ELECTRICAL NETWORKS AND REVERSIBLE MARKOV CHAINS

STEVEN P. LALLEY

The object of this part of the course is to show how the mathematics of finite electrical networks can be used in the study of reversible Markov chains on finite and countable state spaces. The basis of the connection is that *harmonic functions* for reversible Markov chains can be interpreted as *voltages* for electrical networks. Thus, techniques for calculating, approximating, and bounding voltages (especially *shorting* and *cutting*) may be applied to problems involving harmonic functions (and, in particular, hitting probabilities).

1. Definitions: Graphs

Graph: A pair $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ consisting of a finite or countable set $\mathscr{V}$ of *vertices* and a a finite or countable set $\mathscr{E}$ of *edges*. Elements of the set $\mathscr{E}$ are *unordered* pairs $\{x, y\}$ of vertices (these are called the *endpoints* of the edge). (Sometimes we will denote an edge by xy rather than $\{x, y\}$.) An edge with endpoint x is said to be *incident* to x. A vertex y such that there is an edge with endpoints x, y is said to be *adjacent* to x, or a *nearest neighbor* of x. For each vertex x, denote by $\mathscr{E}_x = \mathscr{E}(x)$ the set of edges incident to x, and by $\mathscr{V}_x = \mathscr{V}(x)$ the set of vertices adjacent to x. Say that $\mathscr{G}$ is *locally finite* if, for every $x \in \mathscr{V}$, $\mathscr{E}_x$ is a finite set.

Subgraph: A subgraph of $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ is a graph $\mathscr{G}' = (\mathscr{V}', \mathscr{E}')$ such that $\mathscr{V}' \subset \mathscr{V}$ and $\mathscr{E}' \subset \mathscr{E}$. The subgraph $\mathscr{G}'$ of $\mathscr{G}$ is called a *spanning subgraph* if it has the same vertex set.

Isomorphism: An *isomorphism* between two graphs $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ and $\mathscr{G}' = (\mathscr{V}', \mathscr{E}')$ is a pair of bijections $\varphi_V : \mathscr{V} \to \mathscr{V}'$ and $\varphi_E : \mathscr{E} \to \mathscr{E}'$ such that for every edge $e \in \mathscr{E}$, with endpoints x, y, the image $\varphi_E(e)$ has endpoints $\varphi_V(x), \varphi_V(y)$.

Automorphism: An *automorphism* of $\mathscr{G}$ is an isomorphism of $\mathscr{G}$ onto itself. Observe that the set $Aut(\mathscr{G})$ of all automorphisms of a graph is a *group* under the operation of functional composition.

Directed Graph: A pair $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ consisting of a set $\mathscr{V}$ of *vertices* and a set $\mathscr{E}$ of *directed* edges. Elements of the set $\mathscr{E}$ are *ordered* pairs (x, y) of vertices: x is called the *tail* and y the *head* of the edge. (Sometimes we will denote a directed edge by xy instead of (x, y).) Edges are sometimes called *arrows*. *Isomorphisms* and *automorphisms* of directed graphs are defined in a similar fashion as for (undirected) graphs. For each vertex x, denote by $\mathscr{E}_x^-, \mathscr{E}_x^+$ the sets of arrows with tail x and head x, and let $\mathscr{E}_x = \mathscr{E}_x^+ \cup \mathscr{E}_x^-$. Define $\mathscr{V}_x^+, \mathscr{V}_x^-$, and $\mathscr{V}_x$ similarly.

Path: A path in a graph or digraph $\mathscr{G}$ is a sequence $x_0 x_1 x_2 \dots$ of vertices such that there is an edge connecting x_i and x_{i+1} for each i. A path in a multigraph is a sequence $x_1 x_2 \dots$ of vertices such that for each i there is a directed edge with tail x_i and head x_{i+1} . The *length* of a finite path $x_0 x_1 x_2 \dots x_n$ is defined to be n; its *endpoints* are $x_0$and x_n. If the endpoints x_0 and x_n are the same vertex, the path is called *closed*. A path is called *self-avoiding* if all of its vertices are distinct.

Distance: The distance $d(x,y)$ between two vertices x,y is the minimum integer n such that there is a path of length n with endpoints x and y, or ∞ if there is no path with endpoints x,y. The distance between two sets $A,B\subset\mathscr{V}$ is the minimum distance $d(x,y)$, where $x\in A$ and $y\in B$.

Boundaries: Let $(\mathscr{V},\mathscr{E})$ be a directed graph (or multi-graph). For any set $F\subset\mathscr{V}$ of vertices, the in- and out-boundaries of F are defined as follows:

$$\begin{aligned}\partial_{\text{out}}F&=\{y\notin F:(x,y)\in\mathscr{E}\ \text{ for some }\ x\in F\};\\ \partial_{\text{in}}F&=\{y\notin F:(y,x)\in\mathscr{E}\ \text{ for some }\ x\in F\}.\end{aligned}\tag{1}$$

If $F=\{x\}$ is a singleton then $\partial_{out}F=\mathscr{V}_x^+$ and $\partial_{\text{in}}(F)=\mathscr{V}_x^-$.

Connected Graph: A graph (or directed graph) is *connected* if for every pair $x,y\in\mathscr{V}$ there is a path with endpoints x and y (equivalently, $d(x,y)<\infty$).

Cycle: A finite closed path $x_0x_1x_2\dots x_n$ such that no two of the vertices $x_1x_2\dots x_n$ are the same.

Tree: A connected graph with no cycles. In a tree, a vertex with only one incident edge is called a *leaf*. A graph (not necessarily connected) with no cycles is called a *forest*.

Variations: A *multigraph* may have multiple edges connecting the same pair $\{x,y\}$ of vertices. A *graph with loops* may have edges whose endpoints are the same vertex. *Directed multigraphs* and *directed graphs with loops* are defined similarly. Paths, cycles, and distances in multigraphs, directed multigraphs, etc. are defined in much the same way as for graphs and directed graphs. Check the first chapter of BOLLOBAS for further terminology concerning graphs.

2. HITTING PROBABILITIES AND HARMONIC FUNCTIONS

2.1. **Reversible Markov Chains.** Let $\mathscr{G}=(\mathscr{V},\mathscr{E})$ be a graph or digraph. A *Markov chain* on $\mathscr{G}$ is a Markov chain whose state space is the vertex set $\mathscr{V}$ of $\mathscr{G}$, and whose transition probabilities $p(x,y)$ satisfy the condition

$$p(x,y)>0 \qquad \text{if and only if } (x,y)\in\mathscr{E}.$$

It is called *reversible* if there is a function $w:\mathscr{V}\to(0,\infty)$ such that the following *detailed balance equation* holds for all $x,y\in\mathscr{V}$:

$$w_x p(x,y)=w_y p(y,x).\tag{2}$$

Many interesting Markov chains are reversible. The electrical network connection is valid only for reversible Markov chains; hence our interest in them.

Examples:

(1) Any Markov chain on a tree.
(2) Any birth and death chain on the integers.
(3) Any *homogeneous* nearest neighbor random walk on the $d-$dimensional integer lattice $\mathbb{Z}^d$. *Homogeneous* means that for any $x,y,z\in\mathbb{Z}^d$,

$$p(x,y)=p(x+z,y+z).$$

EXERCISE: Find (or construct) the relevant weight functions w.

Until further notice, we will assume that all Markov chains are irreducible, i.e., that for any two vertices there is a positive probability path connecting them. This implies that the underlying graph $\mathscr{G}$ is connected. Notice that if a reversible Markov chain is irreducible then the weight function w is unique up to multiplication by a scalar (proof by induction on the distance from a distinguished vertex). Before giving examples of reversible Markov chains we will state a basic result due to Kolmogorov that partially explains the terminology and gives a useful tool for verifying that a Markov chain is reversible.

Proposition 1. *An irreducible Markov chain with transition probabilities $p(x,y)$ is reversible iff for every closed path $x_0, x_1, x_2, \dots, x_n = x_0$,*

$$\prod_{i=0}^{n} p(x_i, x_{i+1}) = \prod_{i=0}^{n} p(x_{i+1}, x_i).$$

The proof is easy, and left as an exercise (or see any standard textbook). It gives easy proofs that each of the 3 examples above is a reversible Markov chain. We won't have any further use for Kolmogorov's theorem. The following is also true:

Proposition 2. *Suppose that the detailed balance equations hold with weight function w. If $\sum_{x\in\mathscr{V}} w_x < \infty$ then the Markov chain is positive recurrent, and $\pi_x = w_x / \sum_{x\in\mathscr{V}} w_x$ is a stationary probability distribution. Moreover, the detailed balance equations hold with w replaced by π.*

The proof is left as an (easy) exercise.

