^{-1}\,.$$

Now $G(x,y)=0$ for x on the right of y; thus (keeping the diagonal in mind)

$$F(x,y) \le K(x,y) + K(y,x)$$

for all $x, y \in \partial\Gamma$, and therefore

$$\mathcal{E}_F(\partial\Gamma) \leq 2\mathcal{E}_K(\partial\Gamma) .$$

Now apply Theorem 14.4 to $\Lambda = \partial\Gamma$:

$$\mathrm{Cap}_F(\partial\Gamma) \geq \frac{1}{2}\mathrm{Cap}_K(\partial\Gamma) \geq \frac{1}{2}\mathbf{P}(\{V_k\} \text{ hits } \partial\Gamma) = \frac{1}{2}\mathbf{P}(\rho \leftrightarrow \partial\Gamma) .$$

This establishes the upper bound for finite Γ.

The inequality for general Γ follows from the finite case by taking limits. $\square$

Remark. The inequality (34) was recently sharpened by Marchal [68].

The notation $\mathcal{E}$ has appeared twice, once as a functional on *flows* and once as a functional on *measures*. As discussed following Lemma 2.10, measures on the boundary of a tree correspond to flows on the tree; we shall see that the energy of a measure on $\partial\Gamma$ is (up to an additive constant) the same as the energy of the corresponding flow on Γ: Given a measure μ on $\partial\Gamma$, let θ be the corresponding flow: $\theta(uv) = \mu(\xi : v \in \xi)$, where u is the parent of v. Observe that

$$\mathcal{E}(\theta) \;=\; \sum_e r_e \theta(e)^2 \;=\; \sum_e r_e \int_{\partial\Gamma}\int_{\partial\Gamma} \mathbf{1}_{\{\xi \ni e\}}\mathbf{1}_{\{\eta \ni e\}} d\mu(\xi) d\mu(\eta) .$$

Moving the sum inside the integral, the above equals

$$\int_{\partial\Gamma}\int_{\partial\Gamma} \sum_e \mathbf{1}_{\{e \in \xi \cap \eta\}} r_e d\mu(\xi) d\mu(\eta) \;=\; \int_{\partial\Gamma}\int_{\partial\Gamma} \sum_{e \leq \xi \wedge \eta} r_e d\mu(\xi) d\mu(\eta) .$$

By the series law for resistances, we are left with

$$\mathcal{E}(\theta) \;=\; \int_{\partial\Gamma}\int_{\partial\Gamma} \mathcal{R}(\rho \leftrightarrow \xi \wedge \eta) d\mu(\xi) d\mu(\eta) . \tag{36}$$

Now if

$$1/\mathcal{C}(\rho \leftrightarrow v) + 1 = 1/\mathbf{P}(\rho \leftrightarrow v) , \tag{37}$$

then substituting in (36) yields

$$\mathcal{E}_K(\mu) = 1 + \mathcal{E}(\theta) , \tag{38}$$

where $K(\xi, \eta) = 1/\mathbf{P}(\rho \leftrightarrow \xi \wedge \eta)$. By taking infimum on both sides of (38) and applying Thomson's Principle, we can rewrite Theorem 14.2: If the correspondence (37) holds for resistances $\{r_e\}$ and an independent $\{p_e\}$ percolation $\mathbf{P}$, then

$$\frac{1}{1 + \mathcal{R}(\rho \leftrightarrow \infty)} \leq \mathbf{P}(\rho \leftrightarrow \infty) \leq \frac{2}{1 + \mathcal{R}(\rho \leftrightarrow \infty)} . \tag{39}$$

It is easily checked that in the case of Bernoulli(p) percolation, the correspondence (37) is preserved by taking $c_e = (1-p)^{-1} p^{|v|}$, where e is the edge connecting v to its parent. In this case the weighted random walk on the resulting network is $\mathrm{RW}_{1/p}$. Thus, (39) implies that percolation occurs at p if and only if $\mathrm{RW}_{1/p}$ is transient.

Consider a Cantor set Λ in the unit interval and the corresponding tree $\Gamma(\Lambda, b)$. We shall see that simple random walk on this tree is transient iff Λ, considered as a subset of $\mathbf{R}^2$, is non-polar for Brownian motion. In particular, transience of $\Gamma(\Lambda, b)$ is independent of b. The following theorem can be found in Benjamini and Peres (1992) in a special case, and in Pemantle and Peres (1995b) in general.

Theorem 14.6 *Let Γ be a subtree of the b^d-adic tree and let $f : (0,\infty) \to (0,\infty)$ be a non-increasing function with $f(0+) = \infty$. Let Ψ be the canonical map from the boundary of the b^d-adic tree to $[0,1]^d$; Ψ^{-1} is base-b representation of points in $[0,1]^d$. Let $\mathrm{dist}(v,w) = b^{-|v\wedge w|}$ for $v, w \in \partial\Gamma$ and let $\mathrm{dist}(x,y)$ be Euclidean distance for $x, y \in [0,1]^d$. Then*

$$\mathrm{Cap}_f(\partial\Gamma) \asymp \mathrm{Cap}_f(\Psi(\partial\Gamma))\,,$$

where Cap_f stands for capacity in the kernel $F(x,y) = f(\mathrm{dist}(x,y))$. This means there exist constants c and C, depending on b and d only, such that

$$c\,\mathrm{Cap}_f(\Psi(\partial\Gamma)) \le \mathrm{Cap}_f(\partial\Gamma) \le C\,\mathrm{Cap}_f(\Psi(\partial\Gamma))\,.$$

Exercise 14.7 *Consider Bernoulli(p) percolation on an infinite tree Γ. Prove that*

$$\mathbf{P}_p(\textit{component of } \rho \textit{ is transient}) > 0 \;\text{ iff }\; \mathbf{P}_{\{p_e\}}(\rho \leftrightarrow \partial\Gamma) > 0\,,$$

where $p_e = \frac{k}{k+1}p$ when $|e| = k$.

Hint: An infinite tree T is transient iff $\mathrm{Cap}_{|x\wedge y|}(\partial T) > 0$. The kernel $|x\wedge y|$ is obtained by applying $f(r) = -\log_b r$ to the distance between x and y.

Proof of Theorem 14.6 For $v \in \Gamma$, let $\mu(v) = \mu(\xi : \xi \ni v)$. We will prove that $\mathcal{E}_f(\mu) \asymp \mathcal{E}_f(\mu\Psi^{-1})$, i.e.,

$$c(b,d) \le \frac{\mathcal{E}_f(\mu)}{\mathcal{E}_f(\mu\Psi^{-1})} \le C(b,d) \tag{40}$$

for some constants $0 < c(b,d) \le C(b,d) < \infty$, depending on b and d only. This will yield $\mathrm{Cap}_f(\partial\Gamma) \asymp \mathrm{Cap}_f(\Psi(\partial\Gamma))$, proving the theorem.

Let

$$h(k) = \begin{cases} f(b^{-k}) - f(b^{1-k}), & k \ge 1 \\ f(1), & k = 0. \end{cases}$$

In the following, write $u \le w$ if w is a descendant of u. Then

$$\mathcal{E}_f(\mu) = \int_{\partial\Gamma}\int_{\partial\Gamma} \sum_{k=0}^{|x\wedge y|} h(k)\; d\mu(x)d\mu(y) = \sum_{k=0}^{\infty} h(k) \iint_{|x\wedge y|\ge k} d\mu(x)d\mu(y).$$

Breaking up the region of integration and observing that $x \wedge y \geq v$ iff $x \geq v$ and $y \geq v$, the above is equal to

$$\sum_{k=0}^{\infty} h(k) \sum_{|v|=k} \int\int_{x \wedge y \geq v} d\mu(x) d\mu(y) = \sum_{k=0}^{\infty} h(k) \sum_{|v|=k} [\mu(v)]^2 = \sum_{k=0}^{\infty} h(k) S_k,$$

where $S_k = S_k(\mu) = \sum_{|v|=k} [\mu(v)]^2$. Note that

$$\sum_{|v|=k+1} [\mu(v)]^2 \leq \sum_{|v|=k} [\mu(v)]^2 \leq b^d \sum_{|v|=k+1} [\mu(v)]^2 ,$$

i.e., $S_{k+1} \leq S_k \leq b^d S_{k+1}$.

We claim that in $[0,1]^d$,

$$\mathcal{E}_f(\mu\Psi^{-1}) \leq \int_{\Psi(\partial\Gamma)} \int_{\Psi(\partial\Gamma)} \sum_{k=0}^{\infty} h(k) \mathbf{1}_{\{k: b^{1-k} \geq |x-y|\}} \, d\mu\Psi^{-1}(x) d\mu\Psi^{-1}(y) .$$

This holds because for the largest k yielding a non-zero term in the sum above, $b^{-k} < |x-y|$ and thus the sum is bounded below by $f(|x-y|)$.

For vertices v, w at the same level of Γ, set $\chi(v,w) = 1$ iff $\Psi(v)$ and $\Psi(w)$ are the same or adjacent subcubes of $[0,1]^d$, and $\chi(v,w) = 0$ otherwise. Then

$$\mu \times \mu\{(\xi,\eta) \colon |\Psi(\xi) - \Psi(\eta)| \leq b^{1-k}\} \leq \sum_{|v|=k-1} \sum_{|w|=k-1} \mu(v)\mu(w)\chi(v,w) . \tag{41}$$

Now use the standard inequality $2\mu(v)\mu(w) \leq [\mu(v)]^2 + [\mu(w)]^2$ and the fact that the number of cubes adjacent to a given cube is bounded above by 3^d, to deduce that

$$\mu \times \mu\{(\xi,\eta) \colon |\Psi(\xi) - \Psi(\eta)| \leq b^{1-k}\} \leq 3^d S_{k-1} \leq 3^d b^d S_k.$$

It follows that

$$\mathcal{E}_f(\mu\Psi^{-1}) \leq (3b)^d \sum_k h(k) S_k = (3b)^d \mathcal{E}_f(\mu) .$$

For the reverse inequality, choose l so that $b^l \geq \sqrt{d}$. Then $|v \wedge w| = k + l$ implies that $|\Psi(v) - \Psi(w)| \leq b^{-k}$, and consequently

$$\begin{aligned} \mathcal{E}_f(\mu\Psi^{-1}) &\geq \sum_{k=0}^{\infty} f(b^{-k}) \, \mu \times \mu\{|v \wedge w| = k + l\} \\ &= \sum_{k=0}^{\infty} f(b^{-k}) \, [S_{k+l}(\mu) - S_{k+l+1}(\mu)] . \end{aligned}$$

Using summation-by-parts shows that the right-hand side above is equal to

$$\sum_{k=0}^{\infty} h(k) S_{k+l}(\mu) \;\geq\; b^{-dl} \sum_{k=0}^{\infty} h(k) S_k(\mu) \;=\; b^{-dl} \mathcal{E}_f(\mu) .$$

$\square$

15 Intersection-Equivalence

This Chapter follows Peres (1996). Throughout this chapter we work in $[0,1]^d$ and all processes considered are started according to the uniform measure on $[0,1]^d$, unless otherwise indicated.

Lemma 15.1 *If B is a Brownian path (killed at an exponential time for $d=2$), then*

$$\mathbf{P}(B \cap \Lambda \neq \emptyset) \asymp \mathrm{Cap}_g(\Lambda)$$

for any Borel set Λ, where

$$g(r) = \begin{cases} \log^+(r^{-1}) & \textit{if } d = 2 \\ r^{2-d} & \textit{if } d > 2 \end{cases} . \tag{42}$$

Proof. (for $d \geq 3$). Denote by K the Martin kernel, see (35). By Theorem 14.3,

$$\mathbf{P}(B \text{ hits } \Lambda) = \int_{[0,1]^d} \mathbf{P}_0(B \text{ hits } \Lambda - x)dx \geq \frac{1}{2} \int_{[0,1]^d} \mathrm{Cap}_K(\Lambda - x)dx .$$

Because $\mathcal{E}_K(\mu) \leq C_d \mathcal{E}_g(\mu)$ for any measure μ on $[0,1]^d$, the right-hand side above is bounded below by

$$\frac{1}{2C_d} \int_{[0,1]^d} \mathrm{Cap}_g(\Lambda - x)dx = \frac{1}{2C_d} \mathrm{Cap}_g(\Lambda) .$$

The upper-bound is a consequence of the probabilistic potential theory developed by Hunt and Doob. There exists a finite measure ν such that

$$\mathbf{P}_x(B \text{ hits } \Lambda) = \int_\Lambda g(|x-y|)d\nu(y) \qquad \text{and} \qquad \nu(\Lambda) = \mathrm{Cap}_g(\Lambda) .$$

(see, e.g., Chung (1973).) Then

$$\mathbf{P}(B \text{ hits } \Lambda) = \int_{[0,1]^d} \mathbf{P}_x(B \text{ hits } \Lambda)dx = \int_\Lambda \int_{[0,1]^d} g(|x-y|)dx d\nu(y) \leq C_d\, \nu(\Lambda) ,$$

where C_d is a constant depending only on d. Note that this proof extends to any initial distribution π for $B(0)$ with a bounded density; more generally a bounded Greenian potential suffices. $\square$

Shizuo Kakutani, generalizing a question of Paul Lévy, asked which compact sets Λ satisfy $\mathbf{P}(\Lambda \cap B_1 \cap B_2 \neq \emptyset) > 0$, where B_1, B_2 are independent Brownian paths in $\mathbf{R}^d$ ($d = 2$ or 3)?

Evans (1987) and Tongring (1988) gave a partial answer:

$$\text{If } \mathrm{Cap}_{g^2}(\Lambda) > 0 , \text{ then } \mathbf{P}(\Lambda \cap B_1 \cap B_2 \neq \emptyset) > 0 . \tag{43}$$

They also found a necessary condition involving the Hausdorff measure of Λ. Later Fitzsimmons and Salisbury (1989) gave the full answer: $\mathrm{Cap}_{g^2}(\Lambda) > 0$ is necessary as well as sufficient in (43). Furthermore, in dimension 2, their very general results yield the equivalence

$$\mathrm{Cap}_{g^k}(\Lambda) > 0 \quad \Leftrightarrow \quad P(\Lambda \cap B_1 \cap \ldots \cap B_k \neq \emptyset) > 0. \tag{44}$$

This led Chris Bishop to make the following insightful conjecture:

Conjecture 2 (Bishop) *Let B denote a Brownian path. Then for any nonincreasing gauge f and any closed set Λ, the event that $\mathrm{Cap}_f(\Lambda \cap B) > 0$ has positive probability iff $\mathrm{Cap}_{fg}(\Lambda) > 0$.*

We will present a proof of this below. Applying Kakutani's Theorem 14.1 to $\Lambda' = \Lambda \cap B_1$ and B_2 shows that

$$\mathbf{P}(\Lambda \cap B_1 \cap B_2 \neq \emptyset) > 0 \quad \Leftrightarrow \quad \mathrm{Cap}_g(\Lambda \cap B_1) > 0 \text{ with positive probability.} \tag{45}$$

Bishop's Conjecture (with $f = g$) along with (45) imply that

$$\mathrm{Cap}_{g^2}(\Lambda) > 0 \quad \Leftrightarrow \quad \mathbf{P}(\Lambda \cap B_1 \cap B_2 \neq \emptyset) > 0.$$

Hence Bishop's Conjecture and Kakutani's Theorem together give (44).

Theorem 15.2 *Let f be a non-negative and non-increasing function. Consider independent $\{p_e\}$ percolation on the 2^d-ary tree, with $p_e = p_k$ whenever $|e| = k$ and with $p_1 \ldots p_k = 1/f(2^{-k})$. Let $Q_d(f) \subseteq [0,1]^d$ be the set corresponding to $\partial\Gamma$ in $[0,1]^d$, where Γ is the component of the root in this percolation. (This component may be finite, whence $Q_d(f) = \emptyset$.) Then, for any closed set $\Lambda \subseteq [0,1]^d$,*

$$\mathrm{Cap}_f(\Lambda) \asymp \mathbf{P}(\Lambda \cap Q_d(f) \neq \emptyset). \tag{46}$$

For $f = g$ in particular, $Q_d(f)$ is intersection-equivalent to Brownian motion, i.e.,

$$\mathbf{P}(\Lambda \cap Q_d(g) \neq \emptyset) \asymp \mathbf{P}(\Lambda \cap B \neq \emptyset). \tag{47}$$

Proof. By Theorem 14.2,

$$\mathbf{P}(\Lambda \cap Q_d(f) \neq \emptyset) \ = \ \mathbf{P}_{\{p_e\}}(\rho \leftrightarrow \partial\Gamma(\Lambda, 2)) \ \asymp \ \mathrm{Cap}_f(\partial\Gamma(\Lambda, 2)), \tag{48}$$

where the constants in $\asymp$ are universal, namely 1 and 2. Theorem 14.6 with $b = 2$ yields

$$\mathrm{Cap}_f(\partial\Gamma(\Lambda, 2)) \ \asymp \ \mathrm{Cap}_f(\Lambda), \tag{49}$$

where the constants in $\asymp$ depend on d. Combining (48) and (49) establishes (46).

Finally, use (46) and Lemma 15.1 to prove (47). $\square$

Corollary 15.3 *Let f and h be non-negative and non-increasing functions. If a random closed set A in $[0,1]^d$ satisfies*

$$\mathbf{P}(A \cap \Lambda \neq \emptyset) \asymp \mathrm{Cap}_h(\Lambda) \tag{50}$$

for all closed $\Lambda \subseteq [0,1]^d$, then

$$\mathbf{P}(\mathrm{Cap}_f(A \cap \Lambda) > 0) > 0 \ \text{ if and only if } \ \mathrm{Cap}_{fh}(\Lambda) > 0 \tag{51}$$

for all closed $\Lambda \subseteq [0,1]^d$. In particular, Bishop's conjecture is true.

