

Random Walks and Electric Networks

Rich Schwartz

April 11, 2022

Most of the things in these notes can be found in the book *Random Walks and Electric Networks*, by Peter Doyle and Laurie Snell. I highly recommend this excellent book. These notes cover some of the main points in the book, but they do not always do things as they are in the book.

1 Harmonic Functions

Let G be a graph. The function $f : G \rightarrow \mathbf{R}$ (defined on the vertex set of G) is called *harmonic* at a vertex v if

$$f(v) = \frac{1}{k} \sum_{i=1}^k f(w_i). \quad (1)$$

Here k is the degree of v and w_1, \dots, w_k are the vertices incident to v . Notice the similarity between this equation and Equation 9.

Suppose that C is a non-empty and proper subset of vertices of G . Suppose that f is some function on C . We call the function $F : G \rightarrow \mathbf{R}$ a *harmonic extension* of f if $F = f$ on C and F is harmonic on $G - C$. We will show that each f has a unique harmonic extension.

Lemma 1.1 *If F_1 and F_2 are two harmonic extensions of f then $F_1 = F_2$.*

Proof: Let $g = F_1 - F_2$. Note that g is a harmonic extension of the 0 function on C . There must be some vertex v where g achieves its maximum. But $g(v)$ is the average of the values of g at the vertices incident to v . This is only possible if g also takes its max at all the vertices incident to v . Continuing outward from v , we see that g must take its max everywhere. But then this common value must be 0. ♠

Lemma 1.2 *Every function f on C has a harmonic extension.*

Proof: Let V be the vector space of functions defined on $G - C$. Let $\Delta_f : V \rightarrow V$ be the map

$$\Delta_f g(v) = v - \frac{1}{k} \sum g(w_i). \quad (2)$$

As usual k is the degree of v and w_1, \dots, w_k are the vertices incident to v . The map Δ_f is an affine transformation from V into V . (It is matrix multiplication followed by translation which depends on the function f .)

Suppose that $\Delta_f(F_1) = \Delta_f(F_2)$. We extend g to that $g = 0$ on C . Consider $g = F_1 - F_2$. Then, at each vertex $v \in G - C$, the value $g(v)$ is the average of its neighbors. But then the same argument as above shows that g cannot have a nonzero maximum or a nonzero minimum. Hence g is identically 0. (This step uses the fact that C is nonempty, so that $g = 0$ at some vertex.)

We have shown that Δ_f is one to one. But then Δ must be onto. In particular, the 0 function on $G - A - B$ is in the image of Δ_f . That is, there exists $F \in V$ such that $\Delta_f(F)$ is the 0 function. By definition F is a harmonic extension of F . ♠

2 The Electric Flow

We keep the same notation as above. Now we set $C = a \cup b$ where a and b are vertices of G . Let V be a harmonic function on $G - a - b$. By the previous result, the values $V(a)$ and $V(b)$ uniquely determine V . We call V the *voltage function* induced by $V(a)$ and $V(b)$. This function has two nice properties:

- If V_1 and V_2 are voltage functions then so is $V_1 + \lambda V_2$ for any constant λ . In other words, the voltage functions form a vector space.
- If $V(a) = V(b)$ then V is constant.

Let $i_{xy} = V(x) - V(y)$. We call i_{xy} the *electric current* induced by the voltage. For each vertex x let

$$i_x = \sum_{y \leftrightarrow x} i_{xy}. \quad (3)$$

The sum is taken over adjacent vertices. The quantity i_x is the total current flowing out of x . Since V is harmonic on $G - a - b$, we have $i_x = 0$ unless $x = a$ or $x = b$.

Lemma 2.1 $i_a + i_b = 0$.

Proof: We have

$$i_a + i_b = \sum_x i_x = \sum_x \sum_y i_{xy} = 0 \quad (4)$$

The last sum is 0 because $i_{xy} = -i_{yx}$ and we are summing over all pairs. ♠

We define

$$R(G, a, b) = \frac{V(a) - V(b)}{i_a} = \frac{V(b) - V(a)}{i_b}. \quad (5)$$

Put another way, $R(G, a, b)$ is the value such that $i_a = 1$ when

$$V(a) = R(G, a, b), \quad V(b) = 0.$$

Intuitively, $R(G, a, b)$ is large when you have to have a huge voltage differential to induce a unit current. The quantity $R(G, a, b)$ is called the *effective resistance* of G with respect to (a, b) .