2.2. **Harmonic Functions and Hitting Probabilities.**

Assumption 1. Assume henceforth that all Markov chains are irreducible and that their associated graphs or digraphs are locally finite.

Definition 3. Let F be a subset of $\mathscr{V}$. A function $h : F \cup \partial_{\text{out}} F \to \mathbb{R}$ is said to be *harmonic* at a vertex $x \in F$ if

$$h(x) = \sum_y p(x,y) h(y) \tag{3}$$

It is said to be *harmonic in B* if it is harmonic at every $x \in B$, and is said to be *harmonic* if it is harmonic on the entire state space $\mathscr{V}$.

Since the underlying graph is *locally finite* (meaning that no vertex has infinitely many nearest neighbors) the sum in (3) has only finitely many nonzero terms. Observe that any linear combination of harmonic functions is harmonic: this is known as the *Superposition Property*. In particular, the *difference* between two harmonic functions is harmonic. The term "harmonic" is used because the harmonic functions of classical mathematics are those that satisfy the "mean value property".

The next proposition is called the *Maximum Principle* for harmonic functions.

Proposition 4. *Let $B \subset \mathscr{V}$ be a* finite *set of vertice whose out-boundary $\partial_{out} B \not\emptyset$. Suppose that $h : B \cup \partial B \to \mathbb{R}$ is harmonic in B. Then h has no nontrivial local maximum in B, that is, there is no vertex $x \in B$ such that*

$$h(x) \geq h(y) \quad \textit{for all } y \in \mathscr{E}_x^+ \quad \textit{and}$$
$$h(x) > h(y) \quad \textit{for some } y \in \mathscr{E}_x^+.$$

Furthermore, if $\partial_{out} B \neq \emptyset$ then

$$\max_{x \in B} h(x) \leq \max_{x \in \partial B} h(x). \tag{4}$$

If B is connected then the inequality is strict unless h is constant on $B \cup \partial B$.

Proof. The first statement is fairly obvious. If h is harmonic at x then the equation (3) holds, so $h(x)$ is a (weighted) average of the values $h(y)$, where y is an out-neighbor of x. Hence, if $h(x) \geq h(y)$ for all such y then it must be that $h(x) = h(y)$ for all such y.

Because B is finite, h must achieve a maximum m at some $x \in B$. Pick one such x. Then $h(y) = h(x) = m$ for all out-neighbors y of x, and similarly $h(z) = h(y) = m$ for all out-neighbors z of out-neighbors y, and so on. Because the state space is connected (by assumption, all Markov chains considered here are irreducible), there must be a path from x to $\partial_{out} B$ along which $h = m$ (unless $\partial_{out} B = \emptyset$). Since this path ends at $\partial_{out} B$, it follows that there is at least one $y \in \partial_{\text{out}} B$ such that $h(y) = m$. This proves (4). The proof of the last statement is similar and is left as an exercise. □

Corollary 5. *(Uniqueness Theorem) If f, g are harmonic on B, and if $f = g$ on ∂B, then $f = g$ in B, provided B is finite and $\partial_{out} B \neq \emptyset$.*

Proof. The difference $h = f - g$ is harmonic in B and zero on ∂B, so the Maximum Principle implies that $h = 0$ in B. □

For any subset B of $\mathscr{V}$, define $T = T_B$ to be the time of first exit from B, i.e.,

$$T = T_B = \inf\{n \geq 0 : X_n \notin B\}.$$

Proposition 6. *Let $B \subset \mathscr{V}$ be such that $P^x\{T_B < \infty\} = 1$ for every $x \in B$. Suppose that $\partial_{out} B = F \cup G$, where F and G are disjoint sets. Define $h(x) = P^x\{X_T \in G\}$ for $x \in B \cup \partial_{out} B$. Then h is harmonic in B and satisfies the boundary conditions $h \equiv 0$ on F and $h \equiv 1$ on G. Moreover, it is the* unique *bounded harmonic function satisfying these boundary conditions.*

Proof. That h is harmonic in B is an easy consequence of the Markov property. (Condition on the first step X_1.) If $B \cup \partial_{\text{out}} B$ is finite, then there can only be one harmonic function satisfying the boundary conditions $h \equiv 0$ on F and $h \equiv 1$ on G, by the Maximum Principle (Proposition 4) and the Superposition Principle. If B is infinite then uniqueness follows from the following proposition. □

Proposition 7. *If $h : \mathscr{V} \to \mathbb{R}$ is a bounded (nonnegative) function that is harmonic in B, then $h(X_{n \wedge T})$ is a bounded (nonnegative) martingale under any P^x.*

The proof is an elementary exercise in the use of the Markov property and the definition of a martingale (exercise). We won't use the result of Proposition 6 except when B is finite, so don't panic if you don't yet know what a martingale is. But if you do, then here is the argument that shows how uniqueness in Proposition 6 follows from Proposition 7:

Suppose that there were *two* bounded harmonic functions both satisfying the same boundary conditions. Then their difference h would be a harmonic function in B such that $h \equiv 0$ in ∂B, and since $T = T_B < \infty$ a.s., it would follow that $h(X_T) = 0$ a.s. But then, by the optional sampling theorem for bounded martingales, $h(x) = E^x h(X_0) = E^x h(X_{T\wedge n})$ for all $n \geq 0$ and all $x \in B$. Consequently, by the DCT,

$$h(x) = \lim_{n\to\infty} E^x h(X_{T\wedge n}) = E^x h(X_T) = 0,$$

proving that the two harmonic functions were actually identical.

2.3. **Harmonic Functions: Examples.** Proposition 6 provides a road map for solving "Gambler's Ruin" problems for Markov chains. It is, of course, effective only when harmonic functions can be identified, but there are many examples where this can be done.

Example 8. Consider simple random walk S_n on the integers $\mathbb{Z}$. It is trivial to check that *any linear function* $h(x) = ax + b$ is harmonic. Let $B = [m] = \{1,2,\ldots,m\}$, and define $T = T_B$ to be the time of first exit from B, as above. If the initial point of the Markov chain is in $B \cup \partial B$ then with P^x−probability 1 the exit point $S_T = 0$ or $m+1$. To find the probability $h(x)$ that $S_T = x$ we must find a harmonic function such that $h(0) = 0$ and $h(m+1) = 1$. This is easy:

$$h(x) = \frac{x}{m+1}.$$

Example 9. Let X_n be $p - q$ random walk on the integers, with $p > 1/2 > q > 0$. It is routine to check that the following are all harmonic functions:

$$h(x) = A + B\left(\frac{q}{p}\right)^x.$$

To solve the Gambler's Ruin problem here, look for a harmonic function such that $h(0) = 0$ and $h(m+1) = 1$. The one that works is

$$h(x) = \frac{(q/p)^x - 1}{(q/p)^{m+1} - 1}.$$

3. Electrical Networks

3.1. **Electrical Networks and Reversible Markov Chains.** An *electrical network* is a finite graph (or multigraph, or multigraph with loops) $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ together with an assignment of *resistances* $R_e = R(e)$ to the edges. Given a resistance function R, one defines the *conductance* of any edge to be the reciprocal of its resistance:

$$C_e = C(e) = 1/R(e).$$

Note that the electrical network is completely specified by either the resistance function or the conductance function, as one can be recovered from the other. Say that an electrical network is *connected* if the underlying graph $\mathscr{G}$ is connected and $C_e > 0$ for every edge e.

Standing Assumption: Unless otherwise specified, all electrical networks considered are connected and locally finite.

There is a natural one-to-one correspondence between electrical networks and reversible Markov chains (either in discrete or continuous time). Consider a finite graph $\mathscr{G}$ (no loops or multiple edges!). The correspondence between conductance functions and reversible transition probability kernels (for discrete time Markov chains) is given by

$$p(x,y) = \frac{C_{xy}}{C_x}, \tag{5}$$

where xy denotes the edge with endpoints x,y. Note that any function $p(x,y)$ so defined is a transtion probability kernel, and that the detailed balance equations are satisfied with $w_x = C_x$. Moreover, the Markov chain with these transition probabilities is irreducible, since the electrical network is assumed to be connected. Conversely, given a transition probability kernel satisfying the detailed balance equations (2, one may unambiguously define conductances by

$$C_{xy} = w(x)p(x,y) = w(y)p(y,x). \tag{6}$$

(For electrical networks on multigraphs, or multigraphs with loops, it is also possible to define in a natural way a corresponding reversible Markov chain, but we won't have any need for this: we will only use electrical networks on multigraphs as technical aids for "solving" electrical networks on graphs.)