Proof. Enlarge the probability space where A is defined to include independent limit sets of fractal percolations $Q_d(f)$ and $\widetilde{Q}_d(h)$. By Theorem 15.2

$$\mathbf{P}(A \cap \Lambda \cap Q_d(f) \neq \emptyset \,|\, A) > 0 \ \text{ if and only if } \ \mathrm{Cap}_f(A \cap \Lambda) > 0\,,$$

it follows that

$$\mathbf{P}(\ \mathrm{Cap}_f(A \cap \Lambda) > 0\ ) > 0 \ \text{ if and only if } \ \mathbf{P}(A \cap \Lambda \cap Q_d(f) \neq \emptyset) > 0\,. \tag{52}$$

Conditioning on $Q_d(f)$ and then using (50) with $\Lambda \cap Q_d(f)$ in place of Λ gives

$$\mathbf{P}(A \cap \Lambda \cap Q_d(f) \neq \emptyset) > 0 \ \text{ if and only if } \ \mathbf{P}(\ \mathrm{Cap}_h(\Lambda \cap Q_d(f)) > 0\ ) > 0. \tag{53}$$

Conditioning on $Q_d(f)$ and applying Theorem 15.2 yields

$$\mathbf{P}(\ \mathrm{Cap}_h(\Lambda \cap Q_d(f)) > 0\ ) > 0 \ \text{ if and only if } \ \mathbf{P}(\Lambda \cap Q_d(f) \cap \widetilde{Q}_d(h) \neq \emptyset) > 0\,. \tag{54}$$

Since $Q_d(f) \cap \widetilde{Q}_d(h)$ has the same distribution as $Q_d(fh)$, Theorem 15.2 implies that

$$\mathbf{P}(\Lambda \cap Q_d(f) \cap \widetilde{Q}_d(h) \neq \emptyset) > 0 \ \text{ if and only if } \ \mathrm{Cap}_{fh}(\Lambda) > 0. \tag{55}$$

Combining (52),(53),(54), and (55) proves (51). $\square$

Corollary 15.4 *Suppose $\{A_i\}$ are independent random closed sets in $[0,1]^d$ satisfying*

$$\mathbf{P}(A_i \cap \Lambda \neq \emptyset) \asymp Cap_{g_i}(\Lambda)$$

for all closed $\Lambda \subseteq [0,1]^d$ and some g_i non-negative and non-increasing. Then

$$\mathbf{P}(A_1 \cap \ldots \cap A_k \cap \Lambda \neq \emptyset) > 0 \quad \Leftrightarrow \mathrm{Cap}_{g_1 \cdots g_k}(\Lambda) > 0\,.$$

Example 15.5 *A.s., two independent Brownian paths in $\mathbf{R}^4$ do not intersect.*

This is a well-known result of Dvoretsky, Erdős and Kakutani (1950); we will show how it follows from intersection-equivalence. Let B_1 and B_2 be two independent Brownian paths in $\mathbf{R}^4$, started uniformly in the cube $[0,1]^4$ and intersected with that cube. Each is intersection-equivalent to $Q_4(g)$, and thus

$$\mathbf{P}([0,1]^4 \cap B_1 \cap B_2 \neq \emptyset) \asymp \mathbf{P}(Q_4(g) \cap \widetilde{Q}_4(g) \neq \emptyset)\,, \tag{56}$$

where $\widetilde{Q}_4(g)$ is an independent copy of $Q_4(r^{-2})$. Because $Q_4(g) \cap \widetilde{Q}_4(g)$ has the same distribution as $Q_4(g^2)$,

$$\mathbf{P}([0,1]^4 \cap B_1 \cap B_2 \neq \emptyset) \asymp \mathbf{P}(Q_4(g^2) \neq \emptyset)\,. \tag{57}$$

Since the edge probabilities in the percolation corresponding to $g^2(r) = r^{-4}$ are all $p_k = 1/16$, the tree corresponding to $Q_4(g^2)$ is a critical branching process and thus dies out almost surely:

$$\mathbf{P}(Q_4(g^2) \neq \emptyset) = 0\,. \tag{58}$$

Putting together (57) and (58) shows that the two paths never intersect. △

Corollary 15.6 (Lawler (1982, 1985), Aizenman (1985)) *Let B_1 and B_2 be independent Brownian paths intersected with $[0,1]^d$, considered as sets in $[0,1]^d$. Then*

$$\mathbf{P}(\mathrm{dist}(B_1,B_2) < \epsilon) \asymp \begin{cases} 1 & d \leq 3 \\ \frac{1}{-\log\,\epsilon} & d = 4 \\ \epsilon^{d-4} & d > 4 \end{cases}.$$

Proof. We will prove the cases $d \geq 4$; the other cases are handled similarly. Let g be the Greenian potential (42), and write $Q_d(p)$ instead of $Q_d(g)$, where $p = 2^{2-d}$. For a closed set C and $\epsilon > 0$, let C^ϵ be the set of points within distance ϵ from a point in C. Conditioning on B_2^ϵ and applying Theorem 15.2 gives

$$\mathbf{P}(\mathrm{dist}(B_1,B_2) < \epsilon) \;=\; \mathbf{P}(B_1 \cap B_2^\epsilon \neq \emptyset) \;\asymp\; \mathbf{P}(Q_d(p) \cap B_2^\epsilon \neq \emptyset)\,. \tag{59}$$

Now conditioning on $[Q_d(p)]^\epsilon$ and again applying Theorem 15.2 yields

$$\mathbf{P}(Q_d(p) \cap B_2^\epsilon \neq \emptyset) \;=\; \mathbf{P}([Q_d(p)]^\epsilon \cap B_2 \neq \emptyset) \;\asymp\; \mathbf{P}([Q_d(p)]^\epsilon \cap \widetilde{Q}_d(p) \neq \emptyset)\,, \tag{60}$$

where $\widetilde{Q}_d(p)$ is an independent copy of $Q_d(p)$.

Combining (59) and (60) shows that

$$\mathbf{P}(\mathrm{dist}(B_1,B_2) < \epsilon) \;\asymp\; \mathbf{P}([Q_d(p)]^\epsilon \cap \widetilde{Q}_d(p) \neq \emptyset)\,. \tag{61}$$

Next let $\epsilon/2 < 2^{-k} \leq \epsilon$ and choose ℓ so that $2^\ell \geq \sqrt{d}$. Then $\mathbf{P}([Q_d(p)]^\epsilon \cap \widetilde{Q}_d(p) \neq \emptyset)$ is at most the probability that $Q_d(p)$ and $\widetilde{Q}_d(p)$ both intersect the interior of the same binary cube of side-length $2^{-(k+\ell)}$, and this is bounded below by

$$c^2 \cdot \mathbf{P}(\text{the construction leading to } Q_d(p^2) \text{ survives for } k+\ell \text{ generations})\,, \tag{62}$$

where $c = 1 - q > 0$ is the probability of survival of the (supercritical) branching process associated to the construction of $Q_d(p)$.

The probability in (62) may be estimated via standard branching process arguments, but we use percolation instead. Consider Bernoulli(p^2) percolation on the 2^d-ary tree T and write the probability as $\mathbf{P}_{p^2}(\rho \leftrightarrow T_{k+\ell})$. Since the minimal energy measure on ∂T_k is the uniform measure μ, Theorem 14.2 yields that

$$\frac{1}{\mathbf{P}_p(\rho \leftrightarrow T_k)} \asymp \frac{1}{\mathrm{Cap}_F(T_k)} = \mathcal{E}_F(\mu)\,,$$

where $F(v,w) = p^{-|v\wedge w|}$. We have

$$\mathcal{E}_F(\mu) \;=\; 1 + \sum_{v,w\in T_k}\sum_{j=1}^{k}(p^{-j} - p^{1-j})\mu(v)\mu(w) \;=\; 1 + \sum_{j=1}^{k}\sum_{|v\wedge w|>\geq j}(p^{-j} - p^{1-j})\mu(v)\mu(w)\,.$$

Since $|v \wedge w| \geq j$ if and only if $|v| \geq j$ and $|w| \geq j$,

$$\mathcal{E}_F(\mu) \;=\; 1 + \sum_{j=1}^{k}(p^{-j} - p^{1-j})\left(\sum_{|v|\geq j}\mu(v)\right)^2 \;=\; 1 + \sum_{j=1}^{k}(p^{-j} - p^{1-j})2^{-dj}\,,$$

where the last equality holds because μ is the uniform measure. We conclude that $\mathcal{E}_F(\mu) \asymp \sum_{j=1}^{k}(p2^d)^{-j}$ and

$$\frac{1}{\mathbf{P}_p(\rho \leftrightarrow T_k)} \asymp \begin{cases} k & \text{if } p = 2^{-d} \\ (2^d p)^{-k} & \text{if } p < 2^{-d}\,. \end{cases}$$

Recall that $p = 2^{2-d}$ and hence the probability in (62) is equal to

$$\mathbf{P}_{p^2}(\rho \leftrightarrow T_{k+\ell}) \asymp \begin{cases} (k+\ell)^{-1} \asymp |\log\epsilon|^{-1} & \text{if } d = 4, \text{ because } p^2 = 2^{-d} \text{ for } d = 4\,, \\ 2^{(4-d)(k+\ell)} \asymp \epsilon^{d-4} & \text{if } d > 4\,. \end{cases}$$

For the reverse inequality, recall (61):

$$\mathbf{P}(\mathrm{dist}(B_1, B_2) < \epsilon) \asymp \mathbf{P}([Q_d(p)]^\epsilon \cap \widetilde{Q}_d(p) \neq \emptyset)\,.$$

Let $Q_d^{k-1}(p)$ denote the union of all binary cubes of side-length 2^{1-k} in the $(k-1)$th step of the construction of $Q_d(p)$, and recall that $\epsilon/2 < 2^{-k} \leq \epsilon$. Then $[Q_d(p)]^\epsilon$ is contained in the union of 3^d translates $Q_d^{k-1}(p) + x$ of $Q_d^{k-1}(p)$ and therefore the probability $\mathbf{P}([Q_d(p)]^\epsilon \cap \widetilde{Q}_d(p) \neq \emptyset)$ is bounded above by

$$3^d\mathbf{P}(\text{the construction leading to } Q_d(p^2) \text{ survives to the } (k-1)\text{th generation}).$$

The proof is now concluded by using the previous calculation for this probability. □

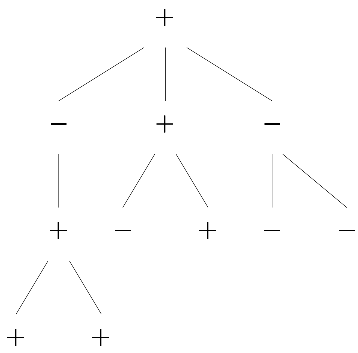


Figure 4: Tree with $+,-$ spins at the vertices.

16 Reconstruction for the Ising Model on a Tree

This chapter follows Evans, Kenyon, Peres and Schulman (1998).

Consider the following broadcast process. At the root ρ of a tree T, a random ± 1 valued "spin" σ_ρ is chosen uniformly. This spin is then propagated, with error, throughout the tree as follows: For a fixed $\epsilon \in (0, 1/2]$, each vertex receives the spin at its parent with probability $1 - \epsilon$, and the opposite spin with probability ϵ. These events at the vertices are statistically independent. This model has been studied in information theory, mathematical genetics and statistical physics; some of the history is described below.

Suppose we are given the spins that arrived at some fixed set of vertices W of the tree. Using the optimal reconstruction strategy (maximum likelihood), the probability of correctly reconstructing the original spin at the root is clearly at least $1/2$; denote this probability by $\frac{1+\Delta}{2}$. We will establish a lower bound for $\Delta = \Delta(T, W, \epsilon)$ in terms of the the effective electrical conductance from the root ρ to W (Theorem 16.2), and an upper bound for Δ which is the maximum flow from ρ to W for certain edge capacities (Theorem 16.3.) When T is an infinite tree, these bounds allow us to determine (in Theorem 16.1) the critical parameter ϵ_c so that, denoting the nth level of T by T_n, we have

$$\lim_{n\to\infty} \Delta(T, T_n, \epsilon) \begin{cases} > 0 & \text{if} \quad \epsilon < \epsilon_c \\ = 0 & \text{if} \quad \epsilon > \epsilon_c \,. \end{cases} \tag{63}$$

As we explain below, vanishing of the above limit is equivalent to extremality of the "free boundary" limiting Gibbs state for the ferromagnetic Ising model. For the special case of regular trees, the problem of determining ϵ_c was open for two decades, and was finally solved in 1995 by Bleher, Ruiz and Zagrebnov [12].

The random spins $\{\sigma_v\}$ that label the vertices of T as described above, can be constructed from independent variables $\{\eta_e\}$ labeling the edges of T, as follows. For each edge e, let $\mathbf{P}[\eta_e = -1] = \epsilon = 1 - \mathbf{P}[\eta_e = 1]$. Let σ_ρ be a uniformly chosen spin, and for any other vertex v let

$$\sigma_v := \sigma_\rho \prod_e \eta_e \,, \tag{64}$$

where the product is over all edges e on the path from ρ to v. Given $\sigma_W = \{\sigma_v : v \in W\}$, the strategy which maximizes the probability of correctly reconstructing σ_ρ, is to decide according to the sign of $\mathbf{E}(\sigma_\rho \mid \sigma_W)$; with this strategy, the difference between the probabilities of correct and incorrect reconstruction is

$$\Delta(T, W, \epsilon) = \mathbf{E}\Big|\mathbf{P}(\sigma_\rho = 1 \mid \sigma_W) - \mathbf{P}(\sigma_\rho = -1 \mid \sigma_W)\Big| . \tag{65}$$

Alternatively, $\Delta(T, W, \epsilon)$ can be interpreted as the *total variation distance* between the conditional distributions of σ_W given $\sigma_\rho = 1$ and given $\sigma_\rho = -1$; see below. The dependence between σ_ρ and σ_W is also captured by the *mutual information*

$$I(\sigma_\rho; \sigma_W) := \sum_{x,y} \mathbf{P}[\sigma_\rho = x, \sigma_W = y] \log \frac{\mathbf{P}[\sigma_\rho = x, \sigma_W = y]}{\mathbf{P}[\sigma_\rho = x]\mathbf{P}[\sigma_W = y]} .$$

Theorem 16.1 *Let T be an infinite tree with root ρ, and suppose its vertices are assigned random spins $\{\sigma_v\}$, using the flip probability $\epsilon < 1/2$ as in (64). Consider the problem of reconstructing σ_ρ from the spins at the n'th level T_n of T.*

(i) *If* $1 - 2\epsilon > \mathrm{br}(T)^{-1/2}$ *then* $\inf_{n \geq 1} \Delta(T, T_n, \epsilon) > 0$ *and* $\inf_{n \geq 1} I(\sigma_\rho; \sigma_{T_n}) > 0$.

(ii) *If* $1 - 2\epsilon < \mathrm{br}(T)^{-1/2}$ *then* $\inf_{n \geq 1} \Delta(T, T_n, \epsilon) = 0$ *and* $\inf_{n \geq 1} I(\sigma_\rho; \sigma_{T_n}) = 0$.

The tail field of the random variables $\{\sigma_v\}_{v \in T}$ contains events with probability strictly between 0 and 1 in case (i), but not in case (ii).

Thus in the notation of (63), $\epsilon_c = (1 - \mathrm{br}(T)^{-1/2})/2$. As mentioned above, this was already known when T is a $b{+}1$ - regular tree (for which $\mathrm{br}(T) = b$). Theorem 16.1 is considerably more general. Simple examples show that at criticality, when $1 - 2\epsilon = \mathrm{br}(T)^{-1/2}$, asymptotic solvability of the reconstruction problem is not determined by the branching number; in this case there is a sharp capacity criterion, proved in [75], that we will not develop here. To see the relevance of the quantity $1 - 2\epsilon$ appearing in Theorem 16.1, note the following equivalent construction of the random variables $\{\sigma_v\}$: Perform independent bond percolation on T with parameter $\gamma = 1 - 2\epsilon$ (the probability of open bonds), and independently assign to each of the resulting percolation clusters a uniform random spin (the same spin is assigned to all vertices in each cluster). This is a special case of the Fortuin-Kasteleyn random cluster representation of the Ising model (see, *e.g.*, [32]); on a tree, it is elementary to verify the equivalence of this representation with the construction (64).

The following two theorems contain estimates of reconstruction probability and mutual information, that imply Theorem 16.1.

Theorem 16.2 *Let T be a tree with root ρ, and let W be a finite set of vertices in T. Given $\epsilon \in (0, 1/2]$, denote $\gamma := 1 - 2\epsilon$, and consider the electrical network obtained by assigning to each edge e of T the resistance $(1 - \gamma^2)\gamma^{-2|e|}$. Then*

$$\left.\begin{array}{l} \Delta(T, W, \epsilon) \\ I(\sigma_\rho; \sigma_W) \end{array}\right\} \geq \frac{1}{1 + \mathcal{R}(\rho \leftrightarrow W)} , \tag{66}$$

where $\mathcal{R}$ denotes effective resistance.

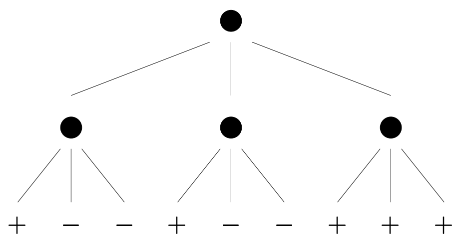


Figure 5: Majority vote can disagree with maximum likelihood.

The proof of this theorem is based on reconstruction by weighted majority vote, i.e., reconstruction according to the sign of an unbiased linear estimator of the root spin. We relate the variance of such an estimator to the energy of a corresponding unit flow from ρ to W. We find it quite surprising that on any infinite tree, reconstruction using such linear estimators has the same threshold as maximum-likelihood reconstruction.

Next, we present an upper bound on Δ and $I(\sigma_\rho; \sigma_W)$. Say that a set of vertices W_1 **separates ρ from W** if any path from ρ to W intersects W_1. For a vertex v of T, denote by $|v|$ the number of edges on the path from v to ρ.

