1 Some Highlights

This chapter surveys a few of the highlights to be encountered in this book, mainly, Chapters 2, 3, 4, 5, 15, and 16. Several of the topics in the book do not appear at all here since they are not as suitable to a quick overview. Also, we concentrate in this overview on trees, since it is easiest to use them to illustrate many of our themes.

1.1 Graph Terminology

For later reference, we introduce in this section the basic notation and terminology for graphs. A ***graph*** is a pair $G = (\mathsf{V}, \mathsf{E})$, where V is a set of ***vertices*** and E is a symmetric irreflexive subset of $\mathsf{V} \times \mathsf{V}$, called the ***edge*** set. ***Irreflexive*** means that E contains no element of the form (x, x). The word ***symmetric*** means that $(x, y) \in \mathsf{E}$ iff $(y, x) \in \mathsf{E}$; here, x and y are called the ***endpoints*** of (x, y). The symmetry assumption is usually phrased by saying that the graph is ***undirected*** or that its edges are ***unoriented***. Without this symmetry assumption, the graph is called ***directed***. If we need to distinguish the two, we write an unoriented edge as $[x, y]$, whereas an oriented edge is written as $\langle x, y \rangle$. An unoriented edge can be thought of as the pair of oriented edges with the same endpoints. If $(x, y) \in \mathsf{E}$, then we call x and y ***adjacent*** or ***neighbors***, and we write $x \sim y$. The ***degree*** of a vertex is the number of its neighbors. If this is finite for each vertex, we call the graph ***locally finite***. If the degree of every vertex is the same number d, then the graph is called ***regular*** or d***-regular***. If x is an endpoint of an edge e, then we also say that x and e are ***incident***, whereas if two edges share an endpoint, then we call those edges ***adjacent***. If we have more than one graph under consideration, we distinguish the vertex and edge sets by writing $\mathsf{V}(G)$ and $\mathsf{E}(G)$. A ***subgraph*** of a graph G is a graph whose vertex set is a subset of $\mathsf{V}(G)$ and whose edge set is a subset of $\mathsf{E}(G)$. One can define the product of two graphs $G_i = (\mathsf{V}_i, \mathsf{E}_i)$ $(i = 1, 2)$ in various ways. The one we use almost exclusively is the ***Cartesian product*** $G = (\mathsf{V}, \mathsf{E})$ with $\mathsf{V} := \mathsf{V}_1 \times \mathsf{V}_2$ and

$$\mathsf{E} := \Big\{ \big((x_1, x_2), (y_1, y_2) \big) \,;\ \big(x_1 = y_1,\ (x_2, y_2) \in \mathsf{E}_2 \big) \text{ or } \big((x_1, y_1) \in \mathsf{E}_1,\ x_2 = y_2 \big) \Big\} \,;$$

this product graph is denoted $G = G_1 \mathbin{\square} G_2$.

A ***path**** in a graph is a sequence of vertices where each successive pair of vertices is an edge in the graph; it is said to ***join*** its first and last vertices. When a path does not pass

* In graph theory, a path is necessarily self-avoiding. What we call a path is called in graph theory a ***walk***. However, to avoid confusion with random walks, we do not adopt that terminology.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

through any vertex (resp., edge) more than once, we will call it ***vertex simple*** (resp., ***edge simple***). We'll just say ***simple*** also to mean vertex simple, which implies edge simple. A finite path with at least one edge and whose first and last vertices are the same is called a ***cycle***. A cycle is called ***simple*** if no pair of vertices are the same except for its first and last ones. A graph is ***connected*** if, for each pair $x \neq y$ of its vertices, there is a path joining x to y. The ***distance*** between x and y is the minimum number of edges among all paths joining x and y, denoted either $d(x, y)$ or $\mathrm{dist}(x, y)$. A graph with no cycles is called a ***forest***; a connected forest is a ***tree***.

If there are numbers (weights) $c(e)$ assigned to the edges e of a graph, the resulting object is called a ***network***. Given a network $G = (\mathsf{V}, \mathsf{E})$ with weights $c(\bullet)$ and a subset K of its vertices, the ***induced subnetwork*** $G \restriction K$ is the subnetwork with vertex set K, edge set $(K \times K) \cap \mathsf{E}$, and weights $c \restriction \big((K \times K) \cap \mathsf{E}\big)$.

Sometimes we work with objects more general than graphs, called multigraphs. A ***multigraph*** is a pair of sets, V and E, together with a pair of maps $\mathsf{E} \to \mathsf{V}$, denoted $e \mapsto e^-$ and $e \mapsto e^+$. The images of e are called the ***endpoints*** of e, the former being its ***tail*** and the latter its ***head***. If $e^- = e^+ = x$, then e is a ***loop*** at x. Edges with the same set of endpoints are called ***parallel*** or ***multiple***. If the multigraph is undirected, then for every edge $e \in \mathsf{E}$, there is an edge $-e \in \mathsf{E}$ such that $(-e)^- = e^+$ and $(-e)^+ = e^-$. For a vertex x of an undirected multigraph, its ***degree*** is $|\{e\,;\ e^- = x\}|$. Sometimes we use ***paths*** of edges rather than of vertices; in this case, the head of each edge must equal the tail of the next edge. Given a subset $K \subseteq \mathsf{V}$, the multigraph G/K obtained by ***identifying*** K to a single vertex $z \notin \mathsf{V}$ is the multigraph whose vertex set is $(\mathsf{V} \setminus K) \cup \{z\}$ and whose edge set is obtained from E by replacing the tail and head maps so that every tail or head that took a value in K now takes the value z. A similar operation is ***contraction*** of an edge e, which is the result of first deleting e and then identifying e^- and e^+; we denote this graph by G/e. A multigraph that is a graph is called a ***simple graph***.

Let $G_1 = (\mathsf{V}_1, \mathsf{E}_1)$ and $G_2 = (\mathsf{V}_2, \mathsf{E}_2)$ be two (multi)graphs. A ***homomorphism*** of G_1 to G_2 is a map $\phi\colon G_1 \to G_2$ such that whenever x and e are incident in G_1, then so are $\phi(x)$ and $\phi(e)$ in G_2. When the graph is directed, then ϕ must also preserve orientation of edges, that is, if the head and tail of e are x and y, respectively, then the head and tail of $\phi(e)$ must be $\phi(x)$ and $\phi(y)$, respectively. If in addition, these graphs come with weight functions c_1 and c_2, so that they are networks, then a ***network homomorphism*** is a graph homomorphism ϕ that satisfies $c_1(e) = c_2(\phi(e))$ for all edges $e \in \mathsf{E}_1$. If ϕ induces bijections of V_1 to V_2 and of E_1 to E_2, then ϕ is called an ***isomorphism***. When $G_1 = G_2$, an isomorphism is called an ***automorphism***. A homomorphism $\phi\colon G_1 \to G_2$ extends to map each subset A of G_1 to a subset $\phi(A)$ of G_2 by mapping all elements of A by ϕ. We also extend ϕ to collections $\mathcal{A}$ of subsets of G_1 by applying ϕ to all elements of $\mathcal{A}$.