3 Energy and Resistance

We keep the same notation as above. Say that a *nice flow* is a function $j : G \times G \rightarrow \mathbf{R}$ such that

1. $j_{xy} = 0$ if x and y do not share an edge.
2. $j_{xy} = -j_{yx}$.
3. If $x \in G - a - b$ then $j_x = 0$. That is, the amount of current flowing into x is the same as the amount flowing out. Here $j_x = \sum j_{xy}$.

If, additionally, $j_a = 1$, we call j_a a *unit flow*. Finally, if $j_a = 0$ we call j a *null flow*.

The nice flows form a vector space \mathbf{V} . The electric flows considered above are examples of nice flows but there are typically many more nice flows than there are electric flows. We'll see examples in the next section.

Here is a technical lemma whose use will become clear in the corollary below.

Lemma 3.1 For any function $\phi : G \rightarrow \mathbf{R}$ and any nice flow ℓ ,

$$\ell_a(\phi(a) - \phi(b)) = \frac{1}{2} \sum_{xy} \ell_{xy}(\phi(x) - \phi(y)).$$

Proof: Both the left hand side and the right hand side are linear functions of ϕ . So, to prove this equality it suffices to prove it on a basis in the vector space of such functions. Suppose that $\phi(a) = 1$ and $\phi(x) = 0$ otherwise. Then the two sides are equal just by definition. Suppose that $\phi(x) = 1$ for some $x \in G - a - b$. Then the left side is obviously 0, and the right side is simply ℓ_x , which is 0. Finally, consider the case when ϕ is identically 1. Then both sides vanish. We've established the identity on a basis, so we're done. ♠

There is a canonical inner product on \mathbf{V} , namely

$$\langle j, k \rangle = \frac{1}{2} \sum_{xy} j_{xy} k_{xy}. \quad (6)$$

The factor of $1/2$ comes from the fact that we're summing over ordered pairs, so each edge appears twice. The quantity $\langle j, j \rangle$ is called the *energy* of j .

Corollary 3.2 Let i be the unit electric flow, j another unit flow, and $k = j - i$. Note that k a null flow. This section is devoted to proving 3 statements:

1. $R(G, a, b) = \langle i, i \rangle$.
2. $\langle i, k \rangle = 0$.
3. $\langle i, i \rangle \leq \langle j, j \rangle$.

Proof: Applying Lemma 3.1 to the case $\ell = i$ and $\phi = V$ (the voltage function) we have

$$\langle i, i \rangle = \frac{1}{2} \sum_{xy} i_{xy} i_{xy} = \frac{1}{2} \sum_{xy} i_{xy} (V(x) - V(y)) =$$

$$i_a(V(a) - V(b)) = R(G, a, b).$$

This proves Statement 1.

Applying Lemma 3.1 to the case $\ell = k$ and $\phi = V$ the same calculation gives

$$\langle k, i \rangle = k_a(V(a) - V(b)) = 0.$$

This proves Statement 2.

Finally,

$$\langle j, j \rangle = \langle i + k, i + k \rangle = \langle i, i \rangle + \langle i, k \rangle + \langle k, i \rangle + \langle k, k \rangle = \langle i, i \rangle + \langle k, k \rangle \geq \langle i, i \rangle.$$

Here we have used the fact that $\langle i, k \rangle = \langle k, i \rangle = 0$ by Statement 2 and symmetry. This proves Statement 3. ♠

Statement 3 of the Corollary says that the unit electric flow minimizes the energy amongst all nice unit flows.

4 Moves on Graphs

Here we describe some rules for how the effective resistance changes when we modify our graph in certain ways.

4.1 Series Law

Suppose $c \in G - a - b$ is a cut vertex of G . Let G_a be the lobe of G which contains a and let G_b be the lobe of G which contains b .

Lemma 4.1 $R(G, a, b) = R(G_a, a, c) + R(G_b, c, b)$.

Proof: Let V be the voltage function on G so that the electric current is the unit flow. Then

$$\begin{aligned} R(G, a, b) &= V(a) - V(b) = \\ &= (V(a) - V(c)) + (V(c) - V(b)) = R(G_a, a, c) + R(G_b, c, b). \end{aligned}$$

What makes this work is that the electric flow out of c is 0. So, by Lemma 2.1 we get the following results: in G_a there is a unit amount of current flowing into c and in G_b there is a unit amount of current flowing out of b . ♠

4.2 Parallel Law

Suppose that $\{a, b\}$ is a cut set for G and that G is the union of two lobes joined together at $a \cup b$. Call these lobes G_1 and G_2 . Let $C = 1/R$. The quantity $C(G, a, b)$ is often called the *effective capacity* of G with respect to a, b .