Notice that the definition of a harmonic function may be given in terms of the conductances C_ε as follows: h is harmonic at $x \in \mathscr{V}$ iff

$$\sum_{e \in \mathscr{E}_x} C(e)(h(x) - h(y)) = 0, \tag{7}$$

where x,y are the endpoints of e. We will see that harmonic functions have interpretations as *electrical potentials.*

3.2. **Flows in Graphs and Multigraphs.** Let $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ be a graph or multigraph (for simplicity we allow no loops). Each edge e has two incident vertices (endpoints) x,y, and so may be *oriented* (or *directed*) in two ways: from x to y, or from y to x. Write $e_{xy} = e(xy)$ and $e_{yx} = e(yx)$ for these two orientations. For each vertex x, define

$$\mathscr{E}_x^+ = \mathscr{E}^+(x) = \{e(xy) : e \in \mathscr{E}_x\} \qquad \text{and}$$
$$\mathscr{E}_x^- = \mathscr{E}^-(x) = \{e(yx) : e \in \mathscr{E}_x\}$$

to be the sets of oriented edges *leaving* and *entering* vertex x, respectively. Observe that each of the sets $\mathscr{E}^+(x)$ and $\mathscr{E}^-(x)$ is in one-to-one correspondence with the set $\mathscr{E}(x)$ of edges incident to x. Define $\mathscr{E}^*$ to be the set of *all* oriented edges of $\mathscr{G}$. Observe that $(\mathscr{V}, \mathscr{E}^*)$ is a directed graph with teh special property that there is a natural pairing of directed edges, namely, $e(xy) \longleftrightarrow e(yx)$.

Flow: A *flow* in $\mathscr{G}$ is a function $J : \mathscr{E}^* \to \mathbb{R}$ satisfying the following two properties: (a) for any edge e with endpoints x,y,

$$J(e(xy)) = -J(e(yx)); \tag{8}$$

and (b) for every vertex x,

$$\sum_{e \in \mathscr{E}(x)} J(e(xy)) = \sum_{e \in \mathscr{E}(x)} J(e(yx)) = 0. \tag{9}$$

If equation (9) holds at all vertices x *except* vertices a and b, then J is called a flow with *source* a (or b) and *sink* b (or a) depending on which case of the following lemma obtains. A function $f:\mathscr{E}^*\to\mathbb{R}$ that satisfies equation (8) will be called *antisymmetric* (in cohomology theory, such a function would be called a *1-form*).

Lemma 10. *Suppose that J is a function that satisfies (8) for every edge e, and satisfies (9) at every vertex x except vertices a and b. Then either*

$$J(a+)\stackrel{\Delta}{=}\sum_{e\in\mathscr{E}(a)}J(e(ay))=J(b-)\stackrel{\Delta}{=}\sum_{e\in\mathscr{E}(b)}J(e(yb))>0 \qquad \textit{or} \tag{10}$$

$$J(a+)=\sum_{e\in\mathscr{E}(a)}J(e(ay))=J(b-)=\sum_{e\in\mathscr{E}(b)}J(e(yb))<0. \tag{11}$$

Proof. Exercise. □

Acyclicity: A *cycle* in $\mathscr{G}$ is a finite sequence $e(x_0x_1)$, $e(x_1x_2)$, …, $e(x_{n-1}x_n)$ of oriented edges such that (i) for each i, the *tail* of $e(x_ix_{i+1})$ is the *head* of $e(x_{i-1}x_i)$; and (ii) $x_0=x_n$. An antisymmetric function $J:\mathscr{E}^*\to\mathbb{R}$ is called *acyclic* if it sums to zero around any directed cycle in $\mathscr{G}$, i.e., for every cycle,

$$\sum_{i=0}^{n-1}J(e(x_ix_{i+1}))=0. \tag{12}$$

Lemma 11. *Every acyclic antisymmetric function is a gradient, i.e., if J is antisymmetric and acyclic then there exists a function $\varphi:\mathscr{V}\to\mathbb{R}$ on the set of vertices (sometimes called the* potential function*) such that for every edge e with incident vertices x,y,*

$$J(e(xy))=\varphi(y)-\varphi(x). \tag{13}$$

Proof. It is enough to consider the case of a *connected* graph or multigraph. Fix a vertex $x_*=x_0$, and define $\varphi(x_*)=0$. For any other vertex y there is a directed path from x_* to y, through vertices $x_0,x_1,\dots,x_n=y$. Define

$$\varphi(y)=\sum_{i=0}^{n-1}J(e(x_ix_{i+1})).$$

That this is a valid definition follows from the hypothesis (12), since this guarantees that for any two directed paths from x_* to y, the sums are equal. The desired relation (13) follows by choosing the right paths. □

3.3. **Electrical Current: Ohm's and Kirkhoff's Laws.** If the two terminals of a 1-volt battery are attached to vertices a and b of an electrical network, then electrical current will flow through the edges (resistors) of the network. The laws governing this current flow are due to Ohm and Kirchhoff; they are experimentally observable facts. According to these laws, there are two functions

$$I:\mathscr{E}^*\longrightarrow\mathbb{R} \qquad \text{and}$$
$$\Psi:\mathscr{E}^*\longrightarrow\mathbb{R},$$

both satisfying (8), that measure the *potential difference* and the *current flow* across directed edges. Assume that the positive terminal of the battery is attached to a and the negative terminal to b.

Ohm's Law: *For every edge e with endpoints x, y,*

$$\Psi(e(xy)) = I(e(xy))C(e). \tag{14}$$

Kirchhoff's Potential Law: *The potential difference function Ψ is acyclic.*

Kirchhoff's Current Law: *The current I is a flow with source a and sink b.*

Proposition 12. *There is a unique function $\varphi : \mathscr{V} \to [0,1]$, called the* voltage *or* electrical potential, *such that*

(a) $\varphi(a) = 1$ *and* $\varphi(b) = 0$.
(b) φ *is harmonic at every vertex* $x \neq a, b$.
(c) For every edge e with endpoints x, y,

$$\Psi(e(xy)) = \varphi(x) - \varphi(y). \tag{15}$$

Proof. Since the potential difference function Ψ is acyclic, by Kirchhoff's first law, Lemma 11 implies that it is the gradient of a real-valued function φ on the vertices. This function may be so chosen that its value at b is 0. Since the attached battery is a 1-volt battery, it follows that $\varphi(a) = 1$. That φ is harmonic except at a and b follows from Ohm's Law and Kirchhoff's second law, because the definition of a flow implies, for every vertex $x \neq a, b$, that

$$\begin{aligned}
\sum_{e \in \mathscr{E}(x)} I(e(xy)) = 0 \qquad &\Longrightarrow \\
\sum_{e \in \mathscr{E}(x)} \Psi(e(xy))/R(e) = 0 \qquad &\Longrightarrow \\
\sum_{e \in \mathscr{E}(x)} \Psi(e(xy))C(e) = 0 \qquad &\Longrightarrow \\
\sum_{e \in \mathscr{E}(x)} (\varphi(x) - \varphi(y))C(e) = 0, &
\end{aligned}$$

and this implies that φ is harmonic at x, by equation (7. Uniqueness now follows from the Maximum Principle for harmonic functions. That φ can only take values between 0 and 1 follows from the maximum principle. □

Corollary 13. *The voltage $\varphi(x)$ at any vertex x is the probability that the Markov chain on $\mathscr{G}$ with transition probabilities (5) and initial point $X_0 = x$ will visit a before b.*

3.4. **Effective Resistance and Escape Probabilities.** Consider the electrical network on a *connected* graph $\mathscr{G}$ with resistance function R and conductance function $C = 1/R$. Denote by $X_0, X_1, X_2, \dots$ the Markov chain with transition probabilities $p(x, y)$ given by (5), and such that under the probability measure P^x the initial state $X_0 = x$ a.s. For any two vertices $a \neq b$, define the escape probability

$$p_{\text{escape}}(a; b) = P^a\{\text{no return to } a \text{ before first visit to b}\}. \tag{16}$$

Define the *effective resistance* $R_{\text{eff}} = R_{\text{eff}}(a, b)$ between the vertices a and b by

$$R_{\text{eff}} = \frac{1}{I(b-)} = \frac{1}{I(a+)} \tag{17}$$

where $I(a+)$ and $I(b-)$ denote the net current flow out of a and into b (recall that these are equal – see Lemma 10). This equation may be thought of as an extension of Ohm's Law, because this is what the resistance would have to be if the network were to be replaced by a single resistor between a and b in such a way that the total current flow and the voltage differential were the same as for the original network.

The next theorem makes clear the importance of the notion of effective resistance. Recall that $C_x = \sum_{e\in\mathscr{E}_x} C(e)$ is the total conductance out of x.

Theorem 14. $p_{\text{escape}}(b;a) = 1/(C_b R_{\text{eff}})$.