1.2 Branching Number

Our trees will usually be ***rooted***, meaning that some vertex is designated as the root, denoted o. We imagine the tree as growing (upward) away from its root. Each vertex then has branches leading to its children, which are its neighbors that are farther from the root. For the purposes of this chapter, we do not allow the possibility of leaves, that is, vertices without children.

How do we assign an average branching number to an arbitrary infinite locally finite tree? If the tree is a binary tree, as in Figure 1.1, then clearly the answer will be 2. But in the general case, since the tree is infinite, no straight average is available. We must take some kind of limit or use some other procedure, but we will be amply rewarded for our efforts.

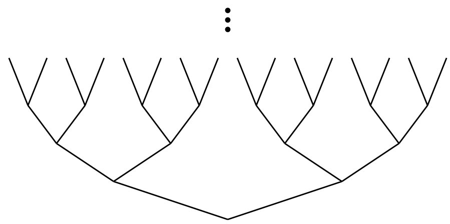

Figure 1.1. The binary tree.

One simple idea is as follows. Let T_n be the set of vertices at distance n from the root, o, called the ***nth level*** of T. Define the ***lower (exponential) growth rate*** of the tree to be

$$\underline{\mathrm{gr}}\, T := \liminf_{n\to\infty} |T_n|^{1/n} \,.$$

This certainly will give the number 2 to the binary tree. One can also define the ***upper (exponential) growth rate***

$$\overline{\mathrm{gr}}\, T := \limsup_{n\to\infty} |T_n|^{1/n}$$

and the ***(exponential) growth rate***

$$\mathrm{gr}\, T := \lim_{n\to\infty} |T_n|^{1/n}$$

when the limit exists. However, notice that these notions of growth barely account for the structure of the tree: only $|T_n|$ matters, not how the vertices at different levels are connected to each other. Of course, if T is ***spherically symmetric***, meaning that for each n, every vertex at distance n from the root has the same number of children (which may depend on n), then there is really no more information in the tree than that contained in the sequence $\langle |T_n|\,;\ n \ge 0\rangle$. For more general trees, however, we will use a different approach.

Consider the tree as a network of pipes and imagine water entering the network at the root. However much water enters a pipe leaves at the other end and splits up among the outgoing pipes (edges). Formally, this means that we consider a nonnegative function θ on the edges of T, called a ***flow***, with the property that for every vertex x other than the root, if x has parent z and children $y_1, \ldots, y_d$, then $\theta\big((z, x)\big) = \sum_{i=1}^{d} \theta\big((x, y_i)\big)$. We say that $\theta(e)$ is the amount of water flowing along e and that the total amount of water flowing from the root to infinity is $\sum_{j=1}^{k} \theta\big((o, x_j)\big)$, where the children of the root o are $x_1, \ldots, x_k$.

Consider the following sort of restriction on a flow: given $\lambda \ge 1$, suppose that the amount of water that can flow through an edge at distance n from o is only λ^{-n}. In other words, if $x \in T_n$ has parent z, then the restriction is that $\theta\big((z, x)\big) \le \lambda^{-n}$. If λ is too big, then perhaps no positive amount of water can flow from the root to infinity. Indeed, consider the binary tree. Then the ***equally splitting flow*** that sends an amount 2^{-n} through each edge at distance n from the root will satisfy the restriction imposed when $\lambda \le 2$ but not for any $\lambda > 2$. In fact, it is intuitively clear that there is no way to get any water to flow when $\lambda > 2$. Obviously, this critical value of 2 for λ is the same as the branching number of the binary tree – if the tree were ternary, then the critical value would be 3. So let us make a general definition: the

branching number of a tree T is the supremum of those λ that admit a positive total amount of water to flow through T; denote this critical value of λ by $\mathrm{br}\,T$.

Let's spend some time on this new concept. For a vertex x other than the root, let $e(x)$ denote the edge that joins x to its parent. The total amount of water flowing is, by definition, $\sum_{x\in T_1}\theta(e(x))$. If we apply the flow condition to each x in T_1, then we see that this sum also equals $\sum_{x\in T_2}\theta(e(x))$. Induction shows, in fact, that it equals $\sum_{x\in T_n}\theta(e(x))$ for every $n \geq 1$. When the flow is constrained in the way we have specified, then this sum is at most $\sum_{x\in T_n}\lambda^{-n} = |T_n|\lambda^{-n}$. Now if we choose $\lambda > \underline{\mathrm{gr}}\,T$, then $\liminf_{n\to\infty}|T_n|\lambda^{-n} = 0$, whence for such λ, no water can flow. Conclusion:

$$\mathrm{br}\,T \leq \underline{\mathrm{gr}}\,T\,. \tag{1.1}$$

Often, as in the case of the binary tree, equality holds here. However, there are many examples of strict inequality.

Before we give an example of strict inequality, here is another example where equality holds in (1.1).

Example 1.1. If T is a tree such that vertices at even distances from o have two children whereas the rest have three children, then $\mathrm{br}\,T = \mathrm{gr}\,T = \sqrt{6}$. Why? It is easy to see that $\mathrm{gr}\,T = \sqrt{6}$, whence by (1.1), it remains to show that $\mathrm{br}\,T \geq \sqrt{6}$. In other words, it remains to show that, given $\lambda < \sqrt{6}$, a positive amount of water can flow to infinity under the constraints described. Indeed, we can use the water flow with amount $6^{-n/2}$ flowing on those edges at distance n from the root when n is even and with amount $6^{-(n-1)/2}/3$ flowing on those edges at distance n from the root when n is odd.

More generally, one can show (Exercise 1.2) that equality holds in (1.1) whenever T is spherically symmetric.

Now we give an example where strict inequality holds in (1.1).

Example 1.2. (The 1–3 Tree) We will construct a tree T embedded in the upper half-plane with o at the origin. We'll have $|T_n| = 2^n$, but we'll connect them in a funny way. List T_n in clockwise order as $\langle x_1^n, \ldots, x_{2^n}^n\rangle$. Let x_k^n have one child if $k \leq 2^{n-1}$ and three children otherwise; see Figure 1.2. Define a ***ray*** in a tree to be an infinite path from the root that doesn't backtrack. If x is a vertex of T that does not have the form $x_{2^n}^n$, then there are only finitely many rays that pass through x. This means that water cannot flow to infinity through such a vertex x when $\lambda > 1$. That leaves only the possibility of water flowing along the single ray consisting of the vertices $x_{2^n}^n$, but that's impossible too. Hence $\mathrm{br}\,T = 1$, yet $\mathrm{gr}\,T = 2$.

