M40007: Introduction to Applied Mathematics

Imperial College London

1 A random journey on the "mini-Tube"

Consider a fictional mini London Tube network as shown in Figure 1.

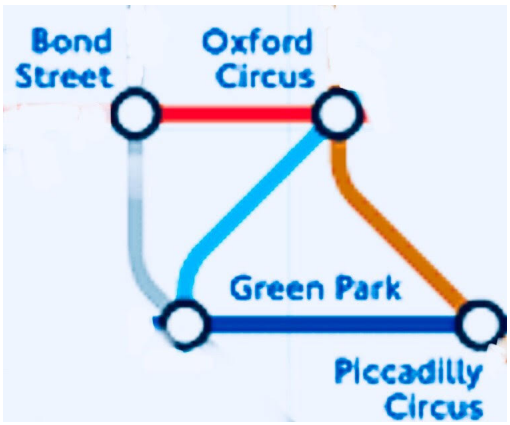

Figure 1: A map of the "mini-Tube" in a fictional London.

We can think of this as the graph with $n = 4$ nodes and $m = 5$ edges relabelled as shown in Figure 2:

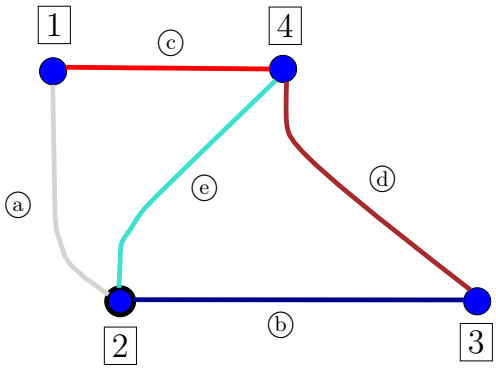

Figure 2: Relabelling of the "mini-Tube".

With this done, it is easy to see that this is precisely the same graph we have been considering earlier in the context of resistive electric circuits!

Suppose a tourist is on a mini-Tube journey and, when she arrives at any mini-Tube station, she decides at random which train to take next, with all possible routes from station station having equal probability, including the possibility of returning in the direction she came.

Question 1: What is the probability, if she starts her journey at Oxford Circus, of reaching Bond Street *before* she reaches Green Park? We will call this probability p_4 since Oxford Circus is node $\boxed{4}$ in Figure 2.

Question 2: What is this same probability if she starts, instead, at Piccadilly Circus? We will call this probability p_3 since Piccadilly Circus is node $\boxed{3}$ in Figure 2.

Question 3: What is the probability that, if she starts her journey at Bond Street and leaves this station, that she reaches Green Park before returning to Bond Street?

One way to estimate these probabilities is to carry out a direct numerical simulation. The following `MATLAB` code allows the tourist to start in `state` $\boxed{3}$ or $\boxed{4}$ – that is, at nodes $\boxed{3}$ or $\boxed{4}$ – and then simulates the random journey $N = 10,000$ times. It then counts the number of `hits` on Bond Street – or node $\boxed{1}$ – and computes the fraction of the total of N journeys that hit Bond Street first (the "hit rate").

```
%Random walk on mini-Tube: simulation of "hitting probability"
rng('shuffle');
N=10000;          %number of journeys
hit=0;           %initialize number of hits
for k=1:N;       %loop through N journeys
    flag=1;      %set a flag: if flag=1 keep going, if flag=0 STOP journey
    state=3;     %choose to be either node 3 or 4 (initial state)
    while flag == 1
        if state == 1
            flag=0;     %stop if state=1
            hit=hit+1;  %add a "hit"
        end;
        if state == 2
            flag=0;     %stop if state=2
        end;
        if state == 3
        toss=randi([1 2],1,1);      %toss a 2-sided coin
            if toss == 1;state=2;end;
            if toss == 2;state=4;end;
        end;
        if state == 4
        toss=randi([1 3],1,1);      %toss a 3-sided coin
            if toss == 1;state=3;end;
            if toss == 2;state=2;end;
            if toss == 3;state=1;end;
        end;
    end
end;
hit/N                               %output the hit rate
```

The results of three different runs of this code, with the `state` variable initiated at node $\boxed{4}$ – meaning that the tourist starts at Oxford Circus – produces the following output:

```
Command Window
>> RandomWalk2

ans =

    0.4002

>> RandomWalk2

ans =

    0.4056

>> RandomWalk2

ans =

    0.4042
```

The evidence points to the required probability being approximately

$$p_4 \approx 0.4, \qquad \text{or} \qquad \frac{2}{5}. \tag{1}$$

The code is run again, three times, but this time with the `state` variable initiated at node $\boxed{3}$. This means the tourist now starts at Piccadilly Circus each time.

```
Command Window
>> RandomWalk2

ans =

    0.2076

>> RandomWalk2

ans =

    0.1988

>> RandomWalk2

ans =

    0.1990
```

The evidence points to the required probability being approximately

$$p_3 \approx 0.2, \qquad \text{or} \qquad \frac{1}{5}. \tag{2}$$

We have estimates for the answer to Questions 1 and 2; we still have to answer Question 3.