Theorem 16.3 *Let W be a finite set of vertices in the tree T. For any set of vertices W_1 that separates the root ρ from W, we have*

$$\Delta(T, W, \epsilon)^2 \leq 2\Big(1 - \prod_{v \in W_1} \sqrt{1 - \gamma^{2|v|}}\Big) \leq 2 \sum_{v \in W_1} \gamma^{2|v|} \tag{67}$$

and

$$I(\sigma_\rho; \sigma_W) \leq \sum_{v \in W_1} I(\sigma_\rho; \sigma_v) \leq \sum_{v \in W_1} \gamma^{2|v|}. \tag{68}$$

In view of the mincut-maxflow theorem, (68) is an upper bound on mutual information in terms of the maximum flow in a capacitated network. Theorem 16.3 is proved by comparing the given tree T to a "stringy tree" $\widehat{T}$ which has an isomorphic set of paths from the root to the vertices of W_1, but these paths are pairwise edge-disjoint. We show that $\Delta(T, W, \epsilon) \leq \Delta(\widehat{T}, W_1, \epsilon)$ by constructing a noisy channel that maps the spins on W_1 in $\widehat{T}$ to the spins on W in T.

Symmetric trees: Recall that a tree T is **spherically symmetric** if for every $n \geq 1$, all vertices in T_n have the same degree. For such a tree, the effective resistance from the root to level n is easily computed, and we infer from Theorems 16.1-16.3 that

$$\Big(2 + 2(1 - \gamma^2) \sum_{k=1}^{n} \frac{\gamma^{-2k}}{|T_k|}\Big)^{-1} \leq I(\sigma_\rho; \sigma_{T_n}) \leq \inf_{k \leq n} |T_k| \gamma^{2k} \tag{69}$$

and $(1 - 2\epsilon_c)^{-2} = \liminf_n |T_n|^{1/n}$.

The example in Figure 5 shows that even on a regular tree, majority vote can disagree with maximum likelihood when the spin configuration σ_{T_n} is given.

Given the boundary data in Figure 5, the root spin σ_ρ is more likely to be -1 than $+1$ provided that ϵ is sufficiently small, since $\sigma_\rho = +1$ requires 4 spin flips, while $\sigma_\rho = -1$ requires only 3 spin flips.

Organization of the rest of the chapter.

Next, we present background on the Ising model and some references to the statistical mechanics and genetics literatures. Then we infer Theorem 16.1 from Theorems 16.2-16.3. After collecting some facts about mutual information and distances between probability measures, we prove the conductance lower bound for reconstruction, Theorem 16.2, and the upper bound, Theorem 16.3. Extensions and unsolved problems are discussed at the end of the chapter.

Background

Let G be a finite graph with vertex set V. In the *ferromagnetic Ising model* with no external field on G, the *interaction strength* $J > 0$ and the *temperature* $t > 0$ determine a **Gibbs distribution** $\mathcal{G} = \mathcal{G}_{J,t}$ on $\{\pm 1\}^V$ which is defined by

$$\mathcal{G}(\sigma) = Z(t)^{-1} \exp(\sum_{u \sim v} J\sigma_u \sigma_v / t)\,, \tag{70}$$

where the normalizing factor $Z(t)$ is called the *partition function*. If the graph G is a tree, then this is equivalent to the Markovian propagation description in the beginning of the chapter, for an appropriate choice of the error parameter ϵ. Indeed, if $u \sim v$ are adjacent vertices in a finite tree with $\sigma_u = \sigma_v$, then flipping all the spins on one side of the edge connecting u and v will multiply the probability in (70) by $e^{-2J/t}$. Thus if we define ϵ by

$$\frac{\epsilon}{1-\epsilon} = e^{-2J/t}\,, \tag{71}$$

then the distributions defined by (64) and (70) coincide. For an infinite graph G, a weak limit point of the Gibbs distributions (70) on finite subgraphs $\{G_n\}$ exhausting G, (possibly with boundary conditions imposed on $\sigma_{\partial G_n}$), is called a (limiting) **Gibbs state** on G. See Georgii [30] for more complete definitions, using the notion of specification.

For any infinite graph with bounded degrees, the limiting Gibbs state is unique at sufficiently high temperatures, i.e., the limit from finite subgraphs exists and does not depend on boundary conditions. When $G = T$ is a tree, this means that

$$\lim_{n\to\infty} \mathbf{E}[\sigma_\rho \mid \sigma_{T_n} \equiv 1] = 0 \tag{72}$$

at high temperatures. Some graphs admit a phase transition: below a certain critical temperature, multiple Gibbs states appear and the limit in (72) is strictly positive. The critical temperature t_c^+ for this transition on a regular tree T was determined in 1974 by Preston [79]; his result was generalized in 1989 by Lyons [59] who showed that $\tanh(J/t_c^+) = \mathrm{br}(T)^{-1}$; in the equivalent Markovian description, the critical parameter ϵ_c^+ for an all + boundary to affect σ_ρ in the limit, satisfies $1 - 2\epsilon_c^+ = \mathrm{br}(T)^{-1}$.

In general, a Gibbs state is extremal (or "pure") iff it has a trivial tail, see Georgii ([30], Theorem 7.7). The tree-indexed Markov chain (64) on an infinite tree T is the limit of the Gibbs distributions (70) on finite subtrees, with no boundary conditions imposed; hence it is called the **free boundary Gibbs state** on T. In 1975 Spitzer

([82], Theorem 4) claimed that on a $b+1$ - regular tree $T^{(b)}$, the free boundary Gibbs states are extremal at any temperature. A counterexample, due to T. Kamae, was published in 1977 (see Higuchi [42]). Kamae showed that the sum of spins on $T_n^{(b)}$, normalized by its L^2 norm, converges to a non-constant tail-measurable function, provided that $1-2\epsilon > b^{-1/2}$. In 1978, this result was put in a broader context by Moore and Snell [69], who showed it followed from the 1966 results of Kesten and Stigum [51] on multi-type branching processes. Moore and Snell noted that it was open whether the free boundary Gibbs state on $T^{(b)}$ is extremal when $b^{-1} < 1-2\epsilon \le b^{-1/2}$. Chayes, Chayes, Sethna and Thouless [14] successfully analyzed a closely related spin-glass model on T_b; by a gauge transformation, this is equivalent to the Ising model with i.i.d. uniform $\{\pm 1\}$ boundary conditions. Although these boundary conditions are quite different from a free boundary, they turn out to have the same critical temperature. Bleher, Ruiz and Zagrebnov [12] adapted the recursive methods of Chayes et al [14] to the extremality problem, and showed that the free boundary Gibbs state on $T^{(b)}$ is extremal whenever $1-2\epsilon \le b^{-1/2}$. Shortly thereafter, a more streamlined argument was found by Ioffe [44]. Theorem 16.1 was first established in [24]. After learning of that result, Ioffe [45] found an elegant alternative proof for the upper bound.

Genetic reconstruction and parsimony

Tree-indexed Markov chains as in the introduction have been studied in the Mathematical Biology literature by Cavender [13], by Steel and Charleston [84], and others. In that literature the two "spins" are often called "colors", and correspond to traits of individuals, species, or DNA sequences. The "broadcasting errors" (color changes along edges) represent mutations, and one attempts to infer traits of ancestors from those of an observable population.

Proof of Theorem 16.1

(i) From $\gamma = 1-2\epsilon > \mathrm{br}(T)^{-1/2}$ it follows that

$$\mathcal{R}(\rho \leftrightarrow \infty) := \sup_n \mathcal{R}(\rho \leftrightarrow T_n) < \infty$$

when each edge e is assigned conductance $\gamma^{2|e|}$; see (39) and Theorem 2.8. Therefore by (66),

$$\inf_{n\ge 1} \Delta(T, T_n, \epsilon) \ge \inf_{n\ge 1} \frac{1}{1+\mathcal{R}(\rho \leftrightarrow T_n)} \ge \frac{1}{1+\mathcal{R}(\rho \leftrightarrow \infty)} > 0$$

and similarly $\inf_{n\ge 1} I(\sigma_\rho; \sigma_{T_n}) > 0$, as asserted. In particular, σ_ρ is not independent of the tail field of $\{\sigma_v\}$, so this tail field is not trivial.

(ii) If $\gamma = 1-2\epsilon < \mathrm{br}(T)^{-1/2}$ then $\inf_\Pi \sum_{v\in\Pi} \gamma^{2|v|} = 0$, so Theorem 16.3 implies that $\inf_{n\ge 1} \Delta(T, T_n, \epsilon) = 0$ and $\inf_{n\ge 1} I(\sigma_\rho; \sigma_{T_n}) = 0$.

Next, fix a finite set of vertices W_0. For each $w \in W_0$ and $n > |w|$, denote by $T_n(w)$ the set of vertices in T_n which connect to ρ via w. Then Lemma 16.4(iii) implies that for sufficiently large n,

$$I(\sigma_{W_0};\sigma_{T_n}) \le \sum_{w\in W_0} I(\sigma_{W_0},\sigma_{T_n(w)}) = \sum_{w\in W_0} I(\sigma_w,\sigma_{T_n(w)}), \tag{73}$$

since the conditional distribution of $\sigma_{T_n(w)}$ given σ_{W_0} is the same as its conditional distribution given σ_w.

For any finite W_0, the right-hand side of (73) tends to 0 as $n \to \infty$; It follows that the tail of $\{\sigma_v\}$ is trivial.

□

Mutual Information: Definition and Properties

Let X, Y be random variables defined on the same probability space which take finitely many values. The **entropy** of X is defined by

$$H(X) := -\sum_x \mathbf{P}[X=x]\log \mathbf{P}[X=x]$$

and the **mutual information** $I(X;Y)$ between X and Y is defined to be

$$I(X;Y) := H(X)+H(Y)-H(X,Y) = \sum_{x,y}\mathbf{P}[X=x,Y=y]\log\frac{\mathbf{P}[X=x,Y=y]}{\mathbf{P}[X=x]\mathbf{P}[Y=y]}.$$

We collect a few basic properties of mutual information in the following lemma. See, e.g., Cover and Thomas [15] §2.

Lemma 16.4 **(i)** *$I(X;Y)\ge 0$, with equality iff X and Y are independent;*

(ii) *Data processing inequality: If $X \mapsto Y \mapsto Z$ form a Markov chain (i.e., X and Z are conditionally independent given Y), then $I(X;Y) \ge I(X;Z)$.*

(iii) *Subadditivity: If $Y_1,\dots,Y_n$ are conditionally independent given X, then $I(X;(Y_1,\dots,Y_n)) \le \sum_{j=1}^n I(X;Y_j)$.*

The assumption of conditional independence in part (iii) cannot be omitted, as is shown by standard examples of 3 dependent random variables which are pairwise independent (e.g., Boolean variables satisfying $X = Y_1 + Y_2 \bmod 2$). Nevertheless, inequality (68) in Theorem 16.3 extends (iii) to a setting where this conditional independence need not hold.

Distances between probability measures

Let ν_+ and ν_- be two probability measures on the same space Ω. (In our application Ω is finite, but it is convenient to use notation that applies more generally.) Set $\nu := \frac{\nu_+ + \nu_-}{2}$ and denote $f = \frac{d\nu_+}{d\nu}$, $g = \frac{d\nu_-}{d\nu}$, so that $f+g \equiv 2$ identically. Suppose that ξ is uniform in $\{\pm 1\}$, and X has distribution ν_ξ. Inferring ξ from X is a basic problem of Bayesian hypothesis testing. (In our application, ξ will be the root spin σ_ρ, and X will be some function of the spin configuration σ_W on a finite vertex set W.)

There are several important notions of distance between ν_+ and ν_-, that can be related to this inference problem:

- **Total variation distance** $D_V(\nu_+,\nu_-) := \frac{1}{2}\int |f-g|\,d\nu$ can be interpreted as the difference between the probabilities of correct and erroneous inference. Indeed, among all functions $\widehat{\xi}$ of the observations, the probability of error $\mathbf{P}[\widehat{\xi} \neq \xi]$ is minimized by taking $\widehat{\xi} = 1$ if $f(X) \geq g(X)$, and $\widehat{\xi} = -1$ otherwise. We then have

$$\Delta := \mathbf{P}[\widehat{\xi} = \xi] - \mathbf{P}[\widehat{\xi} \neq \xi] = \frac{1}{2}\Big(\int \widehat{\xi} f\,d\nu - \int \widehat{\xi} g\,d\nu\Big) = \frac{1}{2}\int |f-g|\,d\nu\,. \quad (74)$$

- χ^2 **distance** $D_\chi(\nu_+,\nu_-) := \frac{1}{2}\{\int (f-g)^2\,d\nu\}^{1/2}$ represents the L^2 norm of the conditional expectation $\mathbf{E}(\xi\,|\,X) = \frac{1}{2}(f(X)-g(X))$.

- **Mutual information** between ξ and X,

$$D_I(\nu_+,\nu_-) := I(\xi;X) = \frac{1}{2}\int (f\log f + g\log g)\,d\nu \quad (75)$$

is a symmetrized version of the Kullback-Leibler divergence (see Vajda [86]).

- **The Hellinger distance**

$$D_H(\nu_+,\nu_-) := \int (\sqrt{f}-\sqrt{g})^2\,d\nu \; = 2\Big(1 - \int \sqrt{fg}\,d\nu\Big)\,. \quad (76)$$

derives its importance from the simple behavior of the Hellinger integrals

$$\mathtt{Int}_H(\nu_+,\nu_-) := \int \sqrt{fg}\,d\nu$$

for product measures:

$$\mathtt{Int}_H(\nu_+\times\mu_+\,,\,\nu_-\times\mu_-) = \mathtt{Int}_H(\nu_+,\nu_-)\mathtt{Int}_H(\mu_+,\mu_-)\,. \quad (77)$$

These distances appear in different sources under different names and with different normalizations. We collect here some well known inequalities between them, that will be useful below. For more on this topic, see, e.g., Le Cam [56] or Vajda [86].

Lemma 16.5 *With the notation above,*

(i) $D_\chi^2 \leq D_V \leq D_\chi \leq \sqrt{D_H}$

(ii) $D_\chi^2 \leq D_I \leq 2D_\chi^2$

(iii) *If* ν_+ *and* ν_- *are measures on* $\mathbb{R}$*, then*

$$\Big\{\int x\,d(\nu_+-\nu_-)\Big\}^2 = \Big\{\int x[f(x)-g(x)]\,d\nu\Big\}^2 \leq 4\int x^2\,d\nu\cdot D_\chi^2\,.$$

Proof.

(i) The left-hand inequality follows from $|f(x)-g(x)| \leq 2$, and the middle inequality from Cauchy-Schwarz. The right-hand inequality follows from the identity $f - g = (\sqrt{f}-\sqrt{g})\cdot(\sqrt{f}+\sqrt{g})$ and the concavity relation $\frac{\sqrt{f}+\sqrt{g}}{2} \leq \sqrt{\frac{f+g}{2}} = 1$.

(ii) Setting $\psi = (f-g)/2$, the assertion follows from the pointwise inequalities

$$\frac{\psi^2}{2} \leq \frac{1+\psi}{2}\log(1+\psi) + \frac{1-\psi}{2}\log(1-\psi) \leq \psi^2\,. \tag{78}$$

Here the left-hand inequality is verified for $\psi \in [0,1)$ by comparing second derivatives, and the right-hand inequality follows from $\log(1+y) \leq y$.

(iii) This is just the Cauchy-Schwarz inequality.

□

Finally, we interpret the data processing inequality in terms of distances. Suppose that we are given transition probabilities on the state space, *i.e.*, a stochastic matrix M (the entries of M are nonnegative and the row sums are all 1). Write $M^*\mu(y) := \sum_x M(x,y)\mu(x)$. Then Lemma 16.4 (ii) implies that

$$D_I(M^*\nu_+, M^*\nu_-) \leq D_I(\nu_+,\nu_-)\,.$$

An analogous inequality holds for total variation:

$$\begin{aligned} D_V(M^*\nu_+, M^*\nu_-) &= \frac{1}{2}\sum_y |M^*\nu_+(y) - M^*\nu_-(y)| \\ &\leq \frac{1}{2}\sum_y\sum_x M(x,y)|\nu_+(x)-\nu_-(x)| \\ &= \frac{1}{2}\sum_x |\nu_+(x)-\nu_-(x)| = D_V(\nu_+,\nu_-)\,. \end{aligned} \tag{79}$$

Conductance lower bounds: Proof of Theorem 16.2

Recall that each edge e was assigned the resistance

$$R(e) := (1-\gamma^2)\gamma^{-2|e|}\,. \tag{80}$$

Say that a set of vertices W is an **antichain** if no vertex in W is a descendant of another.

Lemma 16.6 *Let W be a finite antichain in T. For any unit flow μ from ρ to W, the weighted sum*

$$S_\mu := \sum_{v\in W}\frac{\mu(v)\sigma_v}{\gamma^{|v|}} \tag{81}$$

satisfies $\mathbf{E}[S_\mu \mid \sigma_\rho] = \sigma_\rho$ *and*

$$\mathbf{E}[S_\mu^2] = \mathbf{E}[S_\mu^2 \mid \sigma_\rho] = 1 + \sum_e R(e)\mu(e)^2 \,. \tag{82}$$

Consequently,

$$\min_\mu \mathbf{E}[S_\mu^2] = 1 + \mathcal{R}(\rho \leftrightarrow W)\,, \tag{83}$$

and the minimum is attained precisely when μ *is the unit current flow from* ρ *to* W*.*

Proof. From the product representation (64), we infer that

$$\mathbf{E}[\sigma_v \mid \sigma_\rho] = \sigma_\rho \gamma^{|v|}$$

for any vertex v. The formula for $\mathbf{E}[S_\mu \mid \sigma_\rho]$ follows by linearity. For any two vertices v, w in T, denote by $\mathtt{path}(v,w)$ the path from v to w. Also, write $\mathtt{path}(v)$ for $\mathtt{path}(\rho, v)$. Clearly,

$$\mathbf{E}[\sigma_v \sigma_w] = \gamma^{|\mathtt{path}(v,w)|} = \gamma^{|v|+|w|-2|v\wedge w|}\,, \tag{84}$$

where $v \wedge w$, the **meeting point** of v and w, is the vertex farthest from the root ρ on $\mathtt{path}(v) \cap \mathtt{path}(w)$. The percolation representation can also be invoked to justify (84).