Proof. Suppose that the first step of the Markov chain is to vertex x. Then by our earlier characterization of harmonic functions as hitting probabilities, the conditional probability of escape to b before return to a is $\varphi(x)$. Consequently,

$$\begin{aligned}
p_{\text{escape}}(b;a) &= \sum_x p(b,x)\varphi_x \\
&= \sum_x p(b,x)(\varphi_x - \varphi_b) \\
&= \sum_x C_{bx}(\varphi_x - \varphi_b)/C_b \\
&= \sum_x I_{bx}/C_b \\
&= I_b/C_b \\
&= 1/C_b R_{\text{eff}}.
\end{aligned}$$

□

The preceding theorem is the keystone of the theory to follow. We will use it for two purposes: (1) explicit computations of escape probabilities; and (2) transience/recurrence theorems. First, however, we will develop some further characterizations of the electrical potential and the electrical current flow.

3.5. **Solving Electrical Networks.** For our purposes, "solving" an electrical network means finding the effective resistance between two vertices a and b. (Our interest in computing effective resistances stems from their usefulness in computing escape probabilities.) The basic strategy is to successively replace the network by simpler and simpler networks without changing the effective resistance between a and b. Here it is important to allow multiple edges between 2 vertices (thus, in this section $\mathscr{G}$ is in general a "multigraph").

Say that an electrical network $(\mathscr{G}, R)$ may be *reduced* to the network $(\mathscr{G}', R')$ if a and b are vertices in both networks and the effective resistance between a and b is the same for both $(\mathscr{G}, R)$ and $(\mathscr{G}', R')$. We will show that there are 3 different kinds of elementary reductions that can be made: these are known as the "Series Law", the "Parallel Law", and the "Shorting Law". Informally, these may be stated as follows:

Series Law: Resistences in series add.
Parallel Law: Conductances in parallel add.
Shorting Law: Vertices at the same voltage may be shorted.

More precisely,

Proposition 15. *(Series Law) Suppose that vertices $x_0, x_1, x_2, \dots, x_n$ appear in series, i.e., there is a single resistor of resistance R_i connecting x_i to x_{i+1} for each $i = 1, 2, \dots, n$, and there are no other edges emanating from any of the vertices $x_1, x_2, \dots, x_{n-1}$. Assume that the source a and the sink b are not in the series $x_0, x_1, x_2, \dots, x_n$. Then the network may be reduced to the network in which the vertices $x_1, x_2, \dots, x_{n-1}$ and the edges emanating from them are removed and replaced by a single resistor between x_0 and x_n with resistance $R_1 + R_2 + \dots + R_n$.*

Proposition 16. *(Parallel Law) Suppose that vertices x, y are connected by edges $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ with resistances $R_1, R_2, \dots, R_n$. Then the network may be reduced to the network in which the edges $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are replaced by a single edge of resistance $(R_1^{-1} + R_2^{-1} + \dots + R_n^{-1})^{-1}$ (i.e., an edge whose conductance is the sum of the conductances of the edges replaced).*

Proposition 17. *(Shorting Law) If two vertices x, y have the same voltage, then the network may be reduced to the network in which x, y are replaced by a single vertex z and the edges with a vertex at x or y are rerouted to the vertex z.*

The reason this is called “shorting” is that in effect one is introducing a new wire of *infinite* conductivity (zero resistance) between x and y.

The proofs of all three laws are easy, and should be done as EXERCISES.

3.6. **Examples.** Three simple examples of the above laws follow.

Example 18. Simple Nearest Neighbor Random Walk on the Integers Consider simple random walk on the set of integers $\{0, 1, 2, \dots, m\}$, with reflection at the endpoints $0, m$. This is the reversible Markov chain corresponding to the electrical network with unit resistances between i and $i+1$ for all $i = 0, 1, \dots, m-1$. The series law implies that the effective resistance between 0 and m is m. It follows that the probability of escape from 0 to m is $1/m$. Note that as $m \to \infty$ the escape probability drops to 0; this implies that the simple nearest neighbor random walk on the integers is recurrent.

Example 19. p-q Random Walk on the Integers Consider the network on the graph whose vertex set is the set of positive integers and whose edges are between nearest neighbors. The conductance function is as follows:

$$\begin{aligned} C_{0,1} &= 1; \\ C_{x,x+1} &= \left(\frac{p}{q}\right) C_{x-1,x} \\ &= \left(\frac{p}{q}\right)^x. \end{aligned}$$

The reversible Markov chain corresponding to this conductance function is the p-q random walk on the positive integers. The effective resistance between 0 and b can be found using the series law: the series between 0 and b can be replaced by a single edge from 0 to b of resistance

$$\sum_{x=0}^{b-1} \left(\frac{q}{p}\right)^x = \frac{1-(q/p)^b}{1-(q/p)}.$$

This is obviously the effective resistance. It follows that the escape probability from 0 to b is 1 over the effective resistance, which is

$$\frac{1-(q/p)}{1-(q/p)^b}.$$

Example 20. Simple Random Walk on the Cube Consider the simple nearest neighbor random walk on the vertices of the cube $\mathbb{Z}_2^3$. This is the reversible Markov chain corresponding to the electrical network whose vertices and edges are the vertices and edges of the cube, and for which the conductances of the edges are all 1. We will calculate the effective resistance and the escape probability for two neighboring vertices, $a = 000$ and $b = 001$. Symmetry implies that the voltage function satisfies

$$\begin{aligned} h(010) &= h(100); \\ h(011)) &= h(101). \end{aligned}$$

The shorting law implies that the network may be reduced by shorting these 2 triples of vertices. This leaves a network with multiple edges that may be successively reduced by applying the parallel and series laws in sequence. One finds (I think) that the effective resistance between a and b is 4/3, so the escape probability is 3/4.

4. Variational Principles

Throughout this section, let $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ be a finite graph (no multiple edges or loops) equipped with resistance function R and conductance function C. Let I be the electrical current function and φ the voltage function when a $v-$volt battery is attached to vertices a and b. The goal of this section is to show that I and φ satisfy simple *variational principles*.

Let $\ell^2(\mathscr{V})$ and $\ell^2(\mathscr{E}^*)$ denote the vector spaces of real-valued functions on the sets of vertices and oriented edges, respectively, equipped with the inner products

$$\langle f, g\rangle = \langle f, g\rangle_{\ell^2(\mathscr{V})} = \sum_{x\in\mathscr{V}} f(x)g(x) \qquad \text{and} \tag{18}$$

$$\langle f, g\rangle = \langle f, g\rangle_{\ell^2(\mathscr{E}^*)} = \sum_{e(xy)\in\mathscr{E}^*} f(e(xy))g(e(xy)). \tag{19}$$

Let

$$\ell^2_\pm(\mathscr{E}^*) = \{\text{antisymmetric } f \in \ell^2(\mathscr{E}^*)\} \tag{20}$$

be the subspace of $\ell^2(\mathscr{E}^*)$ consisting of all *antisymmetric* functions (functions that satisfy equation (8)). There is a natural linear transformation from $\ell^2(\mathscr{V})$ to $\ell^2_\pm(\mathscr{E}^*)$, called the *gradient* map ∇, defined by

$$\nabla f(e_{xy}) = f(y) - f(x). \tag{21}$$

Dirichlet Forms: The Dirichlet form associated with the conductance function C is the quadratic form on $\ell^2_\pm(\mathscr{E}^*)$ defined by

$$\mathscr{D}(f, g) = \frac{1}{2}\sum_{e(xy)\in\mathscr{E}^*} f(e_{xy})g(e_{xy})R(e). \tag{22}$$

Observe that each edge $e \in \mathscr{E}$ is represented twice in this sum, once for each orientation; and the two terms of the sum are equal, since both f and g are antisymmetric. (This explains the factor of 1/2.) There is an analogous quadratic form defined on $\ell^2(\mathscr{V})$, which we will call the *dual Dirichlet form*, by

$$\begin{aligned}\mathscr{D}^*(f,g) &= \mathscr{D}(C\nabla f, C\nabla g) \\ &= \frac{1}{2}\sum_{e(xy)\in\mathscr{E}^*} \nabla f(e(xy))\nabla g(e(xy))C(e).\end{aligned} \tag{23}$$

Theorem 21. *(Dirichlet's Principle, First Form) Let φ be the electrical potential when a battery of voltage 1 is attached to the vertices a and b, i.e., φ is unique bounded function on $\mathscr{V}$ that is harmonic on $\mathscr{V}-\{a,b\}$ and satisfies the boundary conditions $\varphi(a)=1$ and $\varphi(b)=0$. Then φ minimizes the Dirichlet form $\mathscr{D}^*(h,h)$ within the class of functions $h\in\ell^2(\mathscr{V})$ satisfying the boundary conditions $h(a)=1$ and $h(b)=0$.*

Remark 1. Note: The uniqueness of the minimizing function φ is part of the assertion.