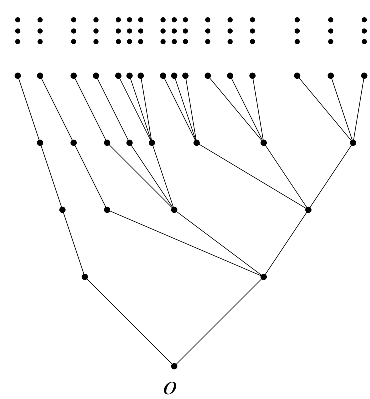

Figure 1.2. A tree with branching number 1 and growth rate 2.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

Example 1.3. If $T^{(1)}$ and $T^{(2)}$ are trees, form a new tree $T^{(1)} \vee T^{(2)}$ from disjoint copies of $T^{(1)}$ and $T^{(2)}$ by joining their roots to a new point taken as the root of $T^{(1)} \vee T^{(2)}$ (Figure 1.3). Then

$$\operatorname{br}\left(T^{(1)} \vee T^{(2)}\right) = \operatorname{br} T^{(1)} \vee \operatorname{br} T^{(2)}$$

since water can flow in the join $T^{(1)} \vee T^{(2)}$ iff water can flow in one of the trees. Here, as usual in probability, we use $a \vee b$ to mean $\max\{a, b\}$ when a and b are real numbers.

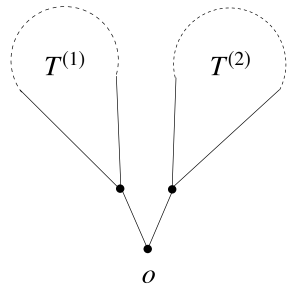

Figure 1.3. Joining two trees.

Although $\underline{\operatorname{gr}}\, T$ is easy to compute, $\operatorname{br} T$ may not be. Nevertheless, it is the branching number that is much more important. Theorems to be described shortly will bear out this assertion. We will develop tools to compute $\operatorname{br} T$ in many common situations.

1.3 Electric Current

We can ask another flow question on trees, this one concerning electrical current. All electrical terms are given precise mathematical definitions in Chapter 2, but for now, we give some bare definitions to sketch the arc of some of the fascinating and surprising connections that we develop later. If positive numbers $c(e)$ are assigned to the edges e of a tree, we may call these numbers ***conductances***, and in that case, the ***energy*** of a flow θ is defined to be $\sum_e \theta(e)^2/c(e)$. We say that electrical ***current flows*** from the root to infinity if there is a nonzero flow with finite energy.

Here's our new flow question: if λ^{-n} is the conductance of edges at distance n from the root of T, will current flow?

Example 1.4. Consider the binary tree. The equally splitting flow has finite energy for every $\lambda < 2$, so in those cases, electrical current does flow. One can show that when $\lambda \geq 2$, not only does the equally splitting flow have infinite energy, but so does every nonzero flow (Exercise 1.4). Thus, current flows in the infinite binary tree iff $\lambda < 2$. Note the slight difference to water flow: when $\lambda = 2$, water can still flow on the binary tree.

In general, there will be a critical value of λ below which current flows and above which it does not. In the example of the binary tree that we just analyzed, this critical value was the same as that for water flow. Is this equality special to nice trees, or does it hold for all trees? We have seen an example of a strange tree (another is in Exercise 1.3), so we might doubt its generality. However, it is indeed a general fact (Lyons, 1990):

Theorem 1.5.* *If $\lambda < \operatorname{br} T$, then electrical current flows, but if $\lambda > \operatorname{br} T$, then it does not.*

---

* This will follow from Theorem 3.5 and the discussion in Section 2.2.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

1.4 Random Walks

There is a striking, but easily established, correspondence between electrical networks and random walks on graphs (or on networks). Namely, given a finite connected graph G with conductances (that is, positive numbers) assigned to the edges, we consider the random walk that can go from a vertex only to an adjacent vertex and whose transition probabilities from a vertex are proportional to the conductances along the edges to be taken. That is, if x is a vertex with neighbors $y_1, \ldots, y_d$ and the conductance of the edge (x, y_i) is c_i, then the transition probability from x to y_j is $p(x, y_j) := c_j / \sum_{i=1}^d c_i$. Now consider two fixed vertices a_0 and a_1 of G. A ***voltage function*** on the vertices is then a function v such that $v(a_i) = i$ for $i = 0, 1$ and for every other vertex $x \neq a_0, a_1$, the equation $v(x) \sum_{i=1}^d c_i = \sum_{i=1}^d c_i v(y_i)$ holds, where the neighbors of x are $y_1, \ldots, y_d$. In other words, $v(x)$ is a weighted average of the values at the neighbors of x. We will see in Section 2.1 that voltage functions exist and are unique. The following proposition provides the basic connection between random walks and electrical networks:

Proposition 1.6. (Voltage as Probability) *For every vertex x, the voltage at x equals the probability that when the corresponding random walk starts at x, it will visit a_1 before it visits a_0.*

The proof of this proposition will be simple: In outline, there is a discrete Laplacian (a difference operator) that will define a notion of harmonic function. Both the voltage and the probability mentioned are harmonic functions of x. The two functions clearly have the same values at a_i (the "boundary" points), and the uniqueness principle holds for this Laplacian, whence the functions agree at all vertices x. This is developed in detail in Section 2.1.

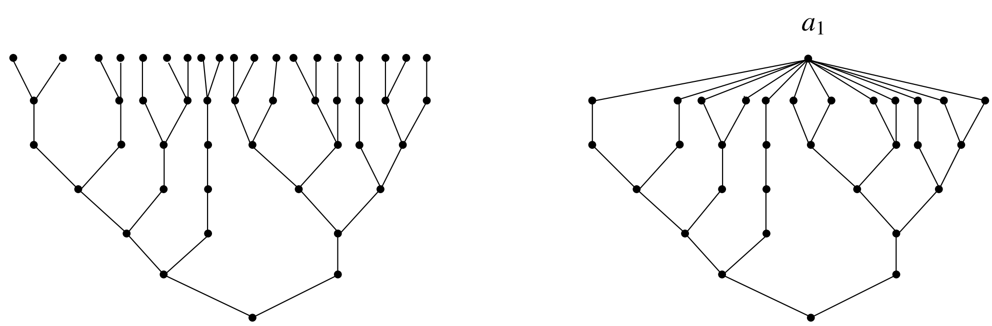

Figure 1.4. Identifying a level to a vertex, a_1.