It is now easy to determine the second moment of S_μ:

$$\mathbf{E}[S_\mu^2] = \sum_{v,w\in W} \frac{\mu(v)\mu(w)}{\gamma^{|v|}\gamma^{|w|}} \mathbf{E}[\sigma_v\sigma_w] = \sum_{v,w\in W} \frac{\mu(v)\mu(w)}{\gamma^{2|v\wedge w|}}\,, \tag{85}$$

Next, insert the identity

$$\gamma^{-2|u|} = 1 + \sum_{e\in\mathtt{path}(u)} R(e)$$

with $u = v\wedge w$, into (85). Changing the order of summation, and using the fact that W is an antichain, we obtain

$$\mathbf{E}[S_\mu^2] = 1 + \sum_e R(e) \sum_{v,w\in W} \mathbf{1}_{\{e\in\mathtt{path}(v\wedge w)\}}\mu(v)\mu(w)\,. \tag{86}$$

Since $\mathtt{path}(v\wedge w) = \mathtt{path}(v)\cap\mathtt{path}(w)$ and

$$\sum_{v,w\in W} \mathbf{1}_{\{e\in\mathtt{path}(v\wedge w)\}}\mu(v)\mu(w) = \Big(\sum_{v\in W}\mathbf{1}_{\{e\in\mathtt{path}(v)\}}\mu(v)\Big)\Big(\sum_{w\in W}\mathbf{1}_{\{e\in\mathtt{path}(w)\}}\mu(w)\Big) = \mu(e)^2,$$

(86) is equivalent to (82). Finally, (83) follows from Thomson's principle. $\square$

Proof of Theorem 16.2: We may assume that W is an antichain. (Otherwise, remove from W all vertices which have an ancestor in W.) Let μ be the unit current flow from ρ to W for the resistances $R(e)$ as in the preceding lemma, and let S_μ be the weighted sum (81). In order to apply Lemma 16.5, denote by ν_+ the conditional

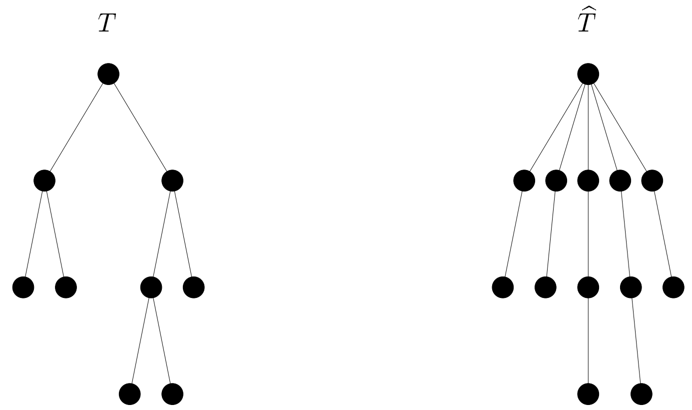


Figure 6: A tree T and the corresponding stringy tree $\widehat{T}$.

distribution of S_μ given that $\sigma_\rho = 1$; define ν_- analogously by conditioning that $\sigma_\rho = -1$, so that $\nu = (\nu_+ + \nu_-)/2$ is the unconditioned distribution of S_μ. We then have by Lemma 16.5(iii) that

$$D_\chi^2(\nu_+, \nu_-) \geq \frac{\left\{\int x d(\nu_+ - \nu_-)\right\}^2}{4\int x^2\, d\nu} = \frac{(\mathbf{E}[S_\mu \,|\, \sigma_\rho = 1] - \mathbf{E}[S_\mu \,|\, \sigma_\rho = -1])^2}{4\mathbf{E}[S_\mu^2]}.$$

Applying Lemma 16.6, we deduce that

$$D_\chi^2(\nu_+, \nu_-) \geq \frac{1}{1 + \mathcal{R}(\rho \leftrightarrow W)}. \tag{87}$$

By Lemma 16.5, the difference $\Delta = \Delta(T, W, \epsilon)$ between the probabilities of correct and incorrect reconstruction, satisfies $\Delta = D_V(\nu_+, \nu_-) \geq D_\chi^2(\nu_+, \nu_-)$, and the mutual information between σ_ρ and σ_W also satisfies $I(\sigma_\rho; \sigma_W) = D_I(\nu_+, \nu_-) \geq D_\chi^2(\nu_+, \nu_-)$. In conjunction with (87), this completes the proof. □

Mincut upper bound: Proof of Theorem 16.3

Definition. A **noisy tree** is a tree with flip probabilities labeling the edges. The *stringy tree* $\widehat{T}$ associated with a finite noisy tree T is the tree which has the same set of root-leaf paths as T but in which these paths act as independent channels. More precisely, for every root-leaf path in T, there exists an identical (in terms of length and flip probabilities on the edges) root-leaf path in $\widehat{T}$, and in addition, all the root-leaf paths in $\widehat{T}$ are edge-disjoint.

Theorem 16.7 *Given a finite noisy tree T with leaves W, let $\widehat{T}$, with leaves $\widehat{W}$ and root $\hat{\rho}$, be the stringy tree associated with T. There is a channel which, for $\xi \in \{\pm 1\}$, transforms the conditional distribution $\sigma_{\widehat{W}} \,|\, (\sigma_{\hat{\rho}} = \xi)$ into the conditional distribution $\sigma_W \,|\, (\sigma_\rho = \xi)$. Equivalently, we say that $\widehat{T}$ dominates T.*

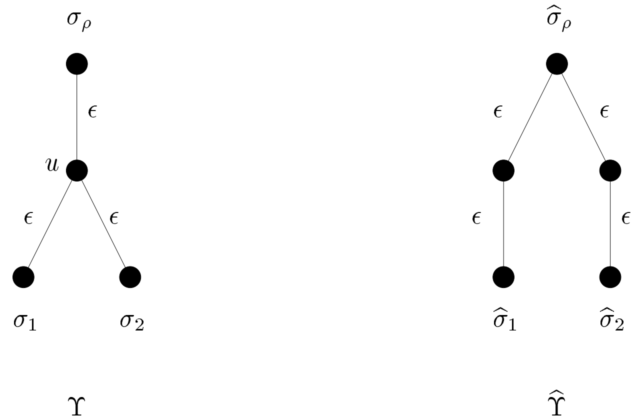


Figure 7: Υ is dominated by $\widehat{\Upsilon}$.

Remark A *channel* is formally defined as a stochastic matrix describing the conditional distribution $\mathbf{P}(Y \mid X)$ of the output variable Y given the input X, see [15]. Often a channel is realized by a relation of the form $Y = f(X, Z)$, where f is a deterministic function and Z is a random variable (representing the "noise") which is independent of X.

Proof: We only establish a key special case of the theorem: namely, that the tree Υ shown in Figure 7, is dominated by the corresponding stringy tree $\widehat{\Upsilon}$. The general case is derived from it by first allowing the flip probabilities to vary from edge to edge, and then applying an inductive argument; see [25] for details.

Given $0 \le \alpha \le 1$, to be specified below, we define the channel as follows:

$$
\begin{aligned}
\sigma_1^* &= \widehat{\sigma}_1 \\
\sigma_2^* &= \begin{cases} \widehat{\sigma}_2 & \text{with probability } \alpha \\ \widehat{\sigma}_1 & \text{with probability } 1-\alpha \end{cases}
\end{aligned}
$$

To prove that $(\widehat{\sigma}_\rho, \sigma_1^*, \sigma_2^*)$ has the same distribution as $(\sigma_\rho, \sigma_1, \sigma_2)$, it suffices to show that the means of corresponding products are equal. (This is a special case of the fact that the characters on any finite Abelian group G form a basis for the vector space of complex functions on G.) By symmetry

$$\mathbf{E}(\sigma_\rho) = \mathbf{E}(\sigma_1) = \mathbf{E}(\sigma_2) = \mathbf{E}(\sigma_\rho\sigma_1\sigma_2) = \mathbf{E}(\widehat{\sigma}_\rho) = \mathbf{E}(\sigma_1^*) = \mathbf{E}(\sigma_2^*) = \mathbf{E}(\widehat{\sigma}_\rho\sigma_1^*\sigma_2^*) = 0$$

and thus we only need to check pair correlations. Clearly, $\mathbf{E}(\widehat{\sigma}_\rho\sigma_1^*) = \mathbf{E}(\sigma_\rho\sigma_1)$ and $\mathbf{E}(\widehat{\sigma}_\rho\widehat{\sigma}_1) = \gamma^2$, whence $\mathbf{E}(\widehat{\sigma}_\rho\sigma_2^*) = \gamma^2 = \mathbf{E}(\sigma_\rho\sigma_2)$ for any choice of α. Finally, since $\mathbf{E}(\sigma_1^*\widehat{\sigma}_2) = \gamma^4 < \gamma^2 = \mathbf{E}(\sigma_1\sigma_2)$ and

$$\mathbf{E}(\sigma_1^*\widehat{\sigma}_1) = 1 > \gamma^2 \,,$$

we can choose $\alpha \in [0,1]$ so that $\mathbf{E}(\sigma_1^*\sigma_2^*) = \mathbf{E}(\sigma_1\sigma_2)$; explicitly,

$$\alpha = (1-\gamma^2)/(1-\gamma^4)\,. \tag{88}$$

This proves that $\widehat{\Upsilon}$ dominates Υ. $\square$

Proof of Theorem 16.3: We first prove (68). Since W_1 separates ρ from W, the data processing inequality (Lemma 16.4 (ii)) yields $I(\sigma_\rho; \sigma_W) \le I(\sigma_\rho; \sigma_{W_1})$. Let T_1 be the tree obtained from T by retaining only W_1 and ancestors of nodes in W_1. Let $\widehat{T_1}$ be the stringy tree associated with T_1. From Theorem 16.7 applied to T_1 and the data processing inequality, we obtain $I(\sigma_\rho; \sigma_{W_1}) \le I(\sigma_{\hat\rho}; \sigma_{\widehat{W}_1})$. Since the spins on leaves of $\widehat{T_1}$ are conditionally independent given $\sigma_{\hat\rho}$, subadditivity (Lemma 16.4 (iii)) gives

$$I(\sigma_{\hat\rho}; \sigma_{\widehat{W}_1}) \le \sum_{\hat v \in \widehat{W}_1} I(\sigma_{\hat\rho}; \sigma_{\hat v}) \,.$$

But due to the definition of the stringy tree, the mutual information between $\sigma_{\hat\rho}$ and $\sigma_{\hat v}$ is identical to the mutual information between σ_ρ and σ_v in T_1, hence the left inequality in (68).

Since $\mathbf{E}(\sigma_\rho \sigma_v) = \gamma^{|v|}$ for each v, the right-hand inequality in (68) follows from the right-hand inequality in (78).

We now turn to the total variation inequality (67). Recall that $\Delta(T, W, \epsilon)$, the difference between the probabilities of correct and incorrect reconstruction, equals $D_V(\nu_+^W, \nu_-^W)$, the total variation distance between the two distributions of the spins on W given $\sigma_\rho = \pm 1$.

By (79), Theorem 16.7, and Lemma 16.5,

$$D_V(\nu_+^W, \nu_-^W) \le D_V(\nu_+^{W_1}, \nu_-^{W_1}) \le D_V(\nu_+^{\widehat{W}_1}, \nu_-^{\widehat{W}_1}) \le \sqrt{D_H(\nu_+^{\widehat{W}_1}, \nu_-^{\widehat{W}_1})} \,.$$

Now, $D_H(\nu_+^{\widehat{W}_1}, \nu_-^{\widehat{W}_1})$ on the stringy tree $\widehat{T_1}$ is easily calculated using the multiplicative property of Hellinger integrals: $\nu_+^{\widehat{W}_1}$ is just the product over $w \in \widehat{W}_1$ of ν_+^w, the distribution of σ_w given $\sigma_\rho = 1$, and similarly $\nu_-^{\widehat{W}_1} = \prod_w \nu_-^w$. Since $\mathtt{Int}_H(\nu_+^w, \nu_-^w) = \sqrt{1 - \gamma^{2|w|}}$, the left-hand inequality in (67) follows; the right-hand inequality there is a consequence of the standard inequality $\prod(1 - x_j) \ge 1 - \sum x_j$. $\square$

Remarks and unsolved problems

1. **Reconstruction at criticality.** It is shown in [12, 44] that on infinite regular trees, $\lim_n \Delta(T, T_n, \epsilon_c) = 0$. On general trees, Theorem 16.2 implies that finite effective resistance from the root to infinity (when each edge at level ℓ is assigned the resistance $(1-2\epsilon)^{-2\ell}$) is sufficient for $\lim_n \Delta(T, T_n, \epsilon) > 0$. In [75], a recursive method is used to show this condition is also necessary.

2. **Multi-colored trees and the Potts model.** The most natural generalization of the two-state tree-indexed Markov chain model studied in this chapter involves multicolored trees, where the coloring propagates according to any finite state tree-indexed Markov chain. For instance, if this Markov chain is defined by a $q \times q$ stochastic matrix where all entries off the main diagonal equal ϵ, then

the q-state Potts model arises. The proof of Theorem 16.2 extends to general Markov chains, and shows that the tail of the tree-indexed chain is nontrivial if $\mathrm{br}(T) > \lambda_2^{-2}$, where λ_2 is the second eigenvalue of the transition matrix (*e.g.* for the q-state Potts model, $\lambda_2 = 1 - q\epsilon$). However, unpublished calculations of E. Mossel indicate that this lower bound is not sharp in general. Furthermore, we do not know a reasonable upper bound on mutual information between root and boundary variables. In particular, it seems that the critical parameter for tail triviality in the Potts model on a regular tree is not known.

3. **An information inequality.** Theorem 16.3 implies that the spins in the ferromagnetic Ising model on a tree satisfy

$$I(\sigma_v;\sigma_W) \leq \sum_{w\in W} I(\sigma_v;\sigma_w)\,,$$

for any vertex v and any finite set of vertices W. Does this inequality hold on other graphs as well?

More generally, are there natural assumptions (*e.g.*, positive association) on random variables $X, Y_1, \ldots Y_n$ that imply the inequality $I(X;(Y_1,\ldots,Y_n)) \leq \sum_{j=1}^n I(X,Y_j)$?

17 Unpredictable Paths in Z and EIT in $\mathbf{Z}^3$

The goal of this chapter is to complete the proof of Theorem 11.1, by exhibiting a probability measure on directed paths in $\mathbf{Z}^3$ that has exponential intersection tails. We construct the required measure in three dimensions from certain nearest-neighbor stochastic processes on $\mathbf{Z}$ that are "less predictable than simple random walk".

For a sequence of random variables $S = \{S_n\}_{n\geq 0}$ taking values in a countable set V, we define its **predictability profile** $\{\mathsf{PRE}_S(k)\}_{k\geq 1}$ by

$$\mathsf{PRE}_S(k) = \sup \mathbf{P}[S_{n+k} = x \mid S_0,\ldots,S_n]\,, \tag{89}$$

where the supremum is over all $x \in V$, all $n \geq 0$, and all histories $S_0,\ldots,S_n$.

Thus $\mathsf{PRE}_S(k)$ is the maximal chance of guessing S correctly k steps into the future, given the past of S. Clearly, the predictability profile of simple random walk on $\mathbf{Z}$ is asymptotic to $Ck^{-1/2}$ for some $C > 0$.

Theorem 17.1 (Benjamini, Pemantle, and Peres 1998) *For any $\alpha < 1$ there exists an integer-valued stochastic process $\{S_n\}_{n\geq 0}$ such that $|S_n - S_{n-1}| = 1$ a.s. for all $n \geq 1$ and*

$$\mathsf{PRE}_S(k) \leq C_\alpha k^{-\alpha} \quad \text{for some } C_\alpha < \infty, \text{ for all } k \geq 1\,. \tag{90}$$

After Theorem 17.1 was proven in BPP (1998), Häggström and Mossel (1998) constructed processes with lower predictability profile. They showed that if f is nondecreasing and $\sum_k (f(k)k)^{-1} < \infty$, then there is a nearest–neighbor process S on $\mathbf{Z}$

with $\mathsf{PRE}_S(k) \le Cf(k)k^{-1}$. (For example, $f(k) = \log^{1+\epsilon}(k)$ satisfies this summability condition.)

Hoffman (1998) proved that this result is sharp: if a nondecreasing function f satisfies $\sum_k (f(k)k)^{-1} = \infty$, then there is no nearest–neighbor process on **Z** with predictability profile bounded by $O(f(k)k^{-1})$.