Theorem 22. *(Dirichlet's Principle, Second Form) Let I be the electrical current flow when a $v-$volt battery is hooked up at a and b. Then among all flows J on the graph $\mathscr{G}$ with source a and sink b and satisfying the "total flow" constraint $J(a+)=I(a+)$,the flow I minimizes the Dirichlet form $\mathscr{D}(J,J)$. Moreover,*

$$\mathscr{D}^*(\varphi,\varphi)=\mathscr{D}(I,I). \tag{24}$$

The proofs will require the following:

Proposition 23. *(Conservation Law) Let $w:\mathscr{V}\to\mathbb{R}$ be any real-valued function on the set of vertices, and let J be a flow on $\mathscr{G}$ with source a and sink b. Then*

$$(w_a-w_b)J(a+)=\frac{1}{2}\sum_{e(xy)\in\mathscr{E}^*}(w_x-w_y)J(e(xy)).$$

Proof. This is a simple computation: just group terms in the sum according to the initial point of the edge. In detail,

$$\begin{aligned}\frac{1}{2}\sum_{\mathscr{E}^*}(w_x-w_y)J(e(xy)) &= \sum_{\mathscr{E}^*} w_x J(e(xy)) \\ &= \sum_{x\in\mathscr{V}}\sum_{e\in\mathscr{E}(x)} w_x J(e(xy)) \\ &= \sum_{\mathscr{E}(a)} w_a J(e(ay)) + \sum_{\mathscr{E}(b)} w_b J(e(by)) \\ &= w_a J(a+) + w_b J(b+) \\ &= (w_a-w_b)J(a+).\end{aligned}$$

□

We will only use this as a technical lemma; however, it is an interesting result in its own right in graph theory. It has the following interpretation: think of the flow as representing a flow of some "commodity" in an economic network. At each node x the value of (a unit amount of) the commodity is determined by the function w_x. Thus, for each directed edge $e(xy)$ the quantity $(w_y - w_x)J(e_{xy})$ is the "value added" to the amount $J(e_{xy})$ of the commodity that passes through the edge. Similarly, $(w_a - w_b)J(a+)$ represents the change in value of the amount $J(a+)$ of the commodity that enters the network at a by the time it leaves the network from b. This should make the conservation law intuitively clear.

Proof of Theorem 21. There is clearly a function h that minimizes the dual Dirichlet form $\mathscr{D}^*(h,h)$ subject to the constraints $h(a) = 1, h(b) = 0$, since this is a finite–dimensional minimization problem. We will argue that this function is harmonic.

Recall that C_x is the sum of the conductances leading out of x. Note that $C_x > 0$ because of our standing assumption that the electrical network is connected. For any directed edge $e = e(xy) \in \mathscr{E}^+(x)$, let $p^x(e) = C(e)/C_x$; this is a probability distribution on the edges leading out of x. The condition that h is harmonic at x is equivalent to the statement that

$$h(x) = \sum_{e(xy)\in\mathscr{E}_x^+} p^x(e)h(y).$$

Fix a vertex $x \neq a,b$, and consider those terms in the sum defining the dual Dirichlet form $\mathscr{D}^*(h,h)$ that involve $h(x)$: they contribute

$$\sum_{\mathscr{E}_x}(h(x)-h(y))^2C(e) = C_x\sum_{\mathscr{E}_x}(h(x)-h(y))^2p^x(e).$$

Here y denotes the other endpoint of e. For fixed values of $h(y)$, $y \neq x$, the value of $h(x)$ that minimizes this is

$$h(x) = \sum_{\mathscr{E}_x} p(e)^x h(y).$$

Hence, if h were *not* harmonic at x, then it would not minimize the energy $\mathscr{D}(h)$ (this argument does not apply at $x = a,b$, though, because at those points h must satisfy the boundary conditions). Thus, h minimizes the dual Dirichlet form only if it is harmonic. But there is only one such harmonic function, by the Maximum Principle, and it is the electrical potential φ.

□

Proof of Theorem 22. Let J be a flow on $\mathscr{G}$ with source a and sink b such that $J(a+) = I(a+)$, where I is the electrical current flow. Define a third flow H by taking the difference of I and J: note that this flow has no source and no sink. Write $J = I + H$ and expand the square in $\Delta^*(J,J)$ to get

$$\begin{aligned} 2\mathscr{D}(J,J) &= \sum_{\mathscr{E}^*} J(e(xy))^2 R_e \\ &= \sum_{\mathscr{E}^*}(I(e_{xy}) + H(e_{xy}))^2 R_e \\ &= \sum_{\mathscr{E}^*} I(e_{xy})^2 R_e + \sum_{\mathscr{E}^*} H(e_{xy})^2 R_e + \sum_{\mathscr{E}^*} I(e_{xy})H(e_{xy})R_e. \end{aligned}$$

To prove that I is the minimizer of the energy, it obviously suffices to show that the third sum is 0. By Ohm's Law, $I(e_{xy})R_e = \varphi(x) - \varphi(y)$; thus, by the Conservation Law for flows proved earlier in this section,

$$\begin{aligned}\sum_{\mathscr{E}^*} H(e_{xy}) I(e_{xy}) R_e &= \sum_{\mathscr{E}^*} H(e_{xy})(\varphi(x)-\varphi(y)) \\ &= (\varphi(a)-\varphi(b)) H(a+) \\ &= 0,\end{aligned}$$

the last equation because H has no source or sink.

That $\mathscr{D}(I,I)=\mathscr{D}^*(\varphi,\varphi)$ follows easily from Ohm's Law. □

The quantity $\mathscr{D}(I,I)$ is called the "total energy dissipated" in the network. It is what you are billed for by the electric company. The total energy dissipated can be expressed very simply in terms of the effective resistance between a and b.

Proposition 24. $\mathscr{D}(I,I)=\mathscr{D}^*(\varphi,\varphi)=I(a+)^2 R_{\text{eff}}(a,b)$.

Proof. That $\mathscr{D}^*(\varphi,\varphi)=\mathscr{D}(I,I)$ was proved above. By Ohm's Law and the conservation law for flows,

$$\begin{aligned}2\mathscr{D}(I,I) &= \sum_{\mathscr{E}^*} I(e_{xy})^2 R_e \\ &= \sum_{\mathscr{E}^*} I(e_{xy})(h(x)-h(y)) C_e R_e \\ &= \sum_{\mathscr{E}^*} I(e_{xy})(h(x)-h(y)) \\ &= 2(h(a)-h(b)) I(a+) \\ &= 2I(a+)^2 R_{\text{eff}}(a,b).\end{aligned}$$

□

Corollary 25. $R_{\text{eff}}(a,b)=\min\{\mathscr{D}(J,J) : J(a+)=J(b-)=1\}$.

5. Rayleigh's Monotonicity Law

The variational principles established earlier are important for various of reasons, not least among them a *comparison principle* for electrical networks due to Lord Rayleigh. It essentially states that if some or all of the resistances in an electrical network are increased then the effective resistance between any 2 vertices cannot decrease. We will only consider ordinary finite graphs, i.e., graphs with no multiple edges or loops.

Theorem 26. *(Rayleigh's Monotonicity Principle) Let R and $\overline{R}$ be two resistance functions on the same finite graph $\mathscr{G}=(\mathscr{V},\mathscr{E})$ satisfying $R_e \le \overline{R}_e$ $\forall e\in\mathscr{E}$. Then for all $a,b\in\mathscr{V}$,*

$$R_{\text{eff}}(a;b) \le \overline{R}_{\text{eff}}(a;b).$$

Proof. Instead of fixing the batteries so that the voltage differential between vertices a and b is 1, we choose voltages $v,\overline{v}>0$ so that the electrical current flows I and $\overline{I}$ satisfy $I(a+)=\overline{I}(a+)=1$. Note that multiplying the voltage by a scalar does not affect the validity of either variational principle.By the second variational principle, the current flows I and $\overline{I}$ minimize the energies

$\mathscr{D}(J < J)$ and $\overline{\mathscr{D}}(J,J)$ respectively, among all flows with source a, sink b, and satisfying $J(a+) = 1$. Recall that

$$\begin{aligned} 2\mathscr{D}(J,J) &= \sum_{\mathscr{E}^*} J(e(xy))^2 R_e \qquad \text{and} \\ 2\overline{\mathscr{D}}(J,J) &= \sum_{\mathscr{E}^*} J(e(xy))^2 \overline{R}(e). \end{aligned}$$

Therefore, for every flow J it must be the case that $\mathscr{D}(J,J) \leq \overline{\mathscr{D}}(J,J)$. It follows that $\mathscr{D}(I,I) \leq \overline{\mathscr{D}}(\overline{I},\overline{I})$. Hence, since $\mathscr{D}(I,I) = I_a^2 R_{\text{eff}}(a;b)$,

$$\begin{aligned} \mathscr{D}(I,I) &\leq \overline{\mathscr{D}}(I,I) \Longrightarrow \\ I_a^2 R_{\text{eff}}(a;b) &\leq \overline{I}_a^2 \overline{R}_{\text{eff}}(a;b) \Longrightarrow \\ R_{\text{eff}}(a;b) &\leq \overline{R}_{\text{eff}}(a;b). \end{aligned}$$

□

Corollary 27. *If some or all the resistances in an electrical network are increased,* except *those leading out of a, then $p_{\text{escape}}(a:b)$ can only decrease.*

This follows immediately from Rayleigh's monotonicity law and the characterization of $p_{\text{escape}}(a:b)$ in terms of effective resistance proved earlier.