What does this say about our trees? Given N, identify all the vertices of level N, that is, T_N, to one vertex, a_1 (see Figure 1.4). Use the root as a_0. Then, according to Proposition 1.6, the voltage at x is the probability that the random walk visits level N before it visits the root when it starts from x. When $N \to \infty$, the limiting voltages are all 0 iff the limiting probabilities are all 0, which is the same thing as saying that on the infinite tree, the probability of visiting the root from any vertex is 1, in other words, the random walk is recurrent. Although we have not yet defined "current," we'll see that no current flows across edges whose endpoints have the same voltage. This will imply, then, that no electrical current flows iff the random walk is recurrent. Contrapositively, electrical current flows iff the random walk is transient. In this

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

way, electrical networks will be a powerful tool to help us decide whether a random walk is recurrent or transient. These ideas are detailed in Section 2.2.

Earlier we considered conductances λ^{-n} on edges at distance n from the root. In this case, the conductances decrease by a factor of λ as the distance increases by 1, so the relative weights at a vertex other than the root are as shown in Figure 1.5. That is, the edge leading back toward the root is λ times as likely to be taken as each edge leading away from the root. Denoting the dependence of the random walk on the parameter λ by RW_λ, we may translate Theorem 1.5 into a probabilistic form (Lyons, 1990):

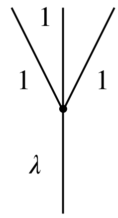

Figure 1.5. The relative weights at a vertex. The tree is growing upwards.

Theorem 1.7.* *If $\lambda < \operatorname{br} T$, then RW_λ is transient, whereas if $\lambda > \operatorname{br} T$, then RW_λ is recurrent.*

Is this form intuitive? Consider a vertex other than the root with, say, d children. If we consider only the distance from o, which increases or decreases at each step of the random walk, a balance at this vertex between increasing and decreasing occurs when $\lambda = d$. If d is constant, then the distance from the root undergoes a random walk with a constant bias (for a fixed λ), so it is easy to see that indeed d is the critical value separating transience from recurrence. What Theorem 1.7 says is that this same heuristic can be used in the general case, provided we substitute the "average" $\operatorname{br} T$ for d.

We will also see how to use electrical networks to prove Pólya's wonderful, seminal theorem that simple random walk on the hypercubic lattice $\mathbb{Z}^d$ is recurrent for $d \leq 2$ and transient for $d \geq 3$.

1.5 Percolation

Suppose that we remove edges at random from a tree, T. To be specific, we keep each edge with some fixed probability p and make these decisions independently for different edges. This random process is called ***percolation***. As we'll see, by Kolmogorov's zero-one law, the probability that an infinite connected component remains in the tree is either 0 or 1. On the other hand, we'll see that this probability is monotonic in p, whence there is a critical value $p_c(T)$ where it changes from 0 to 1. It is also intuitively clear that the "bigger" the tree, the more likely it is that there will be an infinite component for a given p. That is, the "bigger" the tree, the smaller is the critical value p_c. Thus, p_c is vaguely inversely related to a notion of average size or maybe average branching number. Surprisingly, this vague heuristic is precise and general (Lyons, 1990):

* This will be proved as Theorem 3.5.

Theorem 1.8.* *For any tree,* $p_c(T) = 1/\operatorname{br} T$.

What is this telling us? If a vertex x has d children, then the expected number of children remaining after percolation is dp. If dp is "usually" less than 1, then one would not expect that an infinite component would remain, whereas if dp is "usually" greater than 1, then one might guess that an infinite component would be present somewhere. Theorem 1.8 says that this intuition becomes correct when one replaces the "usual" d by $\operatorname{br} T$. Both Theorems 1.5 and 1.8 say that the branching number of a tree is a single number that captures enough of the complexity of a general tree to give the critical value for a stochastic process on the tree. There are other examples as well of this striking phenomenon. Altogether, they make a convincing case that the branching number is indeed the most important single number to attach to an infinite tree.

1.6 Branching Processes

In the preceding section, we looked at existence of infinite components after percolation on a tree. Although this event has probability 0 or 1, if we restrict attention to the connected component of the root, its probability of being infinite is between 0 and 1. These are equivalent ways to approach the issue, since, as we'll see, there is an infinite component somewhere with probability 1 iff the component of the root is infinite with positive probability. But looking at the component of the root also suggests a different stochastic process.

Percolation on a fixed tree produces random trees by random pruning, but there is a way to *grow* trees randomly that was invented by Bienaymé in 1845. Given probabilities p_k adding to 1 ($k = 0, 1, 2, \ldots$), we begin with one individual, and let it reproduce according to these probabilities, that is, it has k children with probability p_k. Each of these children (if there are any) then reproduce independently with the same law, and so on forever or until some generation goes extinct. The family trees produced by such a process are called ***(Bienaymé-)Galton-Watson trees***. A fundamental theorem in the subject (Proposition 5.4) is that extinction is a.s. iff $m \le 1$ and $p_1 < 1$, where $m := \sum_k k p_k$ is the mean number of offspring per individual. This provides further justification for our intuitive understanding of Theorem 1.8. It also raises a natural question: Given that a Galton-Watson family tree is nonextinct (infinite), what is its branching number? All the intuition suggests that it is m a.s., and indeed it is. This was first proved by Hawkes (1981). But here is the idea of a very simple proof (Lyons, 1990).

According to Theorem 1.8, to determine $\operatorname{br} T$, we may determine $p_c(T)$. Thus, let T grow according to a Galton-Watson process, then perform percolation on T, that is, keep edges with probability p. Focus on the component of the root. Looked at as a random tree in itself, this component appears simply as some other Galton-Watson tree; its mean is mp by independence of the growing and the "pruning" (percolation). Hence, the component of the root is infinite with positive probability iff $mp > 1$. This implies that $p_c = 1/m$ a.s. on nonextinction, thus $\operatorname{br} T = m$ a.s. on nonextinction. We'll flesh out the details when we prove Proposition 5.9.

* This will be proved as Theorem 5.15.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

Now let's consider another way to measure the size of Galton-Watson trees. Let Z_n be the size of the nth generation in a Galton-Watson process. How quickly does Z_n grow? It will be easy to calculate that $\mathbf{E}[Z_n] = m^n$. Moreover, a martingale argument will show that the limit $W := \lim_{n\to\infty} Z_n/m^n$ always exists (and is finite). When $1 < m < \infty$, do we have that $W > 0$ a.s. on the event of nonextinction? When $W > 0$, the growth rate of the tree is asymptotically Wm^n; this implies the cruder asymptotic $\operatorname{gr} T = m$. It turns out that indeed $W > 0$ a.s. on the event of nonextinction, under a very mild hypothesis:

The Kesten-Stigum Theorem (1966). *When $1 < m < \infty$, the following are equivalent:*

(i) $W > 0$ *a.s. on the event of nonextinction;*
(ii) $\sum_{k=1}^{\infty} p_k k \log k < \infty$.