We prove Theorem 17.1 using the Ising model on a tree. We follow Häggström and Mossel (1998), who improved the original argument from BPP (1998). The following lemma is the engine behind the proof. Let T be the b–adic tree of depth N, and fix $0 < \epsilon < 1/4$. We will assign to the vertices of T ± 1 labels $\{\sigma(v)\}_{v \in T}$ according to an Ising model (see Chapter 16). For the root ρ, set $\sigma(\rho) = 1$, and for a vertex w with parent v, let

$$\sigma(w) = \begin{cases} \sigma(v) & \text{with probability } 1-\epsilon \\ -\sigma(v) & \text{with probability } \epsilon \end{cases} .$$

Lemma 17.2 *Denote by $Y_N := \sum_{v \in T_N} \sigma(v)$ the sum of the spins at level N. There exists $C_b < \infty$ such that for all $N \ge 1$ and all $x \in \mathbf{Z}$,*

$$\mathbf{P}[Y_N = x] \le \frac{C_b}{\epsilon[b(1-2\epsilon)]^N} .$$

Proof. By decomposing the sum Y_{M+1} into b parts corresponding to the subtrees of depth M rooted at the first level, we get

$$Y_{M+1} = \sum_{j=1}^{b} \sigma(v_j) Y_M^{(j)} ,$$

where $\{\sigma(v_j)\}_{j=1}^b$ are b i.i.d. spins with

$$\sigma(v_j) = \begin{cases} +1 & \text{with probability } 1-\epsilon \\ -1 & \text{with probability } \epsilon \end{cases} ,$$

and $\{Y_M^{(j)}\}_{j=1}^b$ are i.i.d. variables with the distribution of Y_M, independent of these spins. Consequently, the characteristic functions

$$\widehat{Y}_M(\lambda) = \mathbf{E}(e^{i\lambda Y_M})$$

satisfy the recursion

$$\begin{aligned} \widehat{Y}_{M+1}(\lambda) &= ((1-\epsilon)\widehat{Y}_M(\lambda) + \epsilon \widehat{Y}_M(-\lambda))^b \\ &= (\Re \widehat{Y}_M(\lambda) + i(1-2\epsilon)\Im \widehat{Y}_M(\lambda))^b \end{aligned} \tag{91}$$

where $\Re$ denotes real part, and $\Im$ imaginary part. For $\theta_n(\lambda) := \arg \widehat{Y}_n(\lambda)$, define

$$J_n := \Big\{ 0 \le \lambda \le \frac{\pi}{2} : \theta_k(\lambda) < \frac{\pi}{2b}, \ k = 0, \cdots, n-1 \Big\}$$

and

$$I_n := J_n \setminus J_{n+1} .$$

We will evaluate the integral of $\widehat{Y}_N(\lambda)$ over $(0,\pi/2]$ by using the decomposition

$$[0,\frac{\pi}{2}] = \Big(\bigcup_{k=0}^{N-1} I_k\Big)\bigcup J_N\,.$$

Rewrite (91) as

$$\widehat{Y}_{M+1}(\lambda) = |\widehat{Y}_M(\lambda)|^b[\cos\theta_M(\lambda) + i(1-2\epsilon)\sin\theta_M(\lambda)]^b, \tag{92}$$

and infer, for $0\le\theta_M(\lambda)\le\frac{\pi}{2b}$, that

$$\theta_{M+1}(\lambda) = b\arctan\Big((1-2\epsilon)\tan\theta_M(\lambda)\Big).$$

Since arctan is concave in $[0,\infty)$ and $\arctan 0 = 0$,

$$\arctan\Big((1-2\epsilon)\alpha\Big) \ge (1-2\epsilon)\arctan(\alpha)$$

for any $\alpha\ge 0$. Therefore

$$\text{If }\ 0\le\theta_M(\lambda)\le\frac{\pi}{2b}\,,\ \text{then }\ \frac{\pi}{2}\ge b\theta_M(\lambda)\ge\theta_{M+1}(\lambda)\ge b(1-2\epsilon)\theta_M(\lambda). \tag{93}$$

If $\lambda\in I_n$, then applying (93) for $M=n-1$ shows that

$$\frac{\pi}{2}\ge\theta_n(\lambda)\ge\frac{\pi}{2b}\,. \tag{94}$$

Using (92) with $M=n$ together with (94), we find that for $\lambda\in I_n$,

$$|\widehat{Y}_{n+1}(\lambda)| \le \Big(\cos^2(\frac{\pi}{2b}) + (1-2\epsilon)^2\sin^2(\frac{\pi}{2b})\Big)^{\frac{b}{2}} \le \Big(1-2\epsilon\sin^2(\frac{\pi}{2b})\Big)^{\frac{b}{2}} \le e^{-\varrho\epsilon b}\,, \tag{95}$$

where $\varrho := \sin^2(\frac{\pi}{2b})$. Inductive use of (92) for $\lambda\in I_n$ and $N>n$ gives

$$|\widehat{Y}_N(\lambda)| \le e^{-\varrho\epsilon b^{N-n}}. \tag{96}$$

Since $\theta_0(\lambda)\equiv\lambda$, (93) implies that $\theta_k(\lambda)\ge b^k(1-2\epsilon)^k|\lambda|$ for $\lambda\in J_n$ and $k\le n$. Therefore

$$|I_n|\le|J_n|\le\frac{\pi}{2b^n(1-2\epsilon)^n}, \tag{97}$$

By (96),

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|\widehat{Y}_N(\lambda)|\,d\lambda = \frac{2}{\pi}\int_0^{\frac{\pi}{2}}|\widehat{Y}_N(\lambda)|\,d\lambda \le \frac{2}{\pi}\Big(\sum_{k=0}^{N-1}|I_k|e^{-\varrho\epsilon b^{N-k}} + |J_N|\Big).$$

Inserting (97) yields

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}|\widehat{Y}_N(\lambda)|\,d\lambda \le \frac{1}{(1-2\epsilon)^N}\Big(\sum_{k=0}^{N-1}b^{-k}e^{-\varrho\epsilon b^{N-k}} + b^{-N}\Big). \tag{98}$$

In order to evaluate the sum in the right hand side of (98), we define

$$r = \max\{k \;:\; \varrho\epsilon b^{N-k} > 1\}\,.$$

Separating the contributions of $k \geq r$ and $k < r$, we obtain that

$$\sum_{k=r}^{N-1} b^{-k} e^{-\varrho\epsilon b^{N-k}} + b^{-N} \leq \sum_{k=r}^{N} b^{-k} \leq b^{-r} \sum_{k=0}^{\infty} b^{-k} \tag{99}$$

and

$$\sum_{k=0}^{r-1} b^{-k} e^{-\varrho\epsilon b^{N-k}} \leq \sum_{k=0}^{r-1} b^{-k} e^{-b^{r-k}} \leq b^{-r} \sum_{k=0}^{\infty} b^{k} e^{-b^{k}}. \tag{100}$$

Furthermore, since $\varrho\epsilon b^{N-r-1} \leq 1$, we have that

$$b^{-r} \leq \frac{1}{\varrho\epsilon b^{N-1}}\,. \tag{101}$$

Combining (98), (99), (100), and (101) we see that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} |\widehat{Y}_N(\lambda)|\, d\lambda \leq \frac{C_b}{b^N \epsilon (1-2\epsilon)^N}\,,$$

where

$$C_b = \frac{b(\sum_{k=0}^{\infty} b^{-k} + \sum_{k=0}^{\infty} b^{k} e^{-b^{k}})}{\varrho}$$

and ϱ was defined after (95). Using the inversion formula we conclude that

$$\mathbf{P}[Y_N = x] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \widehat{Y}_N(\lambda) e^{-i\lambda x}\, d\lambda \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} |\widehat{Y}_N(\lambda)|\, d\lambda \leq \frac{C_b}{b^N \epsilon (1-2\epsilon)^N}.$$

□

Proof of Theorem 17.1. For all $N > 0$, we will define a process S up to time $M = 2^N$ with the required properties. A process defined for all times will then exist by consistency of the finite dimensional distributions.

Fix a small $\epsilon > 0$. We assign spins $\{\sigma_v\}$ to the vertices of the binary tree T of depth N, according to the Ising model (described before Lemma 17.2) with error rate ϵ, but we take σ_ρ to be random uniform in $\{\pm 1\}$, rather than fixing it. Enumerate the vertices at depth N from left to right as $v_0, v_1, \ldots, v_M$, and set

$$S_n = \sum_{k=1}^{n} \sigma(v_k).$$

We claim that $\{S_n\}$ has the desired predictability profile. To see this, fix $0 \leq n < M$ and $0 < k \leq M - n$. Observe that $S_{n+k} = S_n + \sum_{j=n+1}^{n+k} \sigma(v_j)$. If we now take the unique h satisfying $2^{h+1} \leq k < 2^{h+2}$, there will exist a vertex w at level $N-h$ for which all of the descendants at depth N are in the set $\{v_{n+1}, \ldots, v_{n+k}\}$. It follows (by conditioning on the spins of all v_i which are not descendants of w and on the spin of w) that

$$\sup_{x\in\mathbf{Z}} \mathbf{P}[S_{n+k}=x|S_0,\ldots,S_n] \le \sup_{x\in\mathbf{Z}} \mathbf{P}[Y_h=x]. \tag{102}$$

Applying Lemma 17.2 and (102) we get

$$\mathsf{PRE}_S(k) \le \frac{C_b}{2^h\epsilon(1-2\epsilon)^h}, \tag{103}$$

and the proof is complete. □

The process S serves as a building block for $\mathbf{Z}^d$-valued processes whose predictability profiles are controlled.

Corollary 17.3 *For each $\frac{1}{2}<\alpha<1$, there is a $\mathbf{Z}^d$–valued process $\Phi=\Phi^{\alpha,d}$ such that the random edge sequence $\{\Phi_{n-1}\Phi_n\}_{n\ge1}$ is in Υ_1, and*

$$\forall k\ge1 \quad \mathsf{PRE}_\Phi(k) \le C(\alpha,d)k^{-(d-1)\alpha}. \tag{104}$$

Proof. Let $W_k^r=(S_k^{(r)}+k)/2$ for $r=1,\ldots,d-1$, where $S^{(r)}$ are independent copies of the process described in Theorem 17.1. For $r=1,\ldots,d-1$, define clocks

$$t_r(n) := \lfloor \frac{n+d-1-r}{d-1} \rfloor,$$

and let $D(n):=n-\sum_{r=1}^{d-1}W^r_{t_r(n)}$.

Write $\Phi_n=(W^1_{t_1(n)},\ldots,W^{d-1}_{t_{d-1}(n)},D(n))$. It is then easy to see that

$$\mathsf{PRE}_\Phi(k) \le \left[\mathsf{PRE}_S(\lfloor\frac{k}{d-1}\rfloor)\right]^{d-1} \le \left(\frac{C_\alpha k}{d-1}\right)^{-\alpha(d-1)} \le C(\alpha,d)k^{-\alpha(d-1)}.$$

□

The last ingredient we need to prove that $\mathbf{Z}^3$ admits paths with exponential intersection tails is the following.

Lemma 17.4 *Let $\{\Gamma_n\}$ be a sequence of random variables taking values in a countable set V. If the predictability profile (defined in (89)) of Γ satisfies $\sum_{k=1}^\infty \mathsf{PRE}_\Gamma(k)<\infty$, then there exist $C<\infty$ and $0<\theta<1$ such that for any sequence $\{v_n\}_{n\ge0}$ in V and all $\ell\ge1$,*

$$\mathbf{P}[\#\{n\ge0 \,:\, \Gamma_n=v_n\}\ge\ell] \le C\theta^\ell. \tag{105}$$

Proof. Choose m large enough so that $\sum_{k=1}^\infty \mathsf{PRE}_\Gamma(km)=\beta<1$, whence for any sequence $\{v_n\}_{n\ge0}$,

$$\mathbf{P}\Big[\exists k\ge1 \,:\, \Gamma_{n+km}=v_{n+km}\,\Big|\,\Gamma_0,\ldots,\Gamma_n\Big] \le \beta \quad \text{for all } n\ge0. \tag{106}$$

If n is replaced by a stopping time τ and the σ–field generated by $\Gamma_0,\ldots,\Gamma_n$ is replaced by the usual stopping time σ-field, then (106) remains valid. This can be seen by decomposing the probability according to the value of τ, and checking that the bound holds in each case. Hence, it follows by induction on $r \geq 1$ that for all $j \in \{0, 1, \ldots, m-1\}$,

$$\mathbf{P}[\#\{k \geq 1 \, : \, \Gamma_{j+km} = v_{j+km}\} \geq r] \leq \beta^r \, . \tag{107}$$

If $\#\{n \geq 0 \, : \, \Gamma_n = v_n\} \geq \ell$ then there must be some $j \in \{0, 1, \ldots, m-1\}$ such that

$$\#\{k \geq 1 \, : \, \Gamma_{j+km} = v_{j+km}\} \geq \ell/m - 1 \, .$$

Thus the inequality (105), with $\theta = \beta^{1/m}$ and $C = m\beta^{-1}$, follows from (107). $\square$

Proof of Theorem 11.1 for $d = 3$: The process Φ constructed in Corollary 17.3 with $\alpha > 1/2$ and $d = 3$ satisfies $\sum_k \mathsf{PRE}_\Phi(k) < \infty$, and hence by Lemma 17.4, the distribution μ of the edge sequence $\{\Phi_{n-1}\Phi_n\}_{n=1}^\infty$ has exponential intersection tails. $\square$

18 Tree-Indexed Processes

Label the vertices of a tree Γ by a collection of i.i.d. real random variables $\{X_v\}_{v\in\Gamma}$. Given Γ and the collection $\{X_v\}_{v\in\Gamma}$, we define the **tree-indexed random walk** $\{S_v\}_{v\in\Gamma}$ by

$$S_v = \sum_{w \leq v} X_w \, ,$$

where $w \leq v$ means that v is a descendant of w.

The simple case where Γ is a binary tree and $X_v = \pm 1$ with probabilities p and $1-p$ was considered by Dubins and Freedman (1967).

We want to determine the speed of tree-indexed random walks, or at least recognize when the speed is positive.

There are several ways to define speed for tree-indexed walks and the answers depend on the definition used. Here are three notions of speed.

Definitions of Speed

- **Cloud Speed**

$$s_{\text{cloud}} := \overline{\lim_n} \, \frac{1}{n} \max_{|v|=n} S_v \, ;$$

- **Burst Speed**

$$s_{\text{burst}} := \sup_{\xi\in\partial\Gamma} \, \overline{\lim_{v\in\xi}} \, \frac{S_v}{|v|} \, ;$$

- **Sustainable Speed**

$$s_{\text{sust}} := \sup_{\xi\in\partial\Gamma} \, \underline{\lim_{v\in\xi}} \, \frac{S_v}{|v|} \, ;$$

These speeds are a.s. constant by Kolmogorov's zero-one law. The first two were studied by Benjamini and Peres (1994b), while the third was studied earlier by Lyons and Pemantle (1992).

Assumptions. Throughout this chapter we will assume that each variable

$$X_v \text{ is not a.s. constant, } \mathbf{E}[X_v] = 0 \text{ and } \mathbf{E}[e^{\lambda X_v}] < \infty \text{ for all } \lambda > 0\,. \qquad (108)$$

These assumptions can be relaxed, but they make the ideas of the proofs more transparent.

In general, $s_{\text{cloud}} \geq s_{\text{burst}} \geq s_{\text{sust}}$. The following examples shows that the inequalities may be strict.

Example 18.1 Consider the 3-1 tree Γ in Example 2.6. It follows from Theorem 18.4 below that on this tree

$$s_{\text{cloud}} > 0 \quad \text{but} \quad s_{\text{burst}} = s_{\text{sust}} = 0\,.$$

Example 18.2 Let $n_1 < n_2 < \ldots$ be a sequence of positive integers. Construct a tree Γ as follows: The first n_1 levels of Γ are as in the 3-1 tree. To each vertex v in the n_1-th level of Γ, attach a copy of the first $n_2 - n_1$ levels of the 3–1 tree, with v as its root. Continue by attaching a copy of the first $n_{k+1} - n_k$ levels of the 3-1 tree to each vertex at level n_k of Γ. For any choice $\{n_i\}$, the tree Γ has positive packing dimension; in particular, $\dim_M(\partial\Gamma) = \dim_P(\partial\Gamma) = \log 2$. However, if the n_i increase sufficiently fast, then the Hausdorff dimension of $\partial\Gamma$ is 0, as in the 3-1 tree. Thus in this case Theorem 18.4 yields that $s_{\text{cloud}} \geq s_{\text{burst}} > 0$, but $s_{\text{sust}} = 0$.

Notation. Denote by $\{\tilde{S}_n\}_{n\geq 0}$ the ordinary random walk indexed by the non-negative integers with i.i.d. increments distributed like X_v. Let $I(\cdot)$ be the *rate function* for the random walk $\{\tilde{S}_n\}$, defined by

$$I(a) = \lim_{n\to\infty} \frac{1}{n} \log \mathbf{P}(\tilde{S}_n > na) \quad (a > 0)\,.$$

Theorem 18.3 (Hammersley (1974), Kingman (1975), Biggins (1977)) *Let Γ be a GW tree with mean $m > 1$. Suppose that the vertices of Γ are labeled by random variables X_v that satisfy (108). On the event that Γ survives, a.s. all speeds coincide and equal $s^* := \sup\{s : I(s) \leq \log m\}$.*

Proof. The inequality $s_{\text{cloud}} \leq s^*$ is easy: By the definition of s^*, for any $\epsilon > 0$ there is $\delta > 0$ such that $I(s^* + \epsilon) > \log m + \delta$. Therefore,

$$\mathbf{P}(\tilde{S}_n > n(s^* + \epsilon)) \leq e^{-n(\log m + \delta)} = m^{-n} e^{-n\delta}\,.$$

Consequently,

$$\mathbf{P}(S_v > n(s^* + \epsilon) \text{ for some } v \in \Gamma_n \mid \text{non-extinction }) \leq \frac{m^n}{1-q} m^{-n} e^{-n\delta}\,,$$

where q is the probability of extinction. The proof is concluded by invoking the Borel-Cantelli Lemma.

For the reverse inequality, let $a < s^*$ be given. Using the strict monotonicity of the rate function and the definition of s^*, choose ϵ so that $I(a) + 2\epsilon < \log m$. For each $k \geq 1$ and $M \in [1, \infty]$, we define a new embedded branching process as follows: start from the root of Γ, and take the set of offspring $\Gamma(v, k, M)$ of a vertex v to consist of all its descendants w in Γ that satisfy

- $|w| = |v| + k$ in Γ ;
- $S_w > S_v + ka$.
- $S_u > S_v - M$ for all u on the path from v to w.