Lemma 4.2 $C(G, a, b) = C(G_1, a, c) + C(G_2, c, b)$.

Proof: Let V be the voltage function for G , chosen so that $V(a) = 1$ and $V(b) = 0$. Then the total current out of a is $C(G, a, b)$. But the restriction V_j to G_j is the voltage function on G_j normalized so that $V_j(a) = 1$ and $V_j(b) = 0$. Hence, the total current out of a in G_j is $C(G_j, a, b)$. But the total current out of a in G is the sum of the total current out of a in G_1 and the total current out of a in G_2 . ♠

4.3 Adding Edges

Here is Rayleigh's Theorem.

Theorem 4.3 (Rayleigh) *Let G' be a graph obtained from G by adding an edge. Then $R(G', a, b) \leq R(G, a, b)$.*

Proof: Let i be the unit electric flow on G . Let i' be the unit electric flow on G' . We can interpret i as a nice unit flow on G' , though possible $i \neq i'$. By Statements 1 and 3 proved in the previous section,

$$R(G', a, b) = \langle i', i' \rangle \leq \langle i, i \rangle = R(G, a, b).$$

We're done. ♠

Of course, we can iterate this to obtain the same result when we add finitely many edges.

4.4 Collapsing Edges

Now suppose e is an edge of G whose endpoints are not in $\{a, b\}$. Let G/e be the graph obtained by collapsing e .

Lemma 4.4 $R(G/e, a, b) \leq R(G, a, b)$.

Proof: Let i and \bar{i} respectively be the unit electric flows on G and G/e . Let i' be the flow on G/e which equals i on each edge of $G - e$.

Let x, y be the endpoints of e . Let z be the vertex in G/e obtained by identifying x and y . We have $i'_z = i_x + i_y = 0$. Hence i' is a unit nice flow on G/e . We have

$$R(G/e, a, b) = \langle \bar{i}, \bar{i} \rangle \leq \langle i', i' \rangle = \langle i, i \rangle = R(G, a, b).$$

This completes the proof. ♠

This result may be iterated, showing that the resistance does not increase when we collapse finitely many vertices of $G - a - b$ to a single vertex.

4.5 Subgraphs

Let $H \subset G$ be a connected subgraph which contains the vertices a, b . Here we prove that $R(G, a, b) \leq R(H, a, b)$. We can find a finite sequence of graphs

$$H = G_0, \dots, G_n = G$$

such that each G_i is a subgraph of G_{i+1} and one of two things is true: Either G_{i+1} is obtained by adding an edge to G_i , or G_{i+1} is obtained by adding a vertex v and connecting it to some vertex of G_i by a new edge e .

In the first case, we have already proved that $R(G_i, a, b) \geq R(G_{i+1}, a, b)$. In the second case, we can extend the voltage function on G_i to the voltage function on G_{i+1} by defining $G(v) = G(v')$ where v' is the other endpoint of e . This gives the voltage function on G_{i+1} , and the two voltage functions define the same flow. So, in this case $R(G_i, a, b) = R(G_{i+1}, a, b)$. Stringing together all the inequalities gives us our result.

5 Four Examples

In this section we use the graph moves above to analyze 4 examples. Before we give these examples, we mention that we can speak about effective resistance when a and b are replaced by disjoint subsets A and B of vertices. We define $G_{A,B}$ be the graph obtained by identifying all the vertices of A to a single vertex and all the vertices of B to a single vertex. We call these new vertices a and b and then we proceed as above. That is, we set $R(G, A, B) = R(G_{A,B}, a, b)$.

Paths: Suppose that P_n is the path with n edges. Let v_0, \dots, v_n be the vertices of G . Let $a = v_n$ and $b = v_0$. Let $A = \{v_n\}$ and $B = \{v_0\}$. We define $V(v_j) = j/n$. Then V exactly the voltage function, because $i_a = 1$ and $V(i+1) - V(i) = 1/n$ for all i . Hence $R(P_n, a, b) = 1/n$. Taking a limit as $n \rightarrow \infty$, we could say that the effective resistance of the infinite path with $a = 0$ and $b = \infty$ is 0.