This will be shown in Section 12.2. Although condition (ii) appears technical and suggests some possibly unpleasant analysis, we will enjoy a conceptual proof of the theorem that uses only extremely simple estimates.

1.7 Random Spanning Trees

The fertile and fascinating field of random spanning trees is one of the oldest areas to be studied in this book but one of the newest to be explored in depth. A ***spanning tree*** of a (connected) graph is a subgraph that is connected, contains every vertex of the whole graph, and contains no cycle: see Figure 1.6 for an example. These trees are usually not rooted. The subject of random spanning trees of a graph goes back to Kirchhoff (1847), who showed its surprising relation to electrical networks. (Actually, Kirchhoff did not think probabilistically; rather, he considered quotients of the number of spanning trees with a certain property divided by the total number of spanning trees. See Kirchhoff's effective resistance formula in Section 4.2 and Exercise 4.30.) One of Kirchhoff's results expresses the probability that a uniformly chosen spanning tree will contain a given edge in terms of electrical current in the graph.

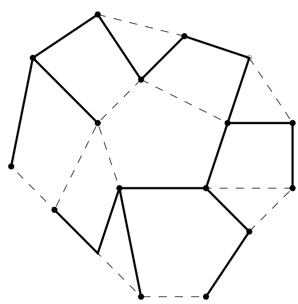

Figure 1.6. A spanning tree in a graph, where the edges of the graph not in the tree are dashed.

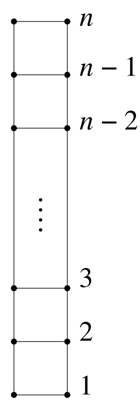

Figure 1.7. A ladder graph.

To get our feet wet, let's begin with a very simple finite graph. Namely, consider the ladder graph of Figure 1.7. Among all spanning trees of this graph, what proportion contain the bottom rung (edge)? In other words, if we were

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

to choose uniformly at random a spanning tree, what is the chance that it would contain the bottom rung? We have illustrated in Figure 1.8 the entire probability spaces for the smallest ladder graphs.

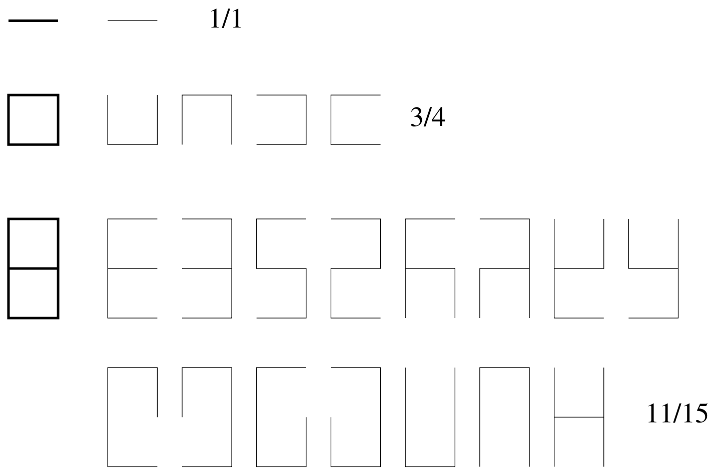

Figure 1.8. The ladder graphs of heights 0, 1, and 2, together with their spanning trees.

As shown, the probabilities in these cases are $1/1$, $3/4$, and $11/15$. The next one is $41/56$. Do you see any pattern? One thing that is fairly evident is that these numbers are decreasing but hardly changing. Amusingly, these numbers are every other term of the continued fraction expansion of $\sqrt{3} - 1 = 0.73^+$ and, in particular, converge to $\sqrt{3} - 1$. In the limit, then, the probability of using the bottom rung is $\sqrt{3} - 1$, and even before taking the limit, this gives an excellent approximation to the probability. How can we easily calculate such numbers? In this case, there is a rather easy recursion to set up and solve, but we will use this example to illustrate the more general theorem of Kirchhoff that we mentioned earlier. In fact, Kirchhoff's theorem will show us why these probabilities are decreasing even before we calculate them.

For the next two paragraphs, we will assume some familiarity with electrical networks; those who do not know these terms will find precise mathematical definitions in Sections 2.1 and 2.2. Suppose that each edge of our graph (any graph – say, the ladder graph) is an electric conductor of unit conductance. Hook up a battery between the endpoints of any edge e – say, the bottom rung (Figure 1.9). Kirchhoff (1847) showed that the proportion of current that flows directly along e is then equal to the probability that e belongs to a randomly chosen spanning tree!

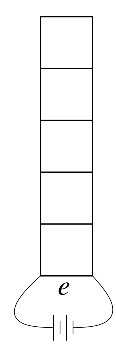

Figure 1.9. A battery is hooked up between the endpoints of e.

Now current flows in two ways: some flows directly across e and some flows through the rest of the network. It is intuitively clear (and justified by Rayleigh's monotonicity principle in Section 2.4) that the higher the ladder, the greater the effective conductance of the ladder minus the bottom rung, hence the less current proportionally will

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

flow along e, whence by Kirchhoff's theorem, the less the chance that a random spanning tree contains the bottom rung. This confirms our observations.

It turns out that generating spanning trees at random according to the uniform measure is of interest to computer scientists, who have developed various algorithms over the years for random generation of spanning trees. In particular, this is closely connected to generating a random state from any Markov chain. See Propp and Wilson (1998) for more on this issue.

Early algorithms for generating a random spanning tree used the matrix-tree theorem, which counts the number of spanning trees in a graph via a determinant (Section 4.4). A better algorithm than these early ones, especially for probabilists, was introduced by Aldous (1990) and Broder (1989). It says that if you start a simple random walk at *any* vertex of a finite (connected) graph G and draw every edge it traverses except when it would complete a cycle (that is, except when it arrives at a previously visited vertex), then when no more edges can be added without creating a cycle, what will be drawn – amazingly – is a uniformly chosen spanning tree of G. (To be precise: if X_n $(n \geq 0)$ is the path of the random walk, then the associated spanning tree is the set of edges $\big\{[X_n, X_{n+1}]\,;\ X_{n+1} \notin \{X_0, X_1, \ldots, X_n\}\big\}$.) This beautiful algorithm is quite efficient and useful for theoretical analysis, yet as a graduate student, Wilson (1996) found an even better one that we'll describe in Section 4.1.

