ORIE 6334 Spectral Graph Theory	October 4, 2016
Lecture 13	
Lecturer: David P. Williamson	*Scribe: Andrew Daw*

In this lecture we continue the investigation into networks of electrical flows from the previous lecture, with a continued focus on effective resistance. This will lead us to findings that are not only interesting in their own right, but also allow us to pick back up our discussion of random walks with a freshness of mind. Specifically, we will define hitting times, commute times, and cover times and use what we have learned from effective resistance to get a better understanding of these quantities.

1 Effective Resistance Continued

Recall from the last lecture that the effective resistance between i and j, $r_{\text{eff}}(i,j)$, is the potential drop between i and j induced by an i-j electrical flow. For such a flow, the corresponding potential p is such that $L_G p = e_i - e_j$, where L_G is the Laplacian of the graph of the electrical network. Let $L_G^\dagger$ be the psuedo-inverse of that matrix. Then, $p = L_G^\dagger(e_i - e_j)$ and

$$r_{\text{eff}}(i,j) = (e_i - e_j)^T L_G^\dagger (e_i - e_j)\,.$$

We will now show a way to calculate flow from spanning trees. Let $\mathcal{T}$ be the set of all spanning trees in G. For a tree $T \in \mathcal{T}$, let

$$r(T) = \prod_{\{i,j\}\in T} r(i,j)$$

where $r(i,j)$ is the resistance on edge $\{i,j\}$. Additionally, let

$$Z = \sum_{T\in\mathcal{T}} \frac{1}{r(T)}\,,$$

which we will use as a normalizing factor. Now, again for a spanning tree T, define a flow f_T by

$$f_T(i,j) = \begin{cases} 1 & \text{if } (i,j) \text{ is on the directed } s\text{-}t \text{ path in } T \\ -1 & \text{if } (j,i) \text{ is on the directed } s\text{-}t \text{ path in } T \\ 0 & \text{otherwise} \end{cases}$$

and note that this definition allows us to maintain skew symmetry.

Theorem 1 *Let f be defined*

$$f = \frac{1}{Z}\sum_{T\in\mathcal{T}} \frac{1}{r(T)} f_T\,.$$

Then f is an s-t electrical flow.

[0]This lecture is derived from Bollobás, *Modern Graph Theory*, II.1; Lau's Lecture 12, `https://cs.uwaterloo.ca/~lapchi/cs798/notes/L12.pdf`; and Motwani and Raghavan, *Randomized Algorithms*, Sections 6.3-6.5.

Proof: First, note that by the definition of each f_T we know that f sends $\frac{1}{Z}\sum_{T\in\mathcal{T}}\frac{1}{r(T)} = \frac{Z}{Z} = 1$ unit of current from s to t. Clearly flow is conserved throughout the network, meaning that the flow into each node is equal to the flow leaving that node, and so Kirchoff's Current Law is obeyed.

Now, we want to show that Kirchoff's Power Law (KPL) is obeyed as well by f, and this will show that f is an s-t electrical flow. Recall that KPL states that for any directed cycle C, $\sum_{(i,j)\in C} r(i,j)f(i,j) = 0$.

Let $\mathcal{S}$ be all s-t cuts S such that $s \in S$ and $t \notin S$, the graph induced on S is connected, and the graph induced on $V - S$ is as well. Then,

$$f(i,j) = \frac{1}{Z}\sum_{S\in\mathcal{S}}\sum_{\substack{T:\delta(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)}f_T(i,j),$$

as this sums over all trees that contribute flow to f on (i,j) - in all cases other than the intersection of the s-t cut and the spanning tree being (i,j), the flow on that arc must be 0 by definition. Now, we can split this summation into the arcs headed into the cut and arcs headed out of the cut. To do so, let $\delta^+(S)$ be the set of arcs directed out of S into $V - S$, and let $\delta^-(S)$ be the set of arcs directed into S from $V - S$. Then,

$$f(i,j) = \frac{1}{Z}\sum_{S\in\mathcal{S}}\left(\sum_{\substack{T:\delta^+(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)} - \sum_{\substack{T:\delta^-(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)}\right).$$