Rooted Binary Trees: Let G_n be n -generations of the rooted binary tree. So, the leaves of G have distance n from the root. The root vertex a has degree 2 and all other non-leaves of G have edgree 3. Let B_n be the set of leaves of G_n . Given a vertex v we let

$$V_n(v) = 2^{-d} - 2^{-n}$$

where d is the distance from v to a . For instance $V(a) = 1 - 2^{-n}$, and $V = 1/2 - 2^{-n}$ on the two neighbors of a . The total current flowing out of a is 1 and you can check that V is harmonic at all vertices of G except a . Also, V takes the same value on all the leaves of G_n . So, V gives rise to a harmonic function on $(G_n)_{a, B_n}$. Hence $E(G_n, a, B_n) = 1 - 2^{-n}$. Taking a limit, we would say that the effective resistance of the infinite binary tree (starting at the root) is 1.

The Infinite Square Grid: Let G_∞ denote the usual infinite graphs of edges of the unit square tiling of the plane. Let G_n be the subgraph of G_∞ consisting of the edges of the $(2n) \times (2n)$ square grid centered at the origin. Let $a = \{0, 0\}$ and let B_n denote the outer cycle of G_{2n} .

Now we prove that

$$\lim_{n \rightarrow \infty} R(G_n, a, B_n) = \infty. \tag{7}$$

Consider the concentric cycles B_1, B_2, \dots, B_n . These cycles are pairwise disjoint. Let G'_n denote the graph obtained by collapsing each of these cycles to a point. The graph G'_n looks like the right side of Figure 1. Figure 1 shows the case $n = 2$. Here we have replaced the multiple edges by a number which indicates the number of edges connecting each vertex to the next.

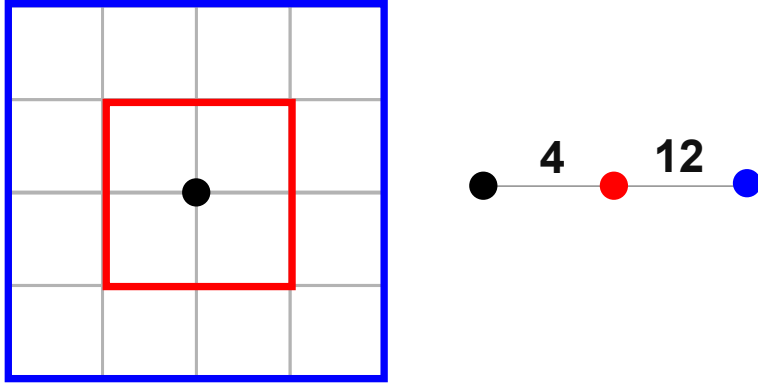


Figure 1: Collapsing the cycles

The collapse of G_n is just a path of length n whose effective resistances are $1/(4 \times 1)$, $1/(4 \times 3)$, $1/(4 \times 5)$, etc. These graphs all appear in parallel. So, by Rayleigh's Theorem, the total effective resistance is

$$R(G_n, A, B_n) \geq \frac{1}{4} \times \left(\frac{1}{1} + \frac{1}{3} + \dots + \frac{1}{2n+1} \right) \quad (8)$$

This series is closely related to the harmonic series, and is easily seen to diverge. This completes the proof. One could say that the resistance to ∞ in the infinite square grid is infinite.

The Infinite Cubical Grid: Now let G_∞ denote the infinite cubical graph in \mathbf{R}^3 . Now we will prove that the resistance from the origin to ∞ is finite in G_∞ . To avoid quite a tedious argument we will not bother making the intermediate finite graph constructions. We're just going to go right to the limit, with the understanding that (with some care) the process could be truncated and we could take an honest limit as in previous cases. Alternatively, it is better to just define effective resistance for infinite graphs.

We are going to build an infinite tree $T_\infty \subset G_\infty$ in layers. By Rayleigh's Theorem, we have

$$R(G_\infty, a, \infty) \leq R(T_\infty, a, \infty).$$

Here ∞ essentially denotes the limit of the (presumed) shells B_n in the finite approximations.

We will build T_∞ in layers. For each $k = 0, \dots, n$, define

$$\Phi_k = \{(100\ i, 100\ j, 1000\ 2^k) \mid i, j \in \{1, \dots, 2^k\}\}$$

The slope of any line segment joining a point of Φ_{k-1} to a point of Φ_k is at most 1. So, we can join each point in Φ_{k-1} to 4 points (making a square) in Φ_k by paths of length at most $C2^k$. So, the point p_{ij} in Φ_{k-1} gets joined to the points in square ij in Φ_k .