Rayleigh's monotonicity law and its corollary can be effective tools in transience/recurrence problems. We shall give two examples of its use in proving the recurrence of certain reversible Markov Chains.

Remark 2. *Simple random walk* on $\mathbb{Z}^2$ is the reversible Markov chain on the integer lattice in 2 dimensions with resistance function R identically 1. Let the vertices of $\mathbb{Z}^2$ be arranged in concentric squares centered at the origin: thus

$$\begin{aligned} S_0 &= \{(0,0)\}; \\ S_k &= \{(m_1,m_2) : |m_1| = k \text{ and } |m_2| \leq k \text{ or } |m_2| = k \text{ and } |m_1| \leq k\}. \end{aligned}$$

We are interested in the probability that simple random walk started at the origin escapes S_k before returning to the origin. To obtain this probability, we modify the network as follows: replace ∂S_k by a single vertex s_*; any edges that connected vertices of S_k to vertices of ∂S_k should be rerouted to s_*; all vertices of $\mathbb{Z}^2$ not in $\cup_0^{k+1} S_j$ should be deleted from the network. The changes described above have no effect on the probability of escape from the origin to the exterior of S_k before returning to the origin. Note that after the modifications the exterior of S_k is just the single vertex s_*; consequently, the escape probability is given by

$$p_{\text{escape}} = \frac{1}{4R_{\text{eff}}}.$$

Here R_{eff} is the effective resistance between the origin and s_*.

To estimate the effective resistance, we will define a resistance function $\overline{R}$ that is dominated by R; by the corollary to Rayleigh's monotonicity law it will follow that $\overline{p_{\text{escape}}} \geq p_{\text{escape}}$. For each $l = 1,2,\ldots$, decrease the resistance between any two vertices of S_l to 0; i.e., "short" the vertices of S_l. Decreasing the resistance to zero is equivalent to increasing the conductance to ∞; by the Dirichlet Principle for the potential function, it must be the case that $\overline{h}$ is constant on each of the sets S_l (otherwise, the energy $\mathscr{D}^*(\overline{h},\overline{h})$ would be ∞). Now by the shorting law, to solve for the effective resistance between the origin and s_* we may replace each of the sets S_l by a single vertex s_l and reroute all the resistors accordingly.

The resulting network may be reduced using the parallel law, then again by the series law. First note that there are $8l-4$ resistors connecting s_{l-1} to s_l, so by the parallel law these may be replaced by a single resistor of resistance $1/(8l-4)$. Then by the series law, the series of vertices from s_0 (the origin) to s_k may be eliminated and the edges replaced by a single resistor of resistance $\sum_{l=1}^{k}(8j-4)^{-1}$ connecting s_0 to s_k. The result is that

$$\overline{R}_{\text{eff}} = \sum_{l=1}^{k}(8j-4)^{-1} \leq R_{\text{eff}}.$$

The last inequality is by Rayleigh's monotonicity law. It follows that as $k \to \infty$, the escape probability for the simple random walk decreases to zero. Consequently, simple random walk on $\mathbb{Z}^2$ is recurrent.

Remark 3. *Simple Random Walk on a Subgraph of* $\mathbb{Z}^2$. Let $\mathscr{G}$ be a subgraph of $\mathbb{Z}^2$: i.e., the vertex set and the edge set of $\mathscr{G}$ are contained in the vertex set and the edge set of the integer lattice $\mathbb{Z}^2$. Thus, $\mathscr{G}$ is obtained by removing edges from the integer lattice. Assume that $\mathscr{G}$ is connected, and that the origin e is in the vertex set of $\mathscr{G}$. The simple random walk on $\mathscr{G}$ is the reversible Markov chain corresponding to the electrical network on $\mathscr{G}$ in which every edge has resistance 1. A We will argue that the simple random walk on $\mathscr{G}$ is recurrent. To show this it suffices to show that the probability of escaping from e to the exterior of the sphere of radius k centered at e before returning to e decreases to zero as $k \to \infty$. As in the previous example, we estimate this probability by replacing the (possibly) infinite network $\mathscr{G}$ by the network in which the boundary of B_k is replaced by a single vertex s_k and all edges leading from B_k to ∂B_k are rerouted to s_k. Then the escape probability may be written in terms of the effective resistance R_{eff} between e and s_k. But the Rayleigh monotonicity law implies that the effective resistance is *no smaller* then that for the integer lattice $\mathbb{Z}^2$, because $\mathscr{G}$ is obtained from $\mathbb{Z}^2$ by increasing the resistances on certain edges (namely, those deleted) to ∞. We have shown that the simple random walk on $\mathbb{Z}^2$ is recurrent. Therefore, the effective resistance between e and s_k for the entire integer lattice increases to ∞ as $k \to \infty$. It follows (the Rayleigh monotonicity law) that the same is true for the network $\mathscr{G}$. This proves that the simple random walk on $\mathscr{G}$ is recurrent.

6. PROBLEMS

Problem 1. 1 Calculate the effective resistance and escape probability for a general birth and death chain on the nonnegative integers. Use this to give necessary and sufficient conditions for such a chain to be transient/recurrent.

Problem 2. 2 The *complete graph* on n vertices, denoted by K_n, is the graph with vertex set $[n] = \{1,2,3,\dots,n\}$ such that for every pair of vertices i,j there is an edge $e(\{ij\})$ with endpoints $i$and j. Consider the electrical network on K_n in which every edge is assigned resistance 1. Calculate the effective resistance between vertices 1 and n.

Problem 3. 3 Consider the graph $\mathscr{G}$ consisting of m copies K^i of K_4 strung out in series, with a single edge e^i connecting K^i to K^{i+1} for each $i = 1,2,\dots n-1$. Let a be any vertex of K^1 *except* the vertex adjacent to K^2, and let b be any vertex of K^n *except* the vertex adjacent to K^{n-1}. Calculate the escape probability $p_{\text{escape}}(a;b)$.

Problem 4. 4 Consider again the Ehrenfest Urn Scheme with a total of $2N$ balls. Let X_n be the number of balls in urn 1 at time n. Calculate the escape probability from N to $2N$, or more generally from $[\alpha N]$ to $[\beta N]$. If possible, find the limit, for fixed α and β, as $N \to \infty$.

Let $\mathscr{G}$ be a finite (ordinary) graph, and let $C_{..}$ be a conductance function on (the edges of) $\mathscr{G}$. Consider the following random process: At each vertex of $\mathscr{G}$ there is a button, which is either black or white. At each (discrete) time $n = 0,1,\dots$ an edge of the graph is chosen at random with probability C_{xy}/C, where C is the sum of the conductances in the network, and the buttons at x and y are interchanged. When a button reaches a it is immediately painted black, and when it reaches b it is immediately painted white.

Problem 5. 5 Let $h(x)$ be the voltage at x when a unit voltage is imposed between a and b (i.e., a 1 volt battery is hooked up to a and b). Show that regardless of the initial configuration of black and white buttons on the vertex set $\mathscr{V}$ the system reaches a statistical equilibrium in which $h(x)$ is the probability of a black button occupies vertex x.

Problem 6. 6 Let I be the electrical current flow in the network when a unit voltage is imposed between a and b. Show that there is a constant $\gamma > 0$ such that for every edge xy, the net flow of black buttons across the edge xy is γI_{xy}.

7. INFINITE NETWORKS:TRANSIENCE AND RECURRENCE

In this section let $\mathscr{G}=(\mathscr{V},\mathscr{E})$ be an *infinite* graph (no multiple edges) that is *locally finite*, i.e., such that every vertex has at most finitely many edges attached. (This assumption is actually extraneous, but allows us to avoid having to use any theorems from Hilbert space theory.) We assume as usual that the graph is connected. As for finite graphs, let $\mathscr{E}^*$ denote the set of *oriented* edges, and for any edge e with endpoints x,y denote by $e(xy)=e_{xy}$ the corresponding oriented edge with x and head y. Let $R_{xy}=R_{yx}$ be resistance attached to the edge e with endpoints x,y, and let $C_{xy}=1/R_{xy}$ be the corresponding conductance. Assume that the resistances are finite and strictly positive. Let X_n be the reversible Markov chain on $\mathscr{V}$ corresponding to the conductance function C_{xy}.