Return for a moment to the ladder graphs. We saw that as the height of the ladder tends to infinity, there is a limiting probability that the bottom rung of the ladder graph belongs to a uniform spanning tree. What about uniform spanning trees in other sequences of growing finite graphs? Suppose that G is an infinite graph. Let G_n be finite (connected) subgraphs with $G_1 \subset G_2 \subset G_3 \subset \cdots$ and $\bigcup G_n = G$. Take the uniform spanning tree probability measure on each G_n. This gives a sequence of probability measures on subsets of edges of G. Does this sequence converge in a reasonable sense? Lyons conjectured that it does, and Pemantle (1991) verified that the weak limit exists. (In other words, if μ_n denotes the uniform spanning tree measure on G_n and B, B' are finite sets of edges in G, then $\lim_n \mu_n[B \subset T_n,\ B' \cap T_n = \varnothing]$ exists, where T_n denotes a random spanning tree in G_n.) This limit is now called the ***free uniform spanning forest**** on G, denoted FUSF or just FSF. Considerations of electrical networks play the dominant role in Pemantle's proof. Pemantle (1991) discovered the astounding fact that on $\mathbb{Z}^d$, the uniform spanning forest is a single tree a.s. if $d \leq 4$, but when $d \geq 5$, there are infinitely many trees a.s.! We'll prove this as Theorem 10.30.

* In graph theory, "spanning forest" usually means a maximal subgraph without cycles, that is, a spanning tree in each connected component. We mean, instead, a subgraph without cycles that contains every vertex.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

1.8 Hausdorff Dimension

We've used water flow on trees to define the branching number, where the amount of water that can flow through an edge at distance n from the root of a tree is constrained to be at most λ^{-n}. There is a useful way to reformulate this via what's known as the max-flow min-cut theorem, proved in Section 3.1. Namely, consider a set Π of edges whose removal leaves the root o in a finite component. We call such a set a ***cutset*** (separating o from infinity). If θ is a flow from o to infinity, then all the water must flow through Π, so one expects that an upper bound on the total that can flow is $\sum_{e(x)\in\Pi} \lambda^{-|x|}$, where $e(x)$ denotes the edge that joins x to its parent, as before, and $|x|$ denotes the distance of a vertex x to the root. This expectation turns out to be correct, so that the most that can flow is

$$\inf\left\{ \sum_{e(x)\in\Pi} \lambda^{-|x|}\,;\ \Pi \text{ is a cutset} \right\}. \tag{1.2}$$

Remarkably, this upper bound is always achievable, that is, there is a flow with this amount in total flowing from the root to infinity; this is the content (in a special case) of the max-flow min-cut theorem. We are going to use this now to understand Hausdorff dimension, but a much more detailed and varied motivation of Hausdorff dimension is given in Chapter 15.

A vertex of degree 1 in a tree is called a ***leaf***. By analogy with the leaves of a finite tree, we call the set of rays of T the ***boundary*** (at infinity) of T, denoted ∂T. (Recall that a ray is an infinite simple path from the root, so ∂T does not include any leaves of T.) Now there is a natural metric on ∂T: if $\xi, \eta \in \partial T$ have exactly n edges in common, define their distance to be $d(\xi, \eta) := e^{-n}$. Thus, if $x \in T$ has more than one child with infinitely many descendants, then the set of rays going through x,

$$B_x := \{\xi \in \partial T\,;\ \xi_{|x|} = x\}, \tag{1.3}$$

has diameter $\operatorname{diam} B_x = e^{-|x|}$. We call a collection $\mathscr{C}$ of subsets of ∂T a ***cover*** if

$$\bigcup_{B\in\mathscr{C}} B = \partial T\,.$$

▷ **Exercise 1.1.**

Let T be an infinite locally finite tree.

(a) **(Kőnig's Lemma)** Show that $\partial T \neq \varnothing$.

(b) Show that ∂T is compact.

Note that

$$\Pi \text{ is a cutset (separating } o \text{ from } \infty) \text{ iff } \{B_x\,;\ e(x) \in \Pi\} \text{ is a cover.} \tag{1.4}$$

The ***Hausdorff dimension*** of ∂T is defined to be

$$\dim \partial T := \sup\left\{ \alpha\,;\ \inf_{\mathscr{C} \text{ a countable cover}} \sum_{B\in\mathscr{C}} (\operatorname{diam} B)^\alpha > 0 \right\}.$$

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

This number is just a disguised version of the branching number.* Indeed,

$$\operatorname{br} T = \sup\left\{\lambda\,;\ \text{water can flow through pipe capacities } \lambda^{-|x|}\right\}.$$

Now use the condition (1.2) to write this as

$$\sup\left\{\lambda\,;\ \inf_{\Pi \text{ a cutset}} \sum_{e(x)\in\Pi} \lambda^{-|x|} > 0\right\}.$$

Replace λ by e^{α} to rewrite it as

$$\exp\sup\left\{\alpha\,;\ \inf_{\Pi \text{ a cutset}} \sum_{e(x)\in\Pi} e^{-\alpha|x|} > 0\right\},$$

and then use the correspondence (1.4) between cutsets and covers to write this as

$$\exp\sup\left\{\alpha\,;\ \inf_{\mathscr{C} \text{ a cover}} \sum_{B\in\mathscr{C}} (\operatorname{diam} B)^{\alpha} > 0\right\}.$$

Now we see the disguise revealed as

$$\operatorname{br} T = \exp\dim\partial T\,.$$

Soon we'll see how this helps us to analyze Hausdorff dimension in Euclidean space.

1.9 Capacity

In Section 1.3, we made the definition

$$\begin{gathered}\text{electrical current flows from the root of an infinite tree}\\ \iff \\ \text{there is a flow with finite energy.}\end{gathered} \tag{1.5}$$

A unit flow on a tree T from the root to infinity is a flow where a total of 1 unit flows from the root. By identifying vertices x with edges $e(x)$, we may identify a unit flow with a function θ on the vertices of T that is 1 at the root and that has the property that for all vertices x,

$$\theta(x) = \sum_i \theta(y_i)\,,$$

where y_i are the children of x. The energy of a flow for the conductances that we've been using as our basic example is then

$$\sum_{x\in T} \theta(x)^2 \lambda^{|x|}\,,$$

* Historically, the branching number was defined by Lyons (1990) only after Furstenberg (1970) considered the Hausdorff dimension of the boundaries of trees, which served as the former's inspiration.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

whence we may write Theorem 1.5 as

$$\operatorname{br} T = \sup\left\{\lambda\,;\ \text{there exists a unit flow } \theta \quad \sum_{x\in T} \theta(x)^2 \lambda^{|x|} < \infty\right\}. \tag{1.6}$$