This is our tree, but the paths we have used are somewhat irregular. We lengthen some of the paths in T_∞ by adding extra vertices so that between Φ_{k-1} and Φ_k , all edges have length 10000×2^k . This process only increases the resistance, by Rayleigh's Theorem. (You have to think a bit about how Rayleigh's Theorem applies in this case.) The new tree T'_∞ is no longer a subgraph of G_∞ but we don't care. What we know is that

$$R(T_\infty, a, \infty) \leq R(T'_\infty, a, \infty).$$

We just have to bound this latter quantity.

Why did we add these edges. Well, by symmetry the voltage function on T'_∞ takes the same value on all vertices that are the same distance from the initial node! So, we can collapse all the vertices at the same distance from the initial node and we get a graph with *the same resistance*. The collapsed graph is just a "path" consisting of 4^k paths, each having length 10000×2^k . The effective resistance of this "path" is

$$\frac{10000 \times 2^k}{4^k} = 10000 \times 2^{-k}.$$

Summing this, we see that the effective resistance to ∞ is at most 20000×2^{-k} . This completes the proof.

So far we have been talking entirely about electric networks, but now we turn to the subject of random walks. In a certain sense, our analysis of the resistance of the square grid and the cubical grid gives a proof of Polya's famous theorems about recurrence of the standard random walk in \mathbf{Z}^2 and the transience of the standard random walk on \mathbf{Z}^3 . What remains to do is to relate what we have done above to the theory of random walks, and then interpret our resistance results in terms of probability.

6 Random Walks on Graphs

Let G be a graph in which every vertex has finite degree. We usually take G to be a finite graph, but sometimes we will consider countable graphs. Suppose that we have given a linear ordering to the edges incident to each vertex of G . A *random walk* on G , starting from the vertex $v \in G$, is a sequence of fair coin flips b_1, b_2, b_3, \dots where the (j) th coin has as many sides as the degree of the vertex v_j . The value of b_1 selects the vertex v_2 adjacent to v , the value of b_2 selects the vertex v_3 adjacent to v_2 , and so on. Here we have set $v_1 = v$. One funny thing about this process is that the number of sides of the coin can vary from vertex to vertex. If you prefer, one can consider random walks on regular graphs, and then one can use the same coin all the time.

Suppose that A and B are two disjoint subsets of vertices. We define $P(v, A, B)$ to be the probability that a random walk starting from v reaches A before it reaches B . Some of you might be satisfied that this notion of probability makes intuitive sense. In this case, just skip the next section. Otherwise, you can read a quick sketch of how this probability is defined in terms of measure theory.

In any case, we are really only going to use a few basic properties of the above function. Suppose that w_1, \dots, w_k are the vertices incident to v and $v \in G - A - B$. Then the basic property is

$$P(v, A, B) = \frac{1}{k} \sum_{i=1}^k P(w_i, A, B). \quad (9)$$

In other words, we have an equal chance of going from v to each w_i , and then we can compute the probability of hitting B before A and just average these probabilities.

Let us relate random walks to electric networks. We have seen already that every function on $A \cup B$ has a unique harmonic extension. Suppose that we set $f = 1$ on A and $f = 0$ on B . Let's define

$$F(v) = P(v, A, B). \quad (10)$$

Note that $F = f$ on $A \cup B$. Equation 9 tells us that F is harmonic for all $v \in G - A - B$. In other words, the probability function $v \rightarrow P(v, A, B)$ we have been considering in previous sections is the unique harmonic extension of the function which is 1 on A and 0 on B .

7 Polya's Theorems

Let G_∞ be the infinite square grid. Now we explain why a random walk on G_∞ , starting at the origin a , returns to the origin with probability 1. Let G_n be the graph made from n consecutive layers of G_∞ . Let B_n denote the outermost layer of G_n . Let P_n denote the probability that a random walk starting at the origin a returns to the origin before hitting B_n . The probability that a random walker returns to the origin in G_∞ is at least as big as the probability that a random walker on G_n hits B_n before returning to a , because in the latter case the process stops and in the former case the process continues and the walker has more changes to return home.

Think about it this way. Suppose we lived in a crazy country where the police throw you in prison if you wander more than 10 miles from your house. Your chances of returning home are increased if the police just disappear. The boundary B_n is basically the prison that you land in if you wander too far from home.

Now we estimate the probability that a random walk starting at a hits B_n before returning to a . Since every random walk starting at a must hit a neighbor of a , the probability that a random walk starting at a returns to the origin is the average of the probabilities that a random walk starting at each of the neighbors of a hits a before hitting B_n . Let b be such a neighbor. The probability in question is $F_n(b)$, where F_n is the harmonic extension on G_n of the function which is 1 on a and 0 on B_n .