Now, let C be a directed cycle in G. Using the form for $f(i,j)$ that we just derived, let's look at the summation in the statement of KPL:

$$\sum_{(i,j)\in C} r(i,j)f(i,j) = \frac{1}{Z}\sum_{(i,j)\in C} r(i,j)\sum_{S\in\mathcal{S}}\left(\sum_{\substack{T:\delta^+(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)} - \sum_{\substack{T:\delta^-(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)}\right).$$

First, let's change the order of the first two summations.

$$\sum_{(i,j)\in C} r(i,j)f(i,j) = \frac{1}{Z}\sum_{S\in\mathcal{S}}\sum_{(i,j)\in C} r(i,j)\left(\sum_{\substack{T:\delta^+(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)} - \sum_{\substack{T:\delta^-(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T)}\right).$$

Now, recall that $r(T)$ is a product of resistances in the spanning tree, which includes $r(i,j)$. In fact, since we are also summing over trees that have a cut only connected by arc (i,j), note that $\frac{r(T)}{r(i,j)} = r(T_1)r(T_2)$ where T_1 is a sub-tree of T that spans S and T_2 is a sub-tree of T that spans $V - S$. Then, we can use this to simplify the expression, so that we have

$$\sum_{(i,j)\in C} r(i,j)f(i,j) = \frac{1}{Z}\sum_{S\in\mathcal{S}}\sum_{(i,j)\in C}\left(\sum_{\substack{T:\delta^+(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T_1)r(T_2)} - \sum_{\substack{T:\delta^-(S)\cap T\\=\{(i,j)\}}}\frac{1}{r(T_1)r(T_2)}\right).$$

Let's now observe that no term in the summation depends on the arc (i,j) other than the summation index for the spanning trees T. That means we are just counting and scaling the elements that satisfy the summation conditions, and so we can simplify further, as follows:

$$\sum_{(i,j)\in C} r(i,j)f(i,j) = \frac{1}{Z}\sum_{S\in\mathcal{S}} \left(|C\cap\delta^+(S)| - |C\cap\delta^+(S)|\right)\sum_{T_1,T_2}\frac{1}{r(T_1)r(T_2)}.$$

Finally, observe that because C is a cycle, the number of times it enters the set S must be equal to the number of times that it exits S, meaning that $|C\cap\delta^+(S)| = |C\cap\delta^+(S)|$. So,

$$\sum_{(i,j)\in C} r(i,j)f(i,j) = 0,$$

and thus we see that KPL is obeyed, and so f is an s-t electrical flow. □

This gives us a new perspective: an s-t electrical flow is the expected value of sampling tree T with probability proportional to $\frac{1}{r(T)}$ and sending 1 unit of flow on a unique s-t path in path in T. This gives an idea for a potential algorithm for computing s-t electrical flows: simply sample a bunch of spanning trees, send flows on their s-t paths, and take the average. While this is interesting, we can't make use of it directly as the field stands currently as we don't know a quick way to sample spanning trees. It does still give us interesting observations though. For example, note that the sampling probabilities do not depend on the choice of s and t. That's a neat observation, and it leads us towards a helpful lemma.

Lemma 2 *For any $\{i,j\}\in E$, let $T\in\mathcal{T}$ be sampled with probability proportional to $\frac{1}{r(T)}$. Then*

$$\Pr[\{i,j\}\in T] = \frac{r_{\mathit{eff}}(i,j)}{r(i,j)}\,.$$

Proof: Let f be an i-j electrical flow, and recall for p being the associated potentials, $r_{\text{eff}}(i,j) = p(i) - p(j)$. For any $T\in\mathcal{T}$, if $\{i,j\}\in\mathcal{T}$ then $f_T(i,j) = 1$ and if $\{i,j\}\notin\mathcal{T}$ then $f_T(i,j) = 0$. Then, using the definition of effective resistance, Ohm's Law, and the prior result, we have

$$\frac{r_{\text{eff}}(i,j)}{r(i,j)} = \frac{p(i)-p(j)}{r(i,j)} = f(i,j) = \sum_{T\in\mathcal{T}}\frac{1}{Z}\frac{1}{r(T)}f_T(i,j) = \sum_{\substack{T\in\mathcal{T}\\ \{i,j\}\in E}}\frac{1}{Z}\frac{1}{r(T)} = \Pr[\{i,j\}\in T]\,.$$

□

We use this lemma for the following theorem.