We can also identify unit flows θ on T with Borel probability measures μ on ∂T via

$$\mu(B_x) = \theta(x)$$

(see Section 15.4). A bit of algebra (Proposition 16.1) will show that (1.6) is equivalent to

$$\operatorname{br} T = \exp\sup\left\{\alpha\,;\ \exists \text{ a probability measure } \mu \text{ on } \partial T \quad \iint \frac{d\mu(\xi)\,d\mu(\eta)}{d(\xi,\eta)^\alpha} < \infty\right\}.$$

For $\alpha > 0$, define the ***α-capacity*** of ∂T to be the reciprocal of the minimum energy of a unit flow for $\lambda = e^\alpha$. When we express this purely in terms of probability measures on the boundary, this will turn out to be the same as the following definition:

$$\operatorname{cap}_\alpha(\partial T)^{-1} := \inf\left\{\iint \frac{d\mu(\xi)\,d\mu(\eta)}{d(\xi,\eta)^\alpha}\,;\ \mu \text{ a probability measure on } \partial T\right\}.$$

Then statement (1.5) says that for $\alpha > 0$,

$$\text{random walk with parameter } \lambda = e^\alpha \text{ is transient} \iff \operatorname{cap}_\alpha(\partial T) > 0\,. \tag{1.7}$$

It follows from Theorem 1.7 that

$$\text{the critical value of } \alpha \text{ for positivity of } \operatorname{cap}_\alpha(\partial T) \text{ is } \dim \partial T. \tag{1.8}$$

Theorem 1.8 told us that these same critical values for random walk, electrical networks, Hausdorff dimension, or capacity are also critical for percolation. But it did not tell us what happens at the critical value, unlike, say, (1.7) does for random walk. This is more subtle to analyze for percolation but is also known (Lyons, 1992):

Theorem 1.9.* (Tree Percolation and Capacity) *For $\alpha > 0$, percolation with parameter $p = e^{-\alpha}$ yields an infinite component a.s. iff $\operatorname{cap}_\alpha(\partial T) > 0$. Moreover,*

$$\operatorname{cap}_\alpha(\partial T) \le \mathbf{P}[\textit{the component of the root is infinite}] \le 2\operatorname{cap}_\alpha(\partial T)\,.$$

Although this appears rather abstract, it is very useful. First of all, when T is spherically symmetric and $p = e^{-\alpha}$, we can calculate the capacities easily (Exercise 16.1):

$$\operatorname{cap}_\alpha(\partial T) = \left(1 + (1-p)\sum_{n=1}^{\infty} \frac{1}{p^n |T_n|}\right)^{-1}.$$

Second, one can use this theorem in combination with (1.7); this allows us to translate problems freely between the domains of random walks and percolation (Lyons, 1992). Third, we describe how it can be used to analyze Brownian motion in the next section.

* This will be proved as Theorem 16.3. The case of the first part of this theorem where all the degrees are uniformly bounded was shown earlier by Fan (1989, 1990).

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press

1.10 Embedding Trees into Euclidean Space

The results described previously, especially those concerning percolation, can be translated to give interesting results on closed sets in Euclidean space. We describe only the simplest such correspondence here.*

Let $b \geq 2$ be an integer. An interval of the form $[k/b^n, (k+1)/b^n]$ for integers k and n is called ***b-adic of order*** n. For a closed nonempty set $E \subseteq [0,1]$, consider the system of b-adic subintervals of $[0,1]$. We'll associate a tree to E as follows: Those intervals whose intersection with E is nonempty will form the vertices of the tree. Two such intervals are connected by an edge iff one contains the other and the ratio of their lengths is b. The root of this tree is $[0,1]$. Denote this tree by $T_{[b]}(E)$. An example is illustrated in Figure 1.10 with $b = 4$. Were it not for the fact that certain numbers have two representations in base b, we could identify $\partial T_{[b]}(E)$ with E. Because of this multiplicity of representation, there are other trees whose boundary we could identify with E. That is, given a tree T, suppose that we associate to each $x \in T_n$ a b-adic interval $I_x \subseteq [0,1]$ of order n in such a way that $|I_x \cap I_y| \leq 1$ for $|x| = |y|$, $x \neq y$, and that I_x is contained in I_z when z is the parent of x. Then the tree T ***codes*** the closed set $E := \bigcap_{n \geq 0} \bigcup_{x \in T_n} I_x$. The difference between ∂T and $\partial T_{[b]}(E)$ is at most countable. As we will see, this implies, for example, that these two boundaries have the same Hausdorff dimension.

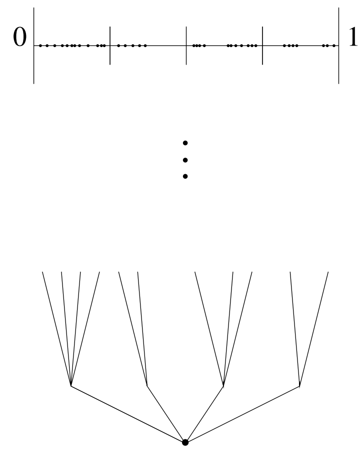

Figure 1.10. Coding by trees.

Hausdorff dimension is defined for subsets of $[0,1]$ just as we defined it for ∂T: A ***cover*** of E is a collection $\mathscr{C}$ of sets whose union contains E, and

$$\dim E := \sup\left\{\alpha\,;\ \inf_{\mathscr{C}\text{ a cover of }E} \sum_{B \in \mathscr{C}} (\operatorname{diam} B)^\alpha > 0\right\},$$

where diam B denotes the (Euclidean) diameter of E. When T codes E, covers of ∂T by sets of the form B_x (as in (1.3)) correspond to covers of E by b-adic intervals, but of diameter $b^{-|x|}$, rather than $e^{-|x|}$. One can show that restricting to covers of E by b-adic intervals does not change the computation of Hausdorff dimension, whence we may conclude (compare the calculation at the end of Section 1.8) that

$$\dim E = \frac{\dim \partial T}{\log b} = \log_b(\operatorname{br} T)\,. \tag{1.9}$$

Example 1.10. Let E be the Cantor middle-thirds set. If $b = 3$, then the binary tree codes E (when the obvious 3-adic intervals are associated to the binary tree), whence (1.9) tells us

* This correspondence was part of Furstenberg's motivation in 1970 for looking at the dimension of the boundary of a tree.

that the Hausdorff dimension of E is $\log_3 2 = \log 2/\log 3$. If we use a different base, b, to code E by a tree T, we will have $\operatorname{br} T = b^{\log_3 2}$.