(Here $M = \infty$ means the last requirement holds automatically.) Since $\mathbf{E}|\Gamma(v, k, \infty)| = m^k \mathbf{P}[\tilde{S}_k > ka]$, the definition of I yields that for sufficiently large k,

$$\mathbf{E}|\Gamma(v, k, \infty)| \geq m^k e^{-k[I(a)+\epsilon]} > 2 .$$

By choosing M large, we can ensure that the embedded process has mean offspring

$$\mathbf{E}|\Gamma(v, k, M)| \geq \frac{1}{2} m^k e^{-k[I(a)+\epsilon]} > 1 .$$

Thus for large k, M, the embedded process is supercritical. Therefore $s_{\text{sust}} > a$ with positive probability. Since

$$\{\Gamma \colon \Gamma \text{ finite or } s_{\text{sust}} \leq a \text{ on } \Gamma\}$$

is an inherited property, Proposition 3.2 implies that $\mathbf{P}[s_{\text{sust}} > a \mid \text{survival}\,] = 1$. Hence, given survival, we have that a.s.,

$$s^* \geq s_{\text{cloud}} \geq s_{\text{burst}} \geq s_{\text{sust}} \geq s^* . \qquad \square$$

We have already encountered two of the following definitions:

- The *upper Minkowski* dimension of $\partial\Gamma$, written $\overline{\dim}_M(\partial\Gamma)$, is $\log \overline{\mathrm{gr}}(\Gamma)$.
- The *Hausdorff* dimension of $\partial\Gamma$, written $\dim_H(\partial\Gamma)$, is $\log \mathrm{br}(\Gamma)$.
- The **Packing** dimension of $\partial\Gamma$, is defined by

  $$\dim_P(\partial\Gamma) \ := \ \inf \{ \sup_i \ \dim_M(\partial\Gamma^{(i)}) \},$$

  where the infimum extends over all countable collections $\{\Gamma^{(i)}\}$ of subtrees of Γ such that $\partial\Gamma \subseteq \bigcup_i \partial\Gamma^{(i)}$.

Theorem 18.4 *Suppose that Γ is an infinite tree without leaves, and the vertices of Γ are labeled by random variables X_v that satisfy (108). Then*

(i) $s_{\text{cloud}} > 0 \Leftrightarrow \overline{\dim}_M(\partial\Gamma) > 0$.

(ii) $s_{\text{burst}} > 0 \Leftrightarrow \dim_P(\partial\Gamma) > 0$.

(iii) $s_{\text{sust}} > 0 \Leftrightarrow \dim_H(\partial\Gamma) > 0$.

Proof. (i) The implication "$\Rightarrow$" is easy: By Cramér's theorem on large deviations, (108) implies that $I(a) > 0$ for any $a > 0$. Therefore

$$\sum_n \mathbf{P}(S_v > na \text{ for some } v \in \Gamma_n) \le \sum_n |\Gamma_n|\, \mathbf{P}(\tilde{S}_n > an) \le \sum_n |\Gamma_n|\, e^{-nI(a)}\,.$$

which is finite since $\overline{\dim}_M(\partial\Gamma) = 0$ means that Γ has subexponential growth. Thus by Borel-Cantelli

$$\mathbf{P}(\{S_v > na \text{ for some } v \in \Gamma_n\} \text{ i.o. }) \;=\; 0$$

for any $a > 0$.

For the implication "$\Leftarrow$", observe that because we assumed Γ has no leaves, there exists at least one descendant in Γ_{2n} for each $v \in \Gamma_n$. Denote the leftmost such descendant by $w(v)$. The $|\Gamma_n|$ paths from vertices $v \in \Gamma_n$ to the corresponding $w(v)$ are disjoint. Since $\overline{\dim}_M(\partial\Gamma) > 0$, if we choose ϵ sufficiently small, then

$$|\Gamma_n| > e^{n[I(2\epsilon)+2\epsilon]} \quad \text{for infinitely many } n \tag{109}$$

By Cramér's theorem, $\mathbf{P}(\tilde{S}_n > 2n\epsilon) > e^{-n[I(2\epsilon)+\epsilon]}$ for large n.

Write $\Gamma_n' = \{v \in \Gamma_n \colon S_v > -n\epsilon\}$. By the Weak Law of Large Numbers,

$$|\Gamma_n|^{-1}\mathbf{E}|\Gamma_n'| = \mathbf{P}(\tilde{S}_n > -n\epsilon) \longrightarrow 1$$

and therefore $\mathbf{P}(|\Gamma_n'| < |\Gamma_n|/2) \longrightarrow 0$. Denote

$$A_n := \{\exists w \in \Gamma_{2n} \;:\; S_w > n\epsilon\}\,.$$

Then

$$\mathbf{P}[A_n^c] \le \mathbf{P}(|\Gamma_n'| < |\Gamma_n|/2) + \mathbf{P}(|\Gamma_n'| \ge |\Gamma_n|/2 \text{ and } S_{w(v)} - S_v \le 2n\epsilon \;\forall\, v \in \Gamma_n')\,.$$

The right-hand side is at most

$$\mathbf{P}(|\Gamma_n'| < |\Gamma_n|/2) + \left(1 - e^{-n[I(2\epsilon)+\epsilon]}\right)^{|\Gamma_n|/2},$$

which tends to zero along a subsequence of n values by (109). Taking stock, we infer that $\mathbf{P}(A_n \text{ i.o.}) \ge \lim_n \mathbf{P}(A_n) = 1$, so $s_{\text{cloud}} \ge \epsilon/2$ a.s.

(ii) The implication "$\Rightarrow$" is easy again: if $\dim_P(\partial\Gamma) = 0$, then given $\epsilon > 0$ we can find a cover $\bigcup_i \partial\Gamma^{(i)}$ of $\partial\Gamma$ with $\overline{\dim}_M(\partial\Gamma^{(i)}) \le \epsilon$ for all i. As in the proof of (i),

$$s_{\mathrm{cloud}}(\Gamma^{(i)}) \le \epsilon'$$

for some ϵ' and all i. Whence

$$s_{\mathrm{burst}}(\Gamma^{(i)}) \le s_{\mathrm{cloud}}(\Gamma^{(i)}) \le \epsilon'$$

for all i and so $s_{\mathrm{burst}}(\Gamma) \le \epsilon'$. Here ϵ' can be made arbitrarily small because ϵ may be taken arbitrarily small.

For the reverse implication "$\Leftarrow$", let $d = \dim_P(\partial\Gamma) > 0$. Pick $\epsilon > 0$ small and let

$$\Gamma' = \{v \in \Gamma \colon \dim_P(\Gamma(v)) > d - \epsilon\};$$

here $\Gamma(v) = \{w \in \Gamma \colon w \le v \text{ or } w \ge v\}$.

Now $\rho \in \Gamma'$, so $\Gamma' \ne \emptyset$ and $\dim_P(\partial\Gamma') > d - \epsilon$. Actually, it is easy to see from the definition of packing dimension that

$$\overline{\dim}_M(\partial\Gamma'(v)) > d - \epsilon \quad \text{for all } v \in \Gamma'.$$

By (i) and the definition of cloud-speed, with probability one we can find for each $v \in \Gamma'$ a vertex $w = f(v) \in \Gamma'(v)$ with $w > v$ and $S_w > |w|\beta$ for some fixed $\beta > 0$. The sequence $\rho, f(\rho), f(f(\rho)), \ldots$ is a sequence of vertices $\{v_j\}_{j\ge 0}$ along a ray of Γ such that

$$\frac{S_{v_i}}{|v_i|} > \beta, \qquad \text{for all } i \ge 1.$$

(iii) was proved by Lyons and Pemantle (1992) in the following sharp form:

$$I(s_{\mathrm{sust}}) = \log \mathrm{br}(\Gamma) = \dim_H(\partial\Gamma).$$

(For the other speed notions there is no analogous exact formula.)

The inequality $I(s_{\mathrm{sust}}) \le \log \mathrm{br}(\Gamma)$ is proved using the first-moment method (see the proof of Theorem 5.4). For the other inequality, fix a so that $I(a) < \dim_H(\partial\Gamma)$ and then choose k such that $\mathbf{P}\left(\tilde{S}_k > ka\right) > \mathrm{br}(\Gamma)^k$. Consider a compressed tree $\Gamma[k]$ whose ℓth level is the $k\ell$th level of Γ, with the induced partial order. It is easy to see that $\dim_H(\partial\Gamma[k]) = k\dim_H(\partial\Gamma)$. Define a general percolation on $\Gamma[k]$ in which the edge $\vec{vw}$ is retained if $S_w - S_v > ka$. This general percolation process is not independent; however, for each fixed k, it is quasi-independent. By proposition 7.1, this percolation survives with positive probability, whence $s_{\mathrm{sust}} \ge a$. It follows that $I(s_{\mathrm{sust}}) \ge \log \mathrm{br}(\Gamma)$. $\square$

Exercise 18.5 *Suppose that Γ is an infinite tree without leaves, and its vertices are labeled by i.i.d. variables $X_v \sim N(0,1)$. Denote $d = \dim_M(\partial\Gamma)$. Prove that*

$$\sqrt{d/2} \le s_{\mathrm{cloud}} \le \sqrt{2d}$$

and both bounds can be achieved.

Hint: Use the ideas in the proof of (i) and optimize, or see [9]. These bounds were sharpened by Benassi (1996).

Consider an infinite tree Γ again, label its vertices by i.i.d. real-valued random variables $\{X_v\}_{v\in\Gamma}$, and let $\{S_v\}_{v\in\Gamma}$ be the corresponding tree-indexed random walk. The following question is mostly open.

Open Problem 2 (Bouncing Rays) *Suppose that there a.s. exists a ray $\xi \in \partial\Gamma$ such that $\liminf_{v\in\xi} S_v > -\infty$. Must there a.s. exist a ray $\xi' \in \partial\Gamma$ with $\lim_{v\in\xi'} S_v = +\infty$?*

The only cases for which the answer is known (Pemantle and Peres 1995a) are when

- $X_v = \pm 1$ with probability 1/2 each, or when
- $X_v \sim N(0,1)$.

In these cases there is an exact capacity criterion on the tree for the property to hold. Even in these special cases the proofs are complicated.

19 Recurrence for Tree-Indexed Markov Chains

This chapter is based on Benjamini and Peres (1994a). For a tree Γ and a vertex v, denote by Γ^v the subtree consisting of v and its descendants. We are given a countable state-space G and a set of transition probabilities $\{p(x,y) \colon x,y \in G\}$. the induced Γ-indexed Markov chain is a collection of G-valued random variables $\{S_v\}_{v\in\Gamma}$, with some initial state $S_\rho := x_0 \in G$ and finite-dimensional distributions specified by the following requirement: if $w \in \Gamma$ and v is the parent of w, then

$$\mathbf{P}\Big(S_w = y \mid S_v = x\,,\ S_u \text{ for } u \notin \Gamma_v\Big) = \mathbf{P}(S_w = y \mid S_v = x) = p(x,y)\,.$$

We may think of the state-space G as a graph, with vertices the elements of G and an edge between x and y iff $p(x,y) > 0$. If $p = \{p(x,y) \colon x,y \in G\}$ is irreducible, i.e., for any $x,y \in G$ there exists an n such that $p^n(x,y) > 0$, then the associated graph is connected.

Definitions. A tree-indexed Markov chain is **recurrent** if it returns infinitely often to its starting point with positive probability:

$$\mathbf{P}(S_v = S_\rho \text{ for infinitely many } v \in \Gamma) > 0\,.$$

A stronger requirement is **ray-recurrence**: $\{S_v\}_{v\in\Gamma}$ is ray-recurrent if

$$\mathbf{P}(\exists\, \xi \in \partial\Gamma :\ S_v = S_\rho \text{ for infinitely many } v \in \xi) > 0\,.$$

In general, recurrence does not imply ray-recurrence (even when $G = \mathbf{Z}^3$). Indeed, the 3-1 tree has exponential growth (which yields recurrence for $G = \mathbf{Z}^d$), yet it has a countable boundary (which precludes ray-recurrence on any transient G).

The probabilities in the definitions of recurrence and ray-recurrence may lie strictly between 0 and 1, even when the indexing tree is a binary tree. If G is a group and the transition probabilities are G-invariant, then there are zero–one laws for both notions of recurrence.

Given a state space G, an irreducible stochastic matrix $p = \{p(x,y) : x, y \in G\}$ and a finite subset F of G, write $\rho(p_F)$ for the spectral radius of the substochastic matrix $p_F = \{p(x,y) : x, y \in F\}$. We then define

$$\rho(G,p) = \sup_{F \text{ finite}} \rho(p_F)\,.$$

Then

$$\mathbf{P}(\exists\, \xi \in \partial\Gamma \text{ with bounded trajectory}) > 0 \quad \Leftrightarrow \quad \mathrm{br}(\Gamma) > \frac{1}{\rho(G,p)}\,.$$

Simple random walk on Z has spectral radius 1, but we can make a quantitative statement on rays with bounded trajectories: For the Γ-indexed simple random walk on $\mathbf{Z}$,

$$\mathrm{br}(\Gamma) > \frac{1}{\cos\left(\pi/(b+1)\right)}$$

is sufficient for the existence of a ray with trajectory in $\{0, 1, \ldots, b-1\}$ to have positive probability, and

$$\mathrm{br}(\Gamma) \geq \frac{1}{\cos\left(\pi/(b+1)\right)}$$

is necessary.

Finally, we note that recurrence of a Γ-indexed Markov chain on G is related to a comparison of the Minkowski dimension of Γ and the spectral radius of G, while ray-recurrence is related to a comparison of packing dimension and spectral radius. In particular, $\dim_M(\partial\Gamma) < -\log[\rho(G,p)]$ implies non-recurrence and $\dim_P(\partial\Gamma) < -\log[\rho(G,p)]$ implies non-ray-recurrence.

More details on the notions described in this chapter, and some amusing examples, can be found in [8, 9]. Benjamini and Schramm [10] give an application of tree-indexed Markov chains to a problem in discrete geometry.

20 Dynamical Percolation

This chapter is based on Häggström, Peres, and Steif (1997).

Consider Bernoulli(p) percolation on an infinite graph G. Recall that each edge is, independently, open with probability p. As before, $\mathbf{P}_{G,p} = \mathbf{P}_p$ will denote this product measure. Write $\mathcal{C}$ for the event that there exists an infinite open cluster. Recall that by Kolmogorov's 0-1 law, the probability of $\mathcal{C}$ is, for fixed G and p, either 0 or 1. As remarked previously, there exists a critical probability $p_c = p_c(G) \in [0,1]$ such that

$$\mathbf{P}_p(\mathcal{C}) = \begin{cases} 0 & \text{for } p < p_c \\ 1 & \text{for } p > p_c. \end{cases}$$

At $p = p_c$ we can have either $\mathbf{P}_p(\mathcal{C}) = 0$ or $\mathbf{P}_p(\mathcal{C}) = 1$, depending on G.

In this chapter we consider a dynamical variant of percolation. Given $p \in (0,1)$, we want the set of open edges to evolve so that at any fixed time $t \geq 0$, the distribution of this set is $\mathbf{P}_p$. The most natural way to accomplish this is to let the distribution at time 0 be given by $\mathbf{P}_p$, and to let each edge change its status (open or closed) according to a continuous time, stationary 2-state Markov chain, independently of all other edges. For an edge e of G, write $\eta_t(e) = 1$ if e is open at time t, and $\eta_t(e) = 0$ otherwise. The entire configuration of open and closed edges at time t, denoted η_t, can then be regarded as an element of $X = \{0,1\}^E$ (where E is the edge set of G). The evolution of η_t is a Markov process, and can be viewed as the simplest type of particle system. Each edge flips (changes its value) at rate

$$\lambda(\eta_t, e) = \begin{cases} p & \text{if } \eta_t(e) = 0 \\ 1-p & \text{if } \eta_t(e) = 1 \end{cases}$$

and the probability that two edges flip simultaneously is 0. Write $\mathbf{\Psi}_{G,p}$ (or $\mathbf{\Psi}_p$) for the underlying probability measure of this Markov process, and write $\mathcal{C}_t$ for the event that there is an infinite cluster of open edges in η_t. Since $\mathbf{P}_p$ is a stationary measure for this Markov process, Fubini's theorem implies that

$$\begin{cases} \mathbf{\Psi}_p(\ \mathcal{C}_t \text{ occurs for Lebesgue a.e. } t) = 1 & \text{if} \quad \mathbf{P}_p(\mathcal{C}) = 1 \\ \mathbf{\Psi}_p(\neg\mathcal{C}_t \text{ occurs for Lebesgue a.e. } t) = 1 & \text{if} \quad \mathbf{P}_p(\mathcal{C}) = 0 \end{cases}$$

where $\neg\mathcal{C}_t$ denotes the complement of $\mathcal{C}_t$. The main question studied here is the following,

Question 20.1 *For which graphs can the quantifier "for a.e. t" in the above statements be replaced by "for every t"?*

For $p \neq p_c$, the answer is all graphs.

Proposition 20.2 *For any graph G we have*

$$\begin{cases} \mathbf{\Psi}_p(\ \mathcal{C}_t \ \textit{occurs for every } t\,) = 1 & \textit{if} \quad p > p_c(G) \\ \mathbf{\Psi}_p(\neg\mathcal{C}_t \ \textit{occurs for every } t) = 1 & \textit{if} \quad p < p_c(G)\,. \end{cases} \tag{110}$$

Notation: For $0 \leq a \leq b < \infty$ and any edge e of a graph G, we abbreviate

$$\inf_{[a,b]} \eta(e) := \inf_{t \in [a,b]} \eta_t(e).$$

and write $\mathcal{C}^{\inf}_{[a,b]}$ for the event that there is an infinite cluster of edges with $\inf_{[a,b]} \eta(e) = 1$. Analogously, define $\sup_{[a,b]} \eta$, and let $\mathcal{C}^{\sup}_{[a,b]}$ be the event that there is an infinite cluster of edges with $\sup_{[a,b]} \eta(e) = 1$.

Proof. (i) Suppose $p > p_c$. Let $0 < \epsilon < p - p_c$ and observe that for every edge e,

$$\mathbf{\Psi}_p\Big\{ \inf_{[0,\epsilon]} \eta(e) = 1 \Big\} = p \exp(-(1-p)\epsilon) > p - \epsilon > p_c\,.$$

Since the events $\Big\{\inf_{[0,\epsilon]}\eta(e) = 1\Big\}$ are mutually independent as e ranges over the edges of G, it follows from the definition of p_c that $\mathbf{\Psi}_p\Big[\mathcal{C}^{\inf}_{[0,\epsilon]}\Big] = 1$ and therefore

$$\mathbf{\Psi}_p\Big(\mathcal{C}_t \text{ occurs for all } t \in [0,\epsilon]\Big) = 1\,.$$

Repeating the argument for the intervals $[k\epsilon, (k+1)\epsilon]$ with integer k and using countable additivity, we obtain the supercritical part of the proposition.
(ii) A similar argument proves that for $p < p_c$ there is never an infinite open cluster. We take $\epsilon \in (0, p_c - p)$ and find that

$$\mathbf{\Psi}_p\Big\{\sup_{[0,\epsilon]}\eta(\mathrm{e}) = 1\Big\} = 1-(1-p)\exp(-p\epsilon) < p + p\epsilon < p_c\,. \tag{111}$$

Therefore $\mathbf{\Psi}_p\Big(\mathcal{C}^{\sup}_{[0,\epsilon]}\Big) = 0$, whence there is a.s. no infinite cluster for any $t \in [0,\epsilon]$. Countable additivity concludes the argument. $\square$

At the critical value $p_c(G)$ the situation is more delicate.