A *flow* on $\mathscr{G}$ with source a (and no sink) is an antisymmetric real-valued function $J_{xy}=J(e(xy))$ on the set $\mathscr{E}^*$ of directed edges that satisfies the flow condition (9) at every vertex x except $x=a$, where

$$J(a+)=\sum_{y\in\mathscr{V}}J_{ay}>0. \tag{25}$$

The quantity $J(a+)$ will be called the *size*of the flow. Define the *energy* of the flow J by

$$\mathscr{D}(J,J)=\frac{1}{2}\sum_{\mathscr{E}^*}J_{xy}^2R_{xy}. \tag{26}$$

Theorem 28. *If there exists a finite energy flow with a single source and no sink on $\mathscr{G}$ then the Markov chain X_n is transient.*

Theorem 29. *If X_n is transient, then for every vertex $a\in\mathscr{V}$ there exists a finite energy flow on $\mathscr{G}$ with source a.*

Theorem 28 sometimes gives an effective means of showing that a Markov chain is transient. To show that a Markov chain is transient it is always sufficient to show the existence of a nonconstant nonnegative superharmonic function; but in practice it is often easier to build flows than superharmonic functions.

7.1. Consequences.

Before beginning the proofs of these theorems, let us point out several consequences. First, the 2 theorems together give another approach to "comparisons" of different Markov chains on the vertex set $\mathscr{V}$. Specifically,

Corollary 30. *Let R_{xy} and $\overline{R}_{xy}$ be resistance functions on the graph $\mathscr{G}$ satisfying $R\le\overline{R}$. Let X_n and $\overline{X}_n$ be the corresponding Markov chains. If X_n is recurrent, then so is $\overline{X}_n$.*

Proof. If there is a finite energy flow for $\overline{R}$ then it is also a finite energy flow for R. □

This corollary may be used to give shorter proofs of some of the theorems proved in the preceding section using the Rayleigh principle.

As stated, Theorem 29 does not appear to be a useful too for establishing recurrence, because it requires that one show that every flow with source a and no sink has infinite energy. However, Theorem 29 has a simple corollary that *is* useful in showing that certain Markov chains are recurrent. Given a distinguished vertex a, say that a subset $\Gamma\subset\mathscr{E}$ is a *cutset* for the network if only finitely many vertices can be reached by paths started at a that contain no edges in Γ.

Example: In the integer lattice $\mathbb{Z}^2$, with the origin as distinguished vertex, the set Γ_m consisting of all edges connecting vertices (x,y) with $\max(x,y)=m$ to vertices (x',y') with $\max(x',y')=m+1$ is a cutset.

Theorem 31. *If there is a sequence of nonoverlapping cutsets Γ_m such that*

$$\sum_{m=1}^{\infty}\left(\sum_{e\in\Gamma_m} C(e)\right)^{-1}=\infty \tag{27}$$

then there is no finite energy flow with source a, and so the corresponding reversible Markov chain is recurrent.

Proof. For each m, let $\mathscr{V}_m$ be the (finite) set of vertices that can be reached from a via continuous paths with no edges in Γ_m; and let Γ_m^+ be the set of oriented edges $e(xy)$ such that $e\in\Gamma_m$ and such that $x\in\mathscr{V}_m$ and $y\notin\mathscr{V}_m$. Let J be a unit flow with source a and no sink (*unit flow* means that $J(a+)=1$). Then for every $m\geq 1$,

$$\sum_{e(xy)\in\Gamma_m^+} J(e(xy))=1$$

by an easy calculation (sum over the vertices $x\in\mathscr{V}_m$, and use the fact that J is a unit flow with source a). Consequently, by the Cauchy-Schwartz inequality,

$$\begin{aligned}\sum_{e(xy)\in\Gamma_m^+} J(e(xy))^2R(e)\sum_{e(xy)\in\Gamma_m^+} C(e) &\geq \left(\sum_{e(xy)\in\Gamma_m^+} J(e(xy))R(e)^{1/2}C(e)^{1/2}\right)^2\\ &=\left(\sum_{e(xy)\in\Gamma_m^+} J(e(xy))\right)^2=1,\end{aligned}$$

so

$$\sum_{e(xy)\in\Gamma_m^+} J(e(xy))^2R(e)\geq 1/\sum_{e(xy)\in\Gamma_m^+} C(e).$$

It follows that if (27) holds then the energy of the flow J must be infinite, because

$$\begin{aligned}2\mathscr{D}(J,J)&=\sum_{\mathscr{E}^*} J(e(xy))^2R(e)\\ &\geq\sum_{m=1}^{\infty}\sum_{e(xy)\in\Gamma_m^+} J(e(xy))^2R(e)\\ &\geq\sum_{m=1}^{\infty}(\sum_{e(xy)\in\Gamma_m^+} C(e))^{-1}.\end{aligned}$$

Therefore, by Theorem 29, the corresponding Markov chain is recurrent. □

7.2. **Examples.** First, we show how the Nash-Williams criterion gives an easy proof of the recurrence of two-dimensional random walks. Then we give two examples showing that the construction of finite-energy flows may be a feasible strategy for establishing transience.

7.2.1. *Nearest neighbor random walk on $\mathbb{Z}^2$.* Let C be any conductance function on the edges of the two-dimensional integer lattice $\mathbb{Z}^2$ that is bounded away from 0 and ∞, and let $X_n$be the corresponding reversible Markov chain. There is a natural sequence Γ_m of nonoverlapping cutsets, namely, Γ_m is the set of all edges connecting vertices vertices (x,y) with $\max(x,y)=m$ to vertices (x',y') with $\max(x',y')=m+1$. The cardinality of Γ_m is $8m+4$, so

$$\sum_{m=1}^{\infty}\left(\sum_{\Gamma_m} C(e)\right)^{-1} \geq \text{constant} \sum_{m=1}^{\infty}\frac{1}{8m+4}=\infty.$$

Therefopre, by the Nash-Williams criterion, the Markov chain X_n is recurrent.

7.2.2. *Nearest neighbor random walk on a homogeneous tree.* The homogeneous tree $\mathscr{T}_d$ of degree d is the tree in which every vertex has exactly d edges attached. For the sake of simplicity, take $d=3$. The tree $\mathscr{T}_3$ may be described as follows: the vertex set $\mathscr{V}$ is the set of all finite *reduced words* (including the empty word e) from the alphabet $\{a,b,c\}$, a *reduced word* being defined as a finite sequence of letters in which no letter (a,b, or c) follows itself. Two words v,w are neighbors iff one may be obtained from the other by attaching a single letter to the end of the other; note that every vertex has exactly 3 neighbors, e.g., aba has neighbors $ab, abab, abac$. The length of the word v is just its distance from the empty word e.

Let $C_{..}$ be any conductance function on the edges of the tree $\mathscr{T}_3$ that is bounded away from 0. Then the resistances are bounded above by a constant ρ. Using this fact, I'll construct a finite energy flow with source e. For any word $v\in\mathscr{V}$ of length ≥ 0 and any letter $i\in\{a,b,c\}$ such that vi is a reduced word, define

$$J_{v,vi}=\frac{1}{3}.\frac{1}{2^{|v|-1}}$$

where $|v|$ is the word length of v. It is clear that J is a flow with source e (see diagram), and

$$\mathscr{D}(J,J)\leq\rho\sum_{n=0}^{\infty}N_n(3^{-1}2^{-n+1})^2$$

where N_n is the number of edges in the tree leading from a word of length n to a word of length $n+1$. It is easily proved by induction that $N_n=(3)(2^{n-1})$. Consequently,

$$\mathscr{D}(J,J)<\infty.$$

7.2.3. *Simple NN RW on $\mathbb{Z}^3$.* For each vertex $x\in\mathbb{Z}^3$ let K_x be the unit cube with center x and edges parallel to the coordinate axes. Two vertices x,y are neighbors iff the cubes K_x,K_y have a common face K_{xy}. The resistances are assumed to be bounded above; thus the conductances are bounded away from 0. (For simple NN RW, the conductances are all 1.) I'll construct a finite energy flow. For neighbors x,y, define

$$J_{xy}=\int\int_{K_{xy}}\nabla\left(\frac{1}{r}\right)\cdot N_{xy}\,dS$$

where $r=$ distance to the origin, dS is the surface area element on the square K_{xy}, and N_{xy} is the unit normal on K_{xy} pointing in the direction x to y.