Here

$$F_n = \frac{V_n}{R_n}, \quad R_n = R(G_n, a, B_n),$$

where V_n is the voltage function. Since the total current flowing out of G_n with respect to the potential V_n is at most 1, we have

$$V_n(b) \geq R_n - 1.$$

But then

$$F_n(b) \geq \frac{R_n - 1}{R_n}.$$

Since $R_n \rightarrow \infty$ as $n \rightarrow \infty$ we see that $F_n(b) \rightarrow 1$. This completes the proof.

Now let's consider the situation for the cubical grid G_∞ . This time the resistance out to infinity is finite. So, there is a nonzero harmonic extension

of the function which is 1 on a and tends to 0 as one tends to ∞ in G_∞ . So, interpreting this fact probabilistically, we see that the probability that the probability that a random walk starts at a and returns to a is less than 1. Hence there is a positive probability that the random walk starts at a and never returns.

8 Measure Theoretic Aside

Here is the way the probability $P(v, A, B)$ is treated from the standpoint of measure theory. Again, if you are happy with the definition already, just skip this part of the notes.

In general, the set of all possible coin flips is the subset $S(G, v)$ of infinite allowable integer sequences. A sequence is *allowable* if, for all j , the j th digit b_j does not exceed the degree of the vertex v_j selected by the previous terms b_1, \dots, b_{j-1} . For regular graphs of degree d , the set $S(G, v_0)$ is just the set of all infinite sequences involving d digits.

We let

$$S = S(G, v) \tag{11}$$

Supposing that we have chosen initial allowable sequence $\beta = (b_1, \dots, b_n)$, the *cylinder set* C_β is the set of all allowable infinite sequences which have start with β . The probability that a finite random walk of length n starting at v will produce β is

$$|C_\beta| = \frac{1}{d_1 \dots d_n}, \tag{12}$$

where d_j is the degree of v_j .

In general, one defines the *outer measure* of a subset $E \subset S$ as follows.

$$\mu^*(E) = \inf_{\mathcal{C}} \sum_{C \in \mathcal{C}} |C|. \tag{13}$$

Here \mathcal{C} is a covering of S by cylinder sets and $|C|$ denotes the probability that C occurs, as above. What we are doing is taking the infimum over all possible coverings.

A subset E is called *measurable* if

$$\mu^*(X - E) + \mu^*(X \cap E) = \mu^*(X) \tag{14}$$

for all other subsets $X \subset S$. In this case, we define $\mu(E) = \mu^*(E)$. This definition looks insane, because it is something we would have to test for *all* other subsets. However, one can check several basic properties:

- Cylinder sets themselves are measurable. This is pretty easy. If we have a covering of X we get coverings of $X \cap E$ just by intersecting the cover with E and we get a covering of $X - E$ by intersecting our covering with $S - E$, which is a finite union of cylinder sets.
- If E is measurable so is $S - E$. This just follows straight from the definition.
- The countable union of measurable sets is again measurable. This is a bit more work, but not too bad.

A set is called a *Borel set* if it is obtained by starting with cylinder sets and taking complements and countable unions finitely many times. From the three properties above, Borel sets are measurable.

The function μ is called a *Borel measure*. The pair (S, μ) is called a *probability space*. Subsets of S are often called *events*. When E is a Borel event, $\mu(E)$ is the probability that E occurs. The measure μ is *countably additive*: If E_1, E_2, E_3, \dots is any countable collection of pairwise disjoint measurable subsets of S , then

$$\mu\left(\bigcup E_j\right) = \sum \mu(E_j). \quad (15)$$

The set $E(v, A, B)$ of random walks which start at v and hit A before B is one of these Borel sets. We define

$$P(v, A, B) = \mu(E(v, A, B)). \quad (16)$$

Let us sketch the derivation of Equation 9. Let $E = E(v, A, B)$ and also define $E_j = E(w_j, A, B)$. Note that E_j is a subset of the space $S(w_j, A, B)$. Finally let $E'_j \subset S$ denote the set of random walks which first go to w_j and then lie in E_j . We have

$$\mu(E_j) = (1/k) \mu_j(E'_j) \quad (17)$$

Here μ_j is the measure on $S(G, w_j)$. Also, from Equation 15 (in the finite case) we have $\mu(E) = \sum_j \mu(E_j)$. Putting these two facts together gives Equation 9.