Theorem 3 (Foster's Theorem)

$$\sum_{\{i,j\}\in E}\frac{r_{\mathit{eff}}(i,j)}{r(i,j)} = n-1.$$

Proof: There are two approaches, one of which was originally presented in lecture and the other was suggested in class as a possible, simpler alternative.
Approach 1: We start by making use of Lemma 2:

$$\sum_{\{i,j\}\in E} \frac{r_{\text{eff}}(i,j)}{r(i,j)} = \sum_{\{i,j\}\in E} \Pr[\{i,j\} \in T].$$

Now we use total probability and condition on selecting each possible T. The conditional probability becomes deterministic and is thus an indicator, and the probability of sampling T is the form proportional to $\frac{1}{r(T)}$ that we saw earlier.

$$\sum_{\{i,j\}\in E} \frac{r_{\text{eff}}(i,j)}{r(i,j)} = \sum_{\{i,j\}\in E} \sum_{T\in\mathcal{T}} \frac{1}{Z} \frac{1}{r(T)} \mathbf{1}\left(\{i,j\} \in T\right).$$

Interchaning the order of summation we get

$$\sum_{\{i,j\}\in E} \frac{r_{\text{eff}}(i,j)}{r(i,j)} = \sum_{T\in\mathcal{T}} \frac{1}{Z} \frac{1}{r(T)} \sum_{\{i,j\}\in E} \mathbf{1}\left(\{i,j\} \in T\right).$$

Now, let's recognize that the innermost sum is counting the number of edges in a spanning tree. By definition, this is always $n-1$. Thus,

$$\sum_{\{i,j\}\in E} \frac{r_{\text{eff}}(i,j)}{r(i,j)} = \sum_{T\in\mathcal{T}} \frac{1}{Z} \frac{1}{r(T)} (n-1) = n-1.$$

This is the desired result.
Approach 2: Here we use the same first step as in the first approach.

$$\sum_{\{i,j\}\in E} \frac{r_{\text{eff}}(i,j)}{r(i,j)} = \sum_{\{i,j\}\in E} \Pr[\{i,j\} \in T].$$

Now, let's observe that this summation is equivalent to the calculation for the expected value for the number of edges in a randomly selected spanning tree. However, as we used above, this is will always be $n-1$ by definition, and so we are done.

$$\sum_{\{i,j\}\in E} \frac{r_{\text{eff}}(i,j)}{r(i,j)} = \mathrm{E}\left[|T|\right] = n-1.$$

□ While it is not provided here, this could also be shown via a third approach that follows from $\operatorname{tr}(L_G^{\dagger} L_G) = n-1$, which we argued was true in the previous lecture for G connected.

2 Random Walks II

We will now use some of the concepts we've gained from studying electrical flows to say new things about random walks. To start, let's defined a few new concepts. Let $s, t \in V$.

- Let the **hitting time** from s to t, $h(s,t)$, be the expected number of steps to go from s to t.
- Let the **commute time** of s and t, $C(s,t)$, be the expected number of steps to go from s to t and back.
- Let $C(i)$ be the expected number of steps to visit all the vertices when starting at i. Then the **cover time** is $C(G) = \max_i C(i)$.

Note that $C(s,t) = h(s,t) + h(t,s)$. As previously mentioned, we can use what we've learned from electrical flows to get new findings about random walks. We start with the first such result in a theorem.

Theorem 4

$$C(s,t) = 2m\, r_{\mathit{eff}}(s,t)$$

for graphs G with $r(i,j) = 1$ for every edge $\{i,j\} \in E$.

Proof: First let's note that $h(t,t) = 0$. Then, for any vertex $i \in V$ such that $i \neq t$,

$$h(i,t) = \sum_{j:\{i,j\}\in E} \frac{1}{d(i)}(1 + h(j,t)) = 1 + \sum_{j:\{i,j\}\in E} \frac{h(j,t)}{d(i)}$$

So,

$$d(i) = d(i)h(i,t) - \sum_{j:\{i,j\}\in E} h(j,t) = \sum_{j:\{i,j\}\in E} (h(i,t) - h(j,t))$$

At this point, we can note a familiar form: this looks quite like a Laplacian. With that in mind, let $p_t(i) = h(i,t)$ and

$$b_t(i) = \begin{cases} d(i) & \text{if } i \neq t \\ d(t) - 2m & \text{if } i = t \end{cases}$$

then $h(i,t)$ is a solution to $L_G p_t = b_t$ when $p_t(t) = 0$. Note that this choice of $b_t(t)$ enforces that $\sum_{i\in V} b_t(i) = 0$. Also note that if p is a solution to $L_G p = b_t$ then so is $p + c \cdot e$ for any $c \in \mathbb{R}$; thus, we are justified in setting $p_t(t) = 0$.