That J is a flow with source at the origin is an immediate consequence of the divergence theorem, because $\text{div}(\nabla(1/r))=0$ everywhere except at the origin (an easy computation). That J has finite energy may be seen as follows: First, there exists a constant $B < \infty$ such that $J_{xy} \leq Br^{-2}$, where $r = \min(|x|, |y|)$. Second, the number of vertices of the lattice $\mathbb{Z}^3$ at distance approximately n from the origin is on the order of n^2. Consequently, for suitable constants B', B'',

$$\begin{aligned} \sum_x \sum_y J_{xy}^2 &\leq \sum_{\text{nonzero vertices}} B' r^{-4} \\ &\leq \sum_{n=1}^{\infty} B'' n^2 n^{-4} \\ &< \infty. \end{aligned}$$

This shows that every reversible, nearest neighbor Markov chain on the 3D integer lattice for which the transition probabilities across nearest neighbor bonds are bounded away from 0 is transient.

7.3. **The Dirichlet Principle Revisited.** Let a be a vertex of the graph $\mathscr{G}$ and let Ω be a *finite* set of vertices containing a. Define $p_{\text{escape}}(a;\Omega)$ to be the probability that if started at a the Markov chain will escape from Ω before returning to a. We will characterize $p_{\text{escape}}(a;\Omega)$ in terms of flows.

Recall that the boundary $\partial\Omega$ of the set Ω consists of those vertices of $\mathscr{G}$ not in Ω that are nearest neighbors of vertices in Ω. Since Ω is finite, so is its boundary, because we have assumed that the graph $\mathscr{G}$ is locally finite. Define a flow on $\bar{\Omega} = \Omega \cup \partial\Omega$ with source a and sink $\partial\Omega$ to be an antisymmetric function J on the set of oriented edges of $\mathscr{G}$ that have at least one endpoint in Ω such that

(a) $\sum_{\mathscr{E}^+(x)} J_{xy} = 0 \; \forall \, x \in \Omega - \{a\}$;
(b) $\sum_{\mathscr{E}^+(a)} J_{ay} = C$.

Define the energy $\mathscr{D}_\Omega(J,J)$ of the flow J in the usual way, namely,

$$\mathscr{D}_\Omega(J,J) = \frac{1}{2}\sum J_{xy}^2 R_{xy},$$

where the sum extends over those directed edges of $\mathscr{G}$ with at least one endpoint in Ω; since there are only finitely many edges involved (recall that Ω is finite and because the graph is locally finite), the energy is always finite.

Proposition 32. *The escape probability p_{escape} is given by*

$$p_{\text{escape}}(a;\Omega) = \frac{1}{C_a \min \mathscr{D}_\Omega(J,J)}, \tag{28}$$

where the min is over all flows J of size 1 on $\bar{\Omega}$ with source a and sink $\partial\Omega$. Moreover, the flow of size 1 with source a that minimizes energy is

$$J_{xy} = \frac{C_{xy}(h(x) - h(y))}{C_a p_{\text{escape}}(a;\Omega)} \tag{29}$$

where h is the unique function defined on $\Omega \cup \partial\Omega$ that is harmonic in the set $\Omega - \{a\}$ and satisfies the boundary conditions

$$h(a) = 1 \qquad \textit{and} \tag{30}$$

$$h(y) = 0 \qquad \textit{for all } y \in \partial\Omega. \tag{31}$$

Proof. Define a new graph $\mathscr{G}_\Omega$ as follows:

(1) Let the vertex set $\mathscr{V}_\Omega = \Omega \cup \{b\}$, where b is a new vertex.
(2) Let the edge set $\mathscr{E}_\Omega$ consist of
  (a) all edges in the edge set $\mathscr{E}$ with both endpoints in Ω; and
  (b) one edge e_{new} with endpoints x and b for each edge e_{old} of $\mathscr{E}$ with one endpoint x and the other in $\partial\Omega$.

Define a conductance function C^Ω on this new graph $\mathscr{G}_\Omega$ by setting $C^\Omega(e) = C(e)$ for every edge e with both endpoints in Ω, and setting $C^\Omega(e_{\text{new}}) = C(e_{\text{old}})$ for every new edge e_{new} that corresponds to an edge e_{old} of $\mathscr{G}$ connecting a vertex in Ω to a vertex in $\partial\Omega$. Let $p^\Omega_{\text{escape}}(a;b)$ be the escape probability for the Markov chain on $\mathscr{G}_\Omega$ associated to the new conductance C^Ω; then clearly

$$p_{\text{escape}}(a;\Omega) = p^\Omega_{\text{escape}}(a;b).$$

But by Theorem 14 and Corollary 25,

$$p^\Omega_{\text{escape}}(a;b) = 1/(C_a R^\Omega_{\text{eff}}(a,b)) = 1/(C_a \min\{\mathscr{D}_\Omega(J,J) : J(a+) = J(b-) = 1\}$$

where the min is over all flows J in $\mathscr{G}_\Omega$ with source a, sink b, and size 1. □

7.4. **Proof of Theorem 28.** Assume that there is a finite energy flow I on $\mathscr{G}$ with source a. Without loss of generality we may assume that this is a *unit flow*, i.e., that $I(a+) = 1$. We will show that

$$P^a\{\text{return to a}\} < 1.$$

Let $\Omega_1 \subseteq \Omega_2 \subseteq \dots$ be finite subsets of $\mathscr{V}$, each containing a, whose union is all of $\mathscr{V}$. Clearly the escape probabilities $p_{\text{escape}}(a;\Omega_n)$ are monotone, and converge to the probability of no return to a. Thus, it suffices to prove that the escape probabilities are bounded away from 0.

By the Dirichlet Principle (the last corollary) the escape probabilities are given by

$$p_{\text{escape}}(a;\Omega_n) = \frac{1}{C_a \min \mathscr{D}_{\Omega_n}(J,J)},$$

where the min is over all unit flows on Ω_n with source a. But this min is clearly no larger than the total energy of the flow I, because restricting I to Ω_n gives a unit flow on Ω_n with source a, and the energy of the restriction is no larger than the total energy of I. Therefore,

$$p_{\text{escape}}(a;\Omega_n) \geq \frac{1}{C_a \mathscr{D}(I,I)} > 0$$

for all n. □

7.5. **Proof of Theorem 29.** Assume that the Markov chain is transient. We will show that, for every vertex $a \in \mathscr{V}$ such that the probability of no return to a is positive, there is a finite energy flow with source a. Let $\Omega_1 \subseteq \Omega_2 \subseteq \dots$ be finite subsets of $\mathscr{V}$, each containing a, and whose union is all of $\mathscr{V}$. Define

$$\begin{aligned} h_n(x) &= P^x\{\text{hit } a \text{ before exiting } \Omega_n\}; \\ h(x) &= P^x\{\text{hit } a \text{ eventually}\}. \end{aligned}$$

Observe that h is harmonic at every $x \in \mathscr{V} - \{a\}$, and that h_n is harmonic at every $x \in \Omega_n - \{a\}$; also, for every x, $h_n(x) \to h(x)$ as $n \to \infty$. Observe that the function h_n is the same function as appearing in the statement of Theorem 21 .

Define flows $I^{(n)}$ on Ω_n and I on $\mathscr{G}$ with source a by

$$\begin{aligned} I^{(n)}_{xy} &= C_{xy}(h_n(x)-h_n(y))/C_a p_{\text{escape}}(a;\Omega_n); \\ I_{xy} &= C_{xy}(h(x)-h(y))/C_a p_* \end{aligned}$$

where $p_* =\downarrow \lim p_{\text{escape}}(a;\Omega_n) = P^a\{\text{no return to } a\} > 0$. The flows $I^{(n)}, I$ have source a and no other sources or sinks (by the harmonicity of h_n and h off a). Moreover, $I^{(n)}_{xy} \to I_{xy}$ as $n \to \infty$, for every directed edge xy.

By Proposition 24, the flows $I^{(n)}$ minimize energy among all unit flows on Ω_n with source a, and

$$p_* \le p_{\text{escape}}(a;\Omega_n) = \frac{1}{C_a \mathscr{D}_{\Omega_n}(I^{(n)}, I^{(n)})}.$$

It follows that for every n,

$$\mathscr{D}_{\Omega_n}(I^{(n)}, I^{(n)}) \le \frac{1}{C_a p_*} < \infty.$$

But the energy of I is given by

$$\begin{aligned} \mathscr{D}(I,I) &= \sum_x \sum_y I^2_{xy} R_{xy} \\ &= \sum_x \sum_y \lim_{n\to\infty} (I^{(n)}_{xy})^2 R_{xy} \\ &\le \liminf_{n\to\infty} \sum_{x\in\Omega_n} \sum_y (I^{(n)}_{xy})^2 R_{xy} \\ &= \liminf_{n\to\infty} \mathscr{D}_{\Omega_n}(I^{(n)}, I^{(n)}) \\ &\le \frac{1}{C_a p_*}, \end{aligned}$$

by the Fatou Lemma. Thus, the flow I has finite energy. □