Capacity in Euclidean space is also defined as we defined it on the boundary of a tree:

$$(\operatorname{cap}_\alpha E)^{-1} := \inf \left\{ \iint \frac{d\mu(x)\, d\mu(y)}{|x-y|^\alpha} \,;\ \mu \text{ a probability measure on } E \right\}.$$

It was shown by Benjamini and Peres (1992) and Pemantle and Peres (1995b) (see Section 16.3) that when T codes E,

$$\frac{1}{2}\operatorname{cap}_\alpha E \le \frac{1}{1-b^{-\alpha}} \operatorname{cap}_{\alpha \log b} \partial T \le 3b \operatorname{cap}_\alpha E\,. \tag{1.10}$$

This means that the percolation criterion Theorem 1.9 can be used in Euclidean space. This, and similar extensions, will allow us in Section 16.4 to analyze Brownian motion in $\mathbb{R}^d$ by replacing the path of Brownian motion by an "intersection-equivalent" random fractal that is much easier to analyze, being an embedding of a Galton-Watson tree. This will allow us to determine whether Brownian motion has double points, triple points, etc., in a very easy fashion.

1.11 Notes

The product of two graphs, G_1 and G_2, with $\mathsf{V} := \mathsf{V}_1 \times \mathsf{V}_2$ and the choice

$$\mathsf{E} := \left\{ \big((x_1, x_2), (y_1, y_2)\big) \,;\ (x_1, y_1) \in \mathsf{E}_1 \text{ and } (x_2, y_2) \in \mathsf{E}_2 \right\}$$

is called the ***tensor product***, since its adjacency matrix is the tensor product of the adjacency matrices corresponding to E_1 and E_2. It is denoted $G = G_1 \times G_2$. The union of $G_1 \,\square\, G_2$ and $G_1 \times G_2$ is denoted $G_1 \boxtimes G_2$. Terminology for graph products is not universal; other terms include "sum" for what we called the Cartesian product and "product" for the tensor product.

Other recent books that cover material related to the topics of this book include *Probability on Graphs* by Geoffrey Grimmett, *Reversible Markov Chains and Random Walks on Graphs* by David Aldous and Jim Fill (preliminary version online), *Coarse Geometry and Randomness* by Itai Benjamini, *Markov Chains and Mixing Times* by David A. Levin, Yuval Peres, and Elizabeth L. Wilmer, *Probability: The Classical Limit Theorems* by Henry McKean, *Random Trees: An Interplay between Combinatorics and Probability* by Michael Drmota, *A Course on the Web Graph* by Anthony Bonato, *Random Graph Dynamics* by Rick Durrett, *Complex Graphs and Networks* by Fan Chung and Linyuan Lu, *The Random-Cluster Model* by Geoffrey Grimmett, *Superfractals* by Michael Fielding Barnsley, *Introduction to Mathematical Methods in Bioinformatics* by Alexander Isaev, *Gaussian Markov Random Fields* by Håvard Rue and Leonhard Held, *Conformally Invariant Processes in the Plane* by Gregory F. Lawler, *Random Networks in Communication* by Massimo Franceschetti and Ronald Meester, *Percolation* by Béla Bollobás and Oliver Riordan, *Probability and Real Trees* by Steven Evans, *Random Trees, Lévy Processes and Spatial Branching Processes* by Thomas Duquesne and Jean-François Le Gall, *Combinatorial Stochastic Processes* by Jim Pitman, *Random Geometric Graphs* by Mathew Penrose, *Random Graphs* by Béla Bollobás, *Random Graphs* by Svante Janson, Tomasz Luczak, and Andrzej Ruciński, *Phylogenetics* by Charles Semple and Mike Steel, *Stochastic Networks and Queues* by Philippe Robert, *Random Walks on Infinite Graphs and Groups* by Wolfgang Woess, *Random Walk: A Modern Introduction* by Gregory F. Lawler and Vlada Limic, *Percolation* by Geoffrey Grimmett, *Noise Sensitivity of Boolean Functions and Percolation* by Christophe Garban and Jeffrey E. Steif, *Stochastic Interacting Systems: Contact, Voter and Exclusion Processes* by Thomas M. Liggett, and *Discrete Groups, Expanding Graphs and Invariant Measures* by Alexander Lubotzky.

1.12 Collected In-Text Exercises

1.1. Let T be an infinite locally finite tree.

(a) **(Kőnig's Lemma)** Show that $\partial T \neq \varnothing$.

(b) Show that ∂T is compact.

1.13 Additional Exercises

1.2. Show that $\operatorname{br} T = \underline{\operatorname{gr}}\, T$ when T is a spherically symmetric tree.

1.3. Here we'll look more closely at the joining construction of Example 1.3. We will put together two spherically symmetric trees $T^{(1)}$ and $T^{(2)}$ such that $\operatorname{br}\big(T^{(1)} \vee T^{(2)}\big) = 1$, yet at the same time, $\underline{\operatorname{gr}}\big(T^{(1)} \vee T^{(2)}\big) > 1$. Let $n_k \uparrow \infty$. Let $T^{(1)}$ (resp., $T^{(2)}$) be a tree such that x has one child (resp., two children) for $n_{2k} \leq |x| \leq n_{2k+1}$ and two (resp., one) otherwise; this is shown schematically in Figure 1.11. If n_k increases sufficiently rapidly, then $\operatorname{br} T^{(1)} = \operatorname{br} T^{(2)} = 1$, so $\operatorname{br}\big(T^{(1)} \vee T^{(2)}\big) = 1$. Prove that if $\langle n_k \rangle$ increases sufficiently rapidly, then $\underline{\operatorname{gr}}\big(T^{(1)} \vee T^{(2)}\big) = \sqrt{2}$. Furthermore, show that the set of possible values of $\underline{\operatorname{gr}}\big(T^{(1)} \vee T^{(2)}\big)$ over *all* sequences $\langle n_k \rangle$ is $\big[\sqrt{2}, 2\big]$.

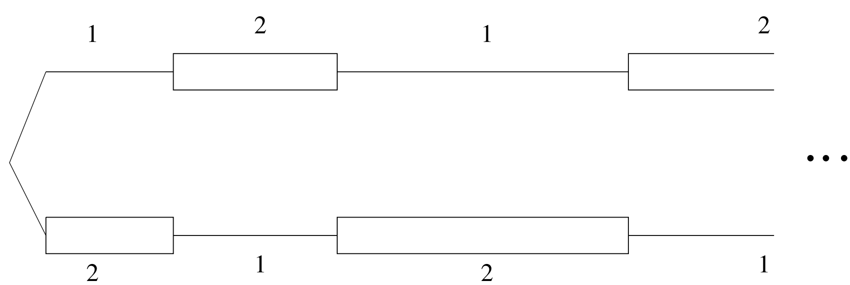

Figure 1.11. A schematic of a tree with branching number 1 and growth rate $\sqrt{2}$.

1.4. Complete Example 1.4 by showing that when λ^{-n} is the conductance of edges at distance n from the root of a binary tree T, current does not flow for $\lambda \geq 2$.

https://doi.org/10.1017/9781316672815.002 Published online by Cambridge University Press