Similarly, if we let $p_s(i) = h(i,s)$ and

$$b_s(i) = \begin{cases} d(i) & \text{if } i \neq s \\ d(s) - 2m & \text{if } i = s \end{cases}$$

then $h(i,s)$ is a solution to $L_G p_s = b_s$ with $p_s(s) = 0$. Then, with both of these in hand,

$$L_G(p_t - p_s) = b_t - b_s = 2m(e_s - e_t)$$

which means that

$$\frac{1}{2m}(p_t - p_s) = L_G^{\dagger}(e_s - e_t)$$

and so $\frac{1}{2m}(p_t - p_s)$ are potentials for an s-t electrical flow. Then,

$$\begin{aligned} r_{\text{eff}}(s,t) &= (e_s - e_t)^T \left(\frac{1}{2m}(p_t - p_s) \right) = \frac{1}{2m}(p_t(s) - p_t(t) - p_s(s) + p_s(t)) \\ &= \frac{1}{2m}(p_t(s) + p_s(t)) = \frac{1}{2m}(h(s,t) + h(t,s)) = \frac{1}{2m}C(s,t) \end{aligned}$$

and so $C(s,t) = 2m\, r_{\text{eff}}(i,j)$. □

This result produces a nice corollary.

Corollary 5 *For any* $\{i,j\} \in E$,

$$C(i,j) \leq 2m$$

since $r_{\mathit{eff}}(i,j) \leq r(i,j) = 1$ *for* $\{i,j\} \in E$.

This theorem and the resulting corollary also lets us say something interesting about the cover time of the graph, and, by comparison, it does not as obviously stem from the electrical flows findings

Theorem 6

$$C(G) \leq 2m(n-1).$$

Proof: Pick any spanning tree T of G. Double all edges and orient each pair in opposite directions. Start at an arbitrary vertex and perform a Eulerian traversal to visit every arc and return to the start; the traversal must also visit every node. Then

$$\mathrm{E}\left[\text{total steps}\right] \leq \sum_{\{i,j\} \in T} (h(i,j) + h(j,i)) = \sum_{\{i,j\} \in T} C(i,j) \leq 2m(n-1).$$

□

However, this bound isn't tight. On the complete graph K_n, $2m(n-1) = \theta(n^3)$ but cover time is $O(n \log n)$. We can do better!

Theorem 7 *Let* $R(G) = \max_{i,j} r_{\mathit{eff}}(i,j)$. *Then,*

$$mR(G) \leq C(G) \leq 2e^3 mR(G) \log n + n$$

Proof: Let s,t be such that $R(G) = r_{\text{eff}}(s,t)$. Then,

$$2mr_{\text{eff}}(s,t) = 2mR(G) = C(s,t) = h(s,t) + (t,s)$$

So,

$$\begin{aligned} C(G) &\geq \max(C(s), C(t)) \\ &\geq \max(h(s,t), h(t,s)) \\ &\geq \frac{c(s,t)}{2} = mR(G) \end{aligned}$$

and so the left inequality is proved.

Now, on to the right inequality. For this, pick some vertex i. Break the walk into $\log n$ parts, each of length $2e^3mR(G)$. No matter where we start a part (say at k),

$$h(k,i) \leq 2mR(G)$$

By Markov's Inequality,

$$\Pr[\text{Number of steps to hit } i \ \geq 2e^3mR(G)] \leq \frac{2mR(G)}{2e^3mR(G)} = e^{-3},$$

which implies that

$$\Pr[\text{Don't hit } i \text{ in any of } \log n \text{ parts}] \leq \left(e^{-3}\right)^{\log n} = \frac{1}{n^3}.$$

This means the probability that there is some i that we don't hit is at most $n \cdot \frac{1}{n^3} = \frac{1}{n^2}$. In this case, we use tree walk of the previous thoerem to visit all vertices in expected time at most $2m(n-1) \leq n^3$ steps. Thus, the expected number of steps is less than $2e^3mP(G)\log n + n^2$. □