Theorem 20.3 *There exists a graph G_1 with the property that at $p = p_c(G)$ we have $\mathbf{P}_{G,p}(\mathcal{C}) = 0$ but $\mathbf{\Psi}_{G,p}\Big(\cup_{t>0}\,\mathcal{C}_t\Big) = 1$. (The latter probability is 0 or 1 for any graph.) There also exists a graph G_2 such that for $p = p_c(G_2)$ we have $\mathbf{P}_{G_2,p}(\mathcal{C}) = 1$, yet $\mathbf{\Psi}_{G_2,p}(\cap_{t>0}\mathcal{C}_t) = 0$.*

The graphs for which percolation problems have been studied most extensively are the lattices $\mathbf{Z}^d$, and trees. On $\mathbf{Z}^2$, the critical value p_c is $1/2$ and $\mathbf{P}_{p_c}(\mathcal{C}) = 0$ (see Kesten (1980)); for $d > 2$ the precise value of $p_c(\mathbf{Z}^d)$ is not known. Hara and Slade (1994) showed that $\mathbf{P}_{p_c}(\mathcal{C}) = 0$ for $\mathbf{Z}^d$ if $d \geq 19$, and it is certainly believed that this holds for all d.

Theorem 20.4 *Let G be either the integer lattice $\mathbf{Z}^d$ with $d \geq 19$ or a regular tree. Then $\mathbf{\Psi}_{G,p_c}(\neg\mathcal{C}_t$ occurs for every $t) = 1$.*

Remark. It is not known whether $G = \mathbf{Z}^2$ can be included in Theorem 20.4. Let $\theta(p)$ denote the $\mathbf{P}_p$-probability that the origin is in an infinite open cluster. The proof of Theorem 20.4 for $G = Z^d$ with $d \geq 19$ uses more information than just $\theta(p_c) = 0$; it also uses that θ has a finite right derivative at p_c. In $\mathbf{Z}^2$ it is known that $\theta(p_c) = 0$, but Kesten and Zhang proved that the right derivative of θ is infinite at p_c.

Next, we consider dynamical percolation on general trees. In Chapter 14, we proved R. Lyons' criterion for $\mathbf{P}_p(\mathcal{C}) > 0$ in terms of effective electrical resistance (see (39)); effective resistance is easy to calculate on trees using the parallel and series laws. Here we obtain such a criterion for dynamical percolation.

For an infinite tree Γ with root ρ, as before we write Γ_n for the set of vertices at distance exactly n from ρ, the nth level of Γ. Recall that a tree is spherically symmetric if all vertices on the same level have equally many children.

Theorem 20.5 *Let $\{\eta_t\}$ be a dynamical percolation process with parameter $0 < p < 1$ on an infinite tree Γ. Assign each edge between levels $n-1$ and n of Γ the resistance p^{-n}/n. If in the resulting resistor network the effective resistance from the root to infinity is finite, then $\mathbf{\Psi}_{\Gamma,p}$-a.s. there exist times $t > 0$ such that Γ has an infinite open cluster, while if this resistance is infinite, then a.s. there are no such times. In particular, if Γ is spherically symmetric, then*

$$\mathbf{\Psi}_{\Gamma,p}(\cup_{t>0}\mathcal{C}_t) = 1 \text{ if and only if } \sum_{n=1}^{\infty} \frac{p^{-n}}{n|\Gamma_n|} < \infty\,. \tag{112}$$

Recall R. Lyons' criterion for the percolation probability on a general tree Γ to be positive: Suppose that $0 < p < 1$ and assign each edge between levels $n-1$ and n resistance p^{-n}. Then $\mathbf{P}_{\Gamma,p}(\mathcal{C}) > 0$ iff the resulting effective resistance from the root to infinity is finite. Thus a spherically symmetric tree Γ with $p = p_c(\Gamma) \in (0,1)$, has $\mathbf{\Psi}_{\Gamma,p}(\cup_{t>0}\mathcal{C}_t) = 1$ but $\mathbf{P}_{\Gamma,p}(\mathcal{C}) = 0$ iff the series in (112) converges but $\sum_{n=1}^{\infty} \frac{p^{-n}}{|\Gamma_n|} = \infty$.

In the course of the proof of Theorem 20.5, we obtain bounds for the probability that there exists a time $t \in [0,1]$ for which there is an open path in η_t from the root to the nth level Γ_n. For example, on the regular tree $\mathbf{T}^k$ with $p = 1/k$, this probability is bounded between constant multiples of $1/\log n$. (The probability under $\mathbf{P}_{1/k}$ that an open path exists from ρ to the nth level of $\mathbf{T}^k$, is bounded between constant multiples of $1/n$; this follows from Kolmogorov's theorem on critical branching processes, see Athreya and Ney (1972).) For a general tree these bounds, given in Theorem 20.9, can be expressed in terms of the effective resistance from the root to Γ_n, and the ratio of the upper and lower bounds is an absolute constant.

For a graph with $\mathbf{\Psi}_{G,p}(\cup_{t>0}\mathcal{C}_t) = 1$ but $\mathbf{P}_{G,p}(\mathcal{C}) = 0$, the set of percolating times at criticality has zero Lebesgue measure, so it is natural to ask for its Hausdorff dimension. For spherically symmetric trees there is a complete answer.

Theorem 20.6 *Let $p \in (0,1)$ and let Γ be a spherically symmetric tree. If the set of times $\{t \in [0,\infty) : \mathcal{C}_t \text{ occurs}\}$ is a.s. nonempty, then $\mathbf{\Psi}_p$-a.s. this set has Hausdorff dimension*

$$\sup\Big\{\, \alpha \in [0,1] \;:\; \sum_{n=1}^{\infty} \frac{p^{-n} n^{\alpha-1}}{|\Gamma_n|} < \infty \Big\}.$$

(*Note that this series converges for $\alpha = 0$ by* (112).)

Here are some interesting trees with $\mathbf{\Psi}_{T,p}(\cup_{t>0}\mathcal{C}_t) = 1$ but $\mathbf{P}_{T,p}(\mathcal{C}) = 0$:

Example 20.7 Let Γ be the spherically symmetric tree where each vertex on level n has 4 children if $n = 1, 2, 4 \ldots$ is a power of 2, and 2 children otherwise. Then it is easily seen that $n2^n \le |\Gamma_n| \le 2n2^n$ for all $n > 0$. Combining Theorem 20.6 with the result of R. Lyons quoted after Theorem 20.5, we see that $\mathbf{\Psi}_{1/2}$-a.s. the set of times for which percolation occurs on Γ has Hausdorff dimension 1 but Lebesgue measure 0. $\triangle$

Example 20.8 Let $0 < p, \beta < 1$, and suppose that Γ is a spherically symmetric tree with $|\Gamma_n| = p^{-n} n^{\beta+o(1)}$ as $n \to \infty$. Then Theorem 20.6 implies that $\mathbf{\Psi}_p$-a.s. the set of times for which percolation occurs on Γ has Hausdorff dimension β. $\triangle$

Since we will introduce an auxiliary random killing time τ, we denote the underlying probability measure $\mathbf{P}$ rather than $\mathbf{\Psi}_p$. The event that there is an open path from the root to $\partial\Gamma$ in η_t is denoted $\{\rho \overset{t}{\leftrightarrow} \partial\Gamma\}$.

Theorem 20.9 *Consider dynamical percolation* $\{\eta_t\}$ *with parameter* $0 < p < 1$ *on a tree* Γ *which is either finite or infinite with* $\mathbf{P}_{\Gamma,p}(\mathcal{C}) = 0$*. Let* τ *be a random variable with an exponential distribution of mean 1, which is independent of the process* $\{\eta_t\}$*. Let*

$$h(n) = \frac{p^{-n}}{n+1} \cdot \frac{1-p^{n+1}}{1-p} \quad \text{for } n \geq 0\,. \tag{113}$$

Then the event $A = \{\exists t \in [0,\tau] \,:\, \rho \overset{t}{\leftrightarrow} \partial\Gamma\}$ *satisfies for some constant* C

$$\frac{1}{2}\mathrm{Cap}_h(\partial\Gamma) \leq \mathbf{P}(A) \leq 2C\mathrm{Cap}_h(\partial\Gamma)\,, \tag{114}$$

Remarks:

(i) It is easy to verify that h is increasing and $h(n) \leq p^{-n}$ for all n. These properties also follow from the interpretation of h given in Lemma 20.10(iii) below. In the sequel, we will sometimes write $h(v)$ instead of $h(|v|)$ when v is a vertex.

(ii) The event A is easier to work with than the perhaps more natural event $B = \{\exists t \in [0,1] \,:\, \rho \overset{t}{\leftrightarrow} \partial\Gamma\}$. Noting that $\mathbf{P}(B) \leq \mathbf{P}(A|\tau > 1) \leq \mathbf{P}(A)/e^{-1}$ and $\mathbf{P}(A) \leq \sum_{k=0}^{\infty} e^{-k}\mathbf{P}(B) = \mathbf{P}(B)/(1-e^{-1})$, we obtain

$$\frac{1-e^{-1}}{2}\mathrm{Cap}_h(\partial\Gamma) \leq \mathbf{P}(B) \leq 2eC\mathrm{Cap}_h(\partial\Gamma)\,.$$

We will only prove the lower bound in Theorem 20.9; consult [37] for the other inequality. We will need a lemma concerning the behavior of a pair of paths.
Notation: Denote by $\{v \overset{t}{\leftrightarrow} w\}$ the event that there is an open path in η_t between the vertices v and w. Similarly, when x is a ray of the tree, $\{\rho \overset{t}{\leftrightarrow} x\}$ means that x is open at time t. Thus $\{\rho \overset{t}{\leftrightarrow} \partial\Gamma\} = \bigcup_{x\in\partial\Gamma} \{\rho \overset{t}{\leftrightarrow} x\}$. For $s > 0$ let $T_v(s) := \int_0^s \mathbf{1}_{\{\rho \overset{t}{\leftrightarrow} v\}}\, dt$ be the amount of time in $[0,s]$ when the path from the root to v is open.

Lemma 20.10 *Let* u *and* w *be vertices of* Γ*. With the notation of Theorem 20.9 in force,*

(i) $\mathbf{E}[T_w(\tau)] = p^{|w|}$

(ii) $\mathbf{E}[T_w(\tau) \mid T_w(\tau) > 0] = \mathbf{E}[T_w(\tau) \mid \rho \overset{0}{\leftrightarrow} w] = h(w)p^{|w|}$

(iii) $\mathbf{P}(T_w(\tau) > 0) = h(w)^{-1}$

(iv) $\mathbf{E}[T_u(\tau)T_w(\tau)] = 2h(u \wedge w)p^{|u|+|w|}$

Proof: Let $q = 1 - p$.

(i) This is immediate from Fubini's Theorem.

(ii) The first equality follows from the lack of memory of the exponential distribution. Verifying the second equality requires a calculation:

$$\begin{aligned}
\mathbf{E}[T_w(\tau) \mid \rho \overset{0}{\leftrightarrow} w] &= \int_0^\infty \mathbf{P}(\rho \overset{t}{\leftrightarrow} w \mid \rho \overset{0}{\leftrightarrow} w)\mathbf{P}(\tau > t)\, dt \\
&= \int_0^\infty (p + qe^{-t})^{|w|} e^{-t}\, dt \;=\; \frac{-(p+qe^{-t})^{|w|+1}}{(|w|+1)q}\Big|_{t=0}^{\infty}.
\end{aligned}$$

(iii) The required probability is the ratio of the expectations in (i) and (ii).

(iv) Since the process $\{\eta_t\}$ is reversible,

$$\begin{aligned}
\mathbf{E}[T_u(\tau)T_w(\tau)] &= \mathbf{E}\int_0^\tau\!\!\int_0^\tau \mathbf{1}_{\{\rho \overset{s}{\leftrightarrow} u\}}\mathbf{1}_{\{\rho \overset{t}{\leftrightarrow} w\}}\, dt\, ds \\
&= 2\int_0^\infty\!\!\int_s^\infty \mathbf{P}(\rho \overset{s}{\leftrightarrow} u)\mathbf{P}\Big(\rho \overset{t}{\leftrightarrow} w \mid \rho \overset{s}{\leftrightarrow} u\Big)e^{-t}\, dt\, ds\,. \qquad (115)
\end{aligned}$$

Observe that for $t > s$,

$$\mathbf{P}\Big(\rho \overset{t}{\leftrightarrow} w \mid \rho \overset{s}{\leftrightarrow} u\Big) = p^{|w|-|u\wedge w|}\mathbf{P}\Big(\rho \overset{t}{\leftrightarrow} u \wedge w \mid \rho \overset{s}{\leftrightarrow} u \wedge w\Big).$$

Change variables $\tilde{t} = t - s$ in (115) to get that $\mathbf{E}[T_u(\tau)T_w(\tau)]$ equals

$$\begin{aligned}
&= 2p^{|w|-|u\wedge w|}\int_0^\infty\!\!\int_0^\infty \mathbf{P}(\rho \overset{s}{\leftrightarrow} u)e^{-s-\tilde{t}}\mathbf{P}\Big(\rho \overset{\tilde{t}}{\leftrightarrow} u \wedge w \mid \rho \overset{0}{\leftrightarrow} u \wedge w\Big)\, d\tilde{t}\, ds \\
&= 2p^{|w|-|u\wedge w|}\mathbf{E}[T_u(\tau)] \cdot \mathbf{E}[T_{u\wedge w}(\tau) \mid \rho \overset{0}{\leftrightarrow} u \wedge w]\,.
\end{aligned}$$

Substituting parts (i) and (ii) of the lemma into the last equation proves (iv).

□

Proof of lower bound in Theorem 20.9. We prove the theorem when Γ is a finite tree; the general case then follows by an appropriate limiting procedure. The lower bound on $\mathbf{P}(A)$ is proved via the second moment method. Let μ be a probability measure on $\partial\Gamma$, and consider the random variable

$$Z := \sum_{v \in \partial\Gamma} T_v(\tau)p^{-|v|}\mu(v)\,. \qquad (116)$$

Lemma 20.10(i) implies that $\mathbf{E}(Z) = 1$. Part (iv) of the same lemma gives

$$\mathbf{E}[Z^2] = \sum_{v \in \partial\Gamma} \sum_{w \in \partial\Gamma} \mathbf{E}[T_v(\tau) T_w(\tau)] p^{-|v|-|w|} \mu(v)\mu(w) = 2\mathcal{E}_h(\mu)\,. \tag{117}$$

Using the Cauchy-Schwarz inequality we find that

$$\mathbf{P}(A) \geq \mathbf{P}(Z > 0) \geq \frac{\mathbf{E}[Z]^2}{\mathbf{E}[Z^2]} = \frac{1}{2\mathcal{E}_h(\mu)}\,.$$

Taking the supremum of the right-hand side over all probability measures μ on $\partial\Gamma$ proves the lower bound on $\mathbf{P}[A]$ in (114). □

We include the statement of one result from Peres and Steif (1998).

Theorem 20.11 *Let Γ be an infinite spherically symmetric tree, $p = p_c(\Gamma) \in (0,1)$ and T^k denote the set of times in $[0,\infty)$ when there are at least k infinite clusters. Suppose that $\mathbf{P}_p(\mathcal{C}) = 0$. Let*

$$\alpha_c := \sup\Big\{ \alpha \in [0,1] \,:\, \sum_{n=1}^{\infty} \frac{p^{-n} n^{\alpha-1}}{|\Gamma_n|} < \infty \Big\}\,.$$

Then for all k, the Hausdorff dimension of T^k is

$$\max\{0, 1 - k(1-\alpha_c)\} \quad \mathbf{\Psi}_p\text{–a.s.}\,. \tag{118}$$

21 Stochastic Domination Between Trees

For a tree Γ with total height $N \leq \infty$, label its vertices by i.i.d. random variables $\{X_v\}_{v\in\Gamma}$. If $B \subseteq \mathbf{R}^N$ is a Borel set, we write

$$\mathbf{P}(B;\Gamma) = \mathbf{P}(\exists\, \xi \in \partial\Gamma : (X_v)_{v\in\xi} \in B)\,.$$

For two such trees Γ and Γ' of height $N \leq \infty$, labeled by $\{X_v\}_{v\in\Gamma}$ and $\{X'_v\}_{v\in\Gamma'}$ respectively, we say that Γ' **stochastically dominates** Γ if for any Borel set $B \subseteq \mathbf{R}^N$,

$$\mathbf{P}(B;\Gamma) \leq \mathbf{P}(B;\Gamma')\,.$$

To verify that one tree dominates another, it suffices to consider the case where the X_v are i.i.d. uniform random variables in $[0,1]$, since other random variables can be written as functions of these.

Recall that a tree Γ is spherically symmetric if all vertices in Γ_n have the same number of offspring.

Theorem 21.1 (Pemantle and Peres 1994) *Let Γ' be a spherically symmetric tree and Γ another (arbitrary) tree. Then Γ' stochastically dominates Γ iff $|\Gamma_n| \leq |\Gamma'_n|$ for all $n \geq 1$.*

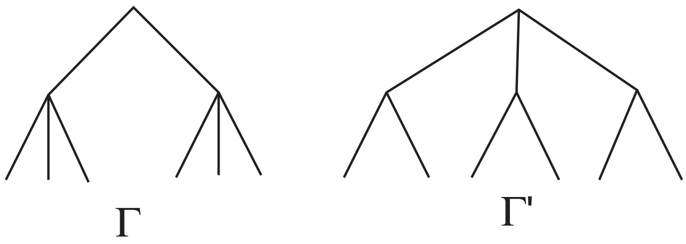


Figure 8: Γ **is dominated by** Γ'**.**

Example 21.2 *Two trees of height 2.*

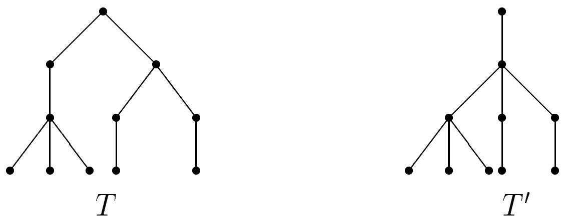
Let Γ be the tree of height 2 in which the root has two offspring and each of these three offspring. Let Γ' be the tree for which the root has three offspring and and each of these two offspring.

Then it is not clear a priori which tree dominates. The result above yields that Γ is dominated by Γ'. $\triangle$

Stochastic domination between trees is well understood only for trees which are either spherically symmetric or have height two. Already for trees of height three, the domination order is somewhat mysterious, as the following example from Pemantle and Peres (1994) demonstrates.

Example 21.3 *Comparison between a tree T and T with vertices glued.*

Consider the trees T and T' in the next figure, where T' is obtained from T by gluing together the vertices in the first generation.



Intuitively, it seems that T should dominate T', but this is not the case. If

$$B^c = ([0,1/2] \times [0,1] \times [0,2/3]) \cup ([1/2,1] \times [0,1/2] \times [0,1])$$

and the X_v are uniform on $[0,1]$, then the probability that $(X_{v_1}, X_{v_2}, X_{v_3}) \in B^c$ for all paths (ρ, v_1, v_2, v_3) in T is $1075/7776$, while the corresponding probability for T' is only $998/7776$. $\triangle$

A consequence of Theorem 21.1 is that, among all trees of height n with $|\Gamma_n| = k$, the tree $T(n,k)$ consisting of k disjoint paths joined at the root is maximal in the stochastic order. If the common law of the X_v is μ and $B \subseteq \mathbf{R}^n$, then $1 -$

$\mathbf{P}(B; T(n,k)) = (1 - \mu^n(B))^k$, where μ^n is n-fold product measure; thus for any tree Γ of height n,

$$1 - \mathbf{P}(B; \Gamma) \geq (1 - \mu^n(B))^k .$$

The definition of $\mathbf{P}(B; \Gamma)$ extends naturally to any graded graph Γ, a finite graph whose vertices are partitioned into levels $1, \ldots, n$ and oriented edges allowed only between vertices in adjacent levels. The following is a natural conjecture.

Conjecture 3 *For any graded graph Γ of height n, let $K(\Gamma)$ be the number of oriented paths that pass through every level of Γ and let X_v be i.i.d. random variables with common law μ. Then for any $B \subseteq \mathbf{R}^n$,*

$$1 - \mathbf{P}(B; \Gamma) \geq (1 - \mu^n(B))^{K(\Gamma)} .$$

If B is upwardly closed (that is, $\mathbf{x} \in B$ and $\mathbf{y} \geq \mathbf{x}$ coordinate-wise imply $\mathbf{y} \in B$), then the conjecture is an easy consequence of the FKG inequality. The case $n = 2$ corresponds to a bipartite graph; Conjecture 3 for this case is due to Sidorenko (1994), who stated it (and proved it in many special cases) in the following analytic form:

Sidorenko's Conjecture: *Let $f : [0,1]^2 \to [0, \infty)$ be a nonnegative bounded measurable function and consider the bipartite graph with vertices $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$. If E is the edge-set of this graph, then*

$$\int \cdots \int \prod_{X_i \sim Y_j} f(x_i, y_j) dx_1 \ldots dx_n dy_1 \ldots dy_m \geq \left(\iint f(x,y) dx dy \right)^{|E|} . \tag{119}$$

For the bipartite graph consisting of three vertices X, Y, Z and two edges XY and XZ, the conjecture reads

$$\iiint f(x,y) f(x,z) dx dy dz \geq \left(\iint f(x,y) dx dy \right)^2$$

and can be easily proved using the Cauchy-Schwarz inequality.

Exercise 21.4 *Prove Sidorenko's conjecture for the bipartite graph with four vertices and three edges, XY, XZ, and WZ. (Hint: use Hölder's inequality with $p = 3$ and $q = 3/2$.)*

Sidorenko proved his conjecture for bipartite graphs with at most one cycle, and for bipartite graphs where one side has at most four vertices. For general finite bipartite graphs, it is still open whether (119) always holds.

We conclude with yet another **problem:** In the statement of Theorem 16.7 we defined an information-theoretic domination relation between trees. It would be quite interesting to compare that relation with the stochastic domination relation studied in this chapter.

Bibliography

[1] M. Aizenman (1985). The intersection of Brownian paths as a case study of a renormalisation method for quantum field theory. *Commun. Math. Phys.* **97**, 91–110.

[2] P. Antal and A. Pisztora (1996). On the chemical distance in supercritical Bernoulli percolation. *Ann. Probab.* **24** 1036–1048.

[3] K. Athreya and P. Ney (1972). *Branching Processes.* Springer-Verlag, New York.

[4] A. Benassi (1996). Arbres et grandes déviations. In *Trees*, B. Chauvin, S. Cohen, A. Rouault (Editors), Birkhäuser.

[5] I. Benjamini, R. Pemantle and Y. Peres (1995). Martin capacity for Markov Chains. *Ann. Probab.* **23**, 1332–1346.

[6] I. Benjamini, R. Pemantle and Y. Peres (1998). Unpredictable paths and percolation. *Ann. Probab.* **26**, 1198–1211.

[7] I. Benjamini and Y. Peres (1992). Random walks on a tree and capacity in the interval. *Ann. IHP Probab. et. Statist.* **28**, 557–592.

[8] I. Benjamini and Y. Peres (1994a). Markov chains indexed by trees. *Ann. Probab.* **22**, 219–243.

[9] I. Benjamini and Y. Peres (1994b). Tree-indexed random walks and first passage percolation. *Probab. Th. Rel. Fields* **98**, 91–112.

[10] I. Benjamini and O. Schramm (1997) Every graph with a positive Cheeger constant contains a tree with a positive Cheeger constant, *Geom. Funct. Anal.* **7**, 403–419.

[11] J. D. Biggins (1977). Chernoff's theorem in the branching random walk. *J. Appl. Prob.* **14**, 630–636.

[12] P. M. Bleher, J. Ruiz and V. A. Zagrebnov (1995). On the purity of the limiting Gibbs state for the Ising model on the Bethe lattice. *J. Stat. Phys.* **79**, 473–482.

[13] J. Cavender (1978). Taxonomy with confidence. *Math. BioSci.* **40**, 271–280.

[14] J. T. Chayes, L. Chayes, J. P. Sethna and D. J. Thouless (1986). A mean field spin glass with short range interactions, *Comm. Math. Phys.* **106**, 41–89.

[15] T. M. Cover and J. A. Thomas (1991). *Elements of Information Theory.* Wiley, New York.

[16] K. L. Chung (1973). Probabilistic approach in potential theory to the equilibrium problem. *Ann. Inst. Fourier, Grenoble* **23**, 313–322.

[17] T. Cox and R. Durrett (1983). Oriented percolation in dimensions $d \geq 4$: bounds and asymptotic formulas. *Math. Proc. Camb. Phil. Soc.* **93**, 151–162.

[18] P. G. Doyle and E. J. Snell (1984). *Random walks and Electrical Networks.* Carus Math. Monographs **22**, Math. Assoc. Amer., Washington, D. C.

[19] L. Dubins and D. Freedman (1967). Random distribution functions. *Proc. fifth Berkeley Symp. on math. stat. and probab.* (LeCam and Neyman, Eds.), Vol. II, Part 1, 183–214.

[20] R. Durrett (1996). *Probability: Theory and Examples*, Second edition. Duxbury Press, Belmont, CA.

[21] A. Dvoretsky, P. Erdős and S. Kakutani (1950). Double points of paths of Brownian motion in n-space. *Acta Sci. Math. Szeged* **12**, 75–81.

[22] A. Dvoretsky, P. Erdős, S. Kakutani and S. J. Taylor (1957). Triple points of Brownian motion in 3-space. *Proc. Camb. Phil. Soc.* **53**, 856–862.

[23] W. Evans (1994). *Information Theory and Noisy Computation.* PhD thesis, Dept. of Computer Science, University of California at Berkeley.

[24] W. Evans, C. Kenyon, Y. Peres and L. J. Schulman (1995). A critical phenomenon in a broadcast process. *Unpublished manuscript.*

[25] W. Evans, C. Kenyon, Y. Peres and L. J. Schulman (1998). Broadcasting on trees and the Ising model. *Preprint.*

[26] W. Evans and L. J. Schulman (1993). Signal propagation, with application to a lower bound on the depth of noisy formulas. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science*, 594–603.

[27] S. Evans (1987). Multiple points in the sample paths of a Levy processes. *Probab. Th. Rel. Fields* **76**, 359–376.

[28] P.J. Fitzsimmons and T. Salisbury (1989). Capacity and energy for multiparameter Markov processes. *Ann. IHP Probab. et. Statist.* **25**, 325–350.

[29] H. Furstenberg (1967). Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Systems Theory* **1**, 1–49.

[30] H. O. Georgii (1988). *Gibbs Measures and Phase Transitions.* W. de Gruyter, Berlin.

[31] G. Grimmett (1989). *Percolation.* New York, Springer Verlag.

[32] G. R. Grimmett, Percolation and disordered systems, in *Lectures in Probability Theory and Statistics, Ecole d'Ete de Probabilites de Saint-Flour* **XXVI**-1996, P. Bernard (ed.), Springer Lecture Notes in Math **1665**, 153–300.

[33] G. R. Grimmett and H. Kesten (1984). *Unpublished.*

[34] G. R. Grimmett, H. Kesten and Y. Zhang (1993). Random walk on the infinite cluster of the percolation model. *Probab. Th. Rel. Fields* **96**, 33–44.

[35] G. R. Grimmett and J. M. Marstrand (1990). The supercritical phase of percolation is well behaved. *Proc. Royal Soc. London Ser. A* **430**, 439–457.

[36] O. Häggström and E. Mossel (1998). Nearest-neighbor walks with low predictability profile and percolation in $2+\epsilon$ dimensions. *Ann. Probab.* **26**, 1212–1231.

[37] O. Häggström, Y. Peres and J. Steif (1997). Dynamical Percolation. *Ann. IHP Probab. et. Statist.* **33**, 497–528.

[38] J.M. Hammersley (1974). Postulates for subadditive processes. *Ann. Probab.* **2**, 652–680.

[39] T. Hara and G. Slade (1994). Mean field behavior and the lace expansion. In *Probability Theory and Phase Transitions* (ed. G. Grimmett), Proceedings of the NATO ASI meeting in Cambridge 1993, Klewer.

[40] J. Hawkes (1981). Trees generated by a simple branching process. *J. London Math. Soc.* **24**, 373–384.

[41] P. Hiemer (1998). Dynamical renormalisation in oriented percolation. *Preprint.*

[42] Y. Higuchi (1977). Remarks on the limiting Gibbs state on a (d+1)-tree. *Publ. RIMS Kyoto Univ.* **13**, 335–348.

[43] C. Hoffman (1998). Unpredictable nearest neighbor processes. *Preprint.*

[44] D. Ioffe (1996a). A note on the extremality of the disordered state for the Ising model on the Bethe lattice. *Lett. Math. Phys.* **37**, 137–143.

[45] D. Ioffe (1996b). A note on the extremality of the disordered state for the Ising model on the Bethe lattice. In *Trees*, B. Chauvin, S. Cohen, A. Rouault (Editors), Birkhäuser.

[46] J. P. Kahane and J. Peyrière (1976). Sur certaines martingales de B. Mandelbrot. *Advances Math.* **22**, 131–145.

[47] S. Kakutani (1944a). On Brownian motion in n-space. *Proc. Imp. Acad. Tokyo* **20**, 648–652.

[48] S. Kakutani (1944b). Two dimensional Brownian motion and harmonic functions. *Proc. Imp. Acad. Tokyo* **20**, 706–714.

[49] H. Kesten (1980). The critical probability of bond percolation on the square lattice equals $\frac{1}{2}$. *Commun. Math. Phys.* **109**, 109–156.

[50] H. Kesten and B. P. Stigum (1966a). A limit theorem for multidimensional Galton-Watson processes. *Ann. Math. Stat.* **37**, 1211–1223.

[51] H. Kesten and B. P. Stigum (1966b). Additional limit theorems for indecomposable multidimensional Galton-Watson processes. *Ann. Math. Stat.* **37**, 1463–1481.

[52] J.F.C. Kingman (1975). The first birth problem for an age-dependent branching process. *Ann. Probab.* **3**, 790–801.

[53] J.-F. Le Gall (1992) Some properties of planar Brownian motion. *École d'été de probabilités de Saint-Flour XX*, Lecture Notes in Math. **1527**, 111–235. Springer, New York.

[54] G. Lawler (1982). The probability of intersection of independent random walks in four dimensions. *Commun. Math. Phys.* **86**, 539–554.

[55] G. Lawler (1985). Intersections of random walks in four dimensions II. *Commun. Math. Phys.* **97**, 583–594.

[56] L. Le Cam (1974). *Notes on Asymptotic Methods in Statistical Decision Theory.* Centre de Rech. Math. Univ. Montréal.

[57] D. Levin and Y. Peres (1998). Energy and Cutsets in Infinite Percolation Clusters. To appear in *Proceedings of the Cortona Workshop on Random Walks and Discrete Potential Theory*, M. Picardello and W. Woess (editors).

[58] T. M. Liggett, R. H. Schonmann and A. M. Stacey (1996). Domination by product measures. *Ann. Probab.* **24**, 1711–1726.

[59] R. Lyons (1989). The Ising model and percolation on trees and tree-like graphs. *Comm. Math. Phys.* **125**, 337–352.

[60] R. Lyons (1990). Random walks and percolation on trees. *Ann. Probab.* **18**, 931–958.

[61] R. Lyons (1992). Random walks, capacity and percolation on trees. *Ann. Probab.* **20**, 2043–2088.

[62] R. Lyons (1995). Random walks and the growth of groups. *C.R. Acad. Sci. Paris* **320**, 1361–1366.

[63] R. Lyons and R. Pemantle (1992). Random walks in a random environment and first-passage percolation on trees. *Ann. Probab.* **20**, 125–136.

[64] R. Lyons, R. Pemantle and Y. Peres (1995). Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. *Ann. Probab.* **23**, 1125–1138.

[65] R. Lyons, R. Pemantle and Y. Peres (1996). Random walks on the lamplighter group, *Ann. Probab.* **24**, 1993–2006.

[66] R. Lyons and Y. Peres (1999). *Probability on Trees and Networks*, Cambridge University Press, in preparation. Current version available at `http://php.indiana.edu/~rdlyons/`.

[67] T. Lyons (1983) A simple criterion for transience of a reversible Markov chain, *Ann. Probab.* **11**, 393–402.

[68] P. Marchal (1998). The Best Bounds in a Theorem of Russell Lyons. *Elect. Comm. Probab.* **3**, 91–94. `http://www.math.washington.edu/~ejpecp/`

[69] T. Moore and J. L. Snell, (1979). A branching process showing a phase transition. *J. Appl. Probab.* **16**, 252–260.

[70] C. St. J.A. Nash-Williams (1959). Random walks and electric currents in networks. *Proc. Cambridge Philos. Soc.* **55**, 181-194

[71] R. Pemantle (1995). Tree-indexed processes. *Stat. Sci.* **10**, 200–213.

[72] R. Pemantle and Y. Peres (1994). Domination between trees and applications to an explosion problem. *Ann. Probab.* **22**, 180–194.

[73] R. Pemantle and Y. Peres (1995a). Critical random walk in random environment on trees. *Ann. Probab.* **23**, 105–140.

[74] R. Pemantle and Y. Peres (1995b). Galton-Watson trees with the same mean have the same polar sets. *Ann. Probab.* **23**, 1102–1124.

[75] R. Pemantle and Y. Peres (1996). Recursions on general trees and the Ising model at critical temperatures. *Unpublished manuscript.*

[76] Y. Peres (1996). Intersection-equivalence of Brownian paths and certain branching processes. *Commun. Math. Phys.* **177**, 417–434.

[77] Y. Peres and J. Steif (1998). The number of infinite clusters in dynamical percolation. *Probab. Th. Rel. Fields* **111**, 141–165.

[78] A. Pisztora (1996). Surface order large deviations for Ising, Potts and percolation models. *Probab. Th. Rel. Fields* **104**, 427–466.

[79] C. J. Preston (1974), *Gibbs States on Countable Sets,* Cambridge Univ. Press.

[80] A. Sidorenko (1994). An analytic approach to extremal problems for graphs and hypergraphs. *Extremal problems for finite sets* (Visegrad, 1991), 423–455, Bolyai Soc. Math. Stud. **3**, Janos Bolyai Math. Soc., Budapest.

[81] P. M. Soardi (1994). *Potential Theory on Infinite Networks.* Springer LNM, Berlin.

[82] F. Spitzer (1975). Markov random fields on an infinite tree. *Ann. Probab.* **3**, 387–394.

[83] M. Steel (1989). Distributions in bicolored evolutionary trees. *Ph.D. Thesis*, Massey University, Palmerston North, New Zealand.

[84] M. Steel and M. Charleston (1995). Five surprising properties of parsimoniously colored trees. *Bull. Math. Biology* **57**, 367–375.

[85] N. Tongring (1988). Which sets contain multiple points of Brownian motion? *Math. Proc. Camb. Phil. Soc.* **103**, 181–187.

[86] I. Vajda (1989). *Theory of Statistical Inference and Information.* Kluwer Academic Publishers, Dordrecht.

[87] B. Virag (1998). The speed of simple random walk on graphs. *Preprint.*