The following MATLAB script simulates the situation where a tourist starts a journey leaving from Bond St – or node $\boxed{1}$ – and calculates the number of times, out of a possible $N = 5000$ journeys, she reaches Green Park – or node $\boxed{2}$ – before returning to Bond Street. This is done by using a `flag` to determine whether or not to terminate the simulation.

```
%Random walk on mini-Tube: simulation of "escape probability"
clear all;rng('shuffle');
N=5000;          %number of journeys
hit=0;           %initialize number of hits on node 2 ("escape node")
for k=1:N;       %loop through N journeys
    flag=1;      %set a flag: if flag=1 keep going, if flag=0 stop journey
    state=1;     %start at node 1 (initial state)
    toss=randi([1 2],1,1);      %toss a 2-sided coin
            if toss == 1;
                state=2;        %go to node 2
                hit=hit+1;      %record a "hit"
                flag=0;         %stop the journey
            end;
            if toss == 2;state=4;end;
    while flag == 1
        if state == 1
            flag=0;     %stop if in state=1
        end;
        if state == 2
            hit=hit+1;  %record a "hit"
            flag=0;     %stop if in state=2
        end;
        if state == 3
        toss=randi([1 2],1,1);      %toss a 2-sided coin
            if toss == 1;state=2;end;
            if toss == 2;state=4;end;
        end;
        if state == 4
        toss=randi([1 3],1,1);      %toss a 3-sided coin
            if toss == 1;state=3;end;
            if toss == 2;state=2;end;
            if toss == 3;state=1;end;
        end;
    end
end;
hit/N                           %output the "hit" rate
```

The results of three runs of the code just shown give:

```
Command Window
>> RandomWalk3

ans =

    0.8024

>> RandomWalk3

ans =

    0.7912

>> RandomWalk3

ans =

    0.8064
```

The probability that the tourist will reach Green Park before returning to Bond Street – or, in other words, the probability p_{esc} that the tourist *escapes* from Bond Street – is

$$p_{\text{esc}} \approx 0.8, \qquad \text{or} \qquad \frac{4}{5}. \tag{3}$$

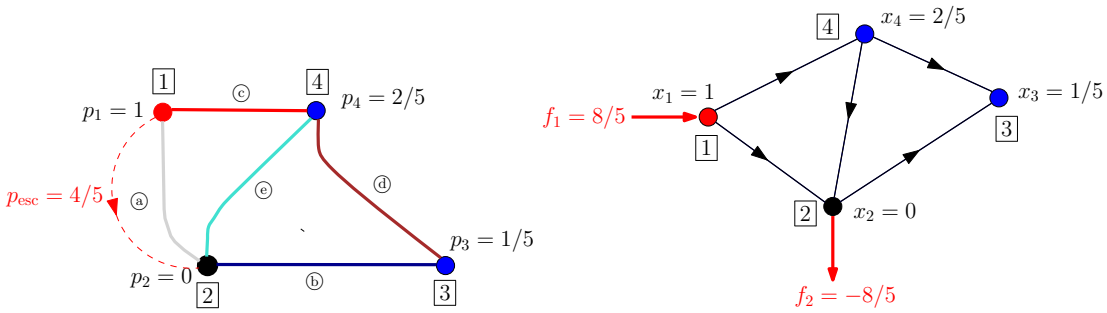

Figure 3: Summary of the results of the mini-Tube journey problem (left) and the resistive circuit (Dirichlet-type) problem (right).

Figure 3 is a juxtaposition of the results of the random mini-Tube journey just simulated and the results of the previous study of a resistive circuit with the same graph. The following observations can be made.

The probability p_4 appears to be $2/5$ which is the same as the value of the voltage x_4 at node $\boxed{4}$. The probability p_3 appears to be $1/5$ which is the same as the value of the voltage x_3 at node $\boxed{3}$. If we define p_1 to be the probability that, starting at Bond Street, i.e. node $\boxed{1}$, the tourist reaches Bond Street before Green Park, i.e. node $\boxed{2}$, then it is clear that

$$p_1 = 1 \tag{4}$$

since the tourist is already at Bond Street, so this outcome is certain. On the other hand, if p_2 to be the probability that, starting at Green Park, i.e. node $\boxed{2}$, the tourist reaches Bond Street before Green Park then it is clear that

$$p_2 = 0 \tag{5}$$

since the tourist is already at Green Park so there is no chance to reach Bond Street first.

It therefore appears to be the case that

$$p_i = x_i, \qquad i = 1, 2, 3, 4. \tag{6}$$

That is, the probability p_i that, starting at station i, the tourist reaches Bond Street (node $\boxed{1}$) before Green Park (node $\boxed{2}$) appears to equal the value of the node voltages of a circuit with node $\boxed{1}$ set to unit voltage and node $\boxed{2}$ grounded.

Another observation is that, while they are not the same, the effective conductance C_{eff} and p_{esc} appear to satisfy the simple relation

$$p_{\text{esc}} = \frac{C_{\text{eff}}}{2}. \tag{7}$$

2 Simple probability theory

Let us attempt an analytical solution to the problem of the random journey of the tourist on the mini-Tube shown in Figure 2. Let p_i be the probability that, starting at node $\boxed{i}$, the tourist making this random journey around the mini-Tube network reaches node $\boxed{1}$ (Bond Street) before node $\boxed{2}$ (Green Park). As discussed already,

$$p_1 = 1, \qquad p_2 = 0. \tag{8}$$

To find p_3 and p_4 we exploit the following law which is a basic one encountered in a first course in probability:

Basic probability law: Let F and G be events such that only *one* of the the two can occur. Let E be some other event. Then

$$P(E) = P(F) \times P(E \text{ given } F) + P(G) \times P(E \text{ given } G). \tag{9}$$

Here $P(E)$ means *the probability of event E occurring*.

Using this basic law of probability we can immediately infer that

$$p_3 = \frac{1}{2}p_2 + \frac{1}{2}p_4. \tag{10}$$

Here E is the event "reaching node $\boxed{1}$ before node $\boxed{2}$"; F is the event "moving to node $\boxed{2}$ from node $\boxed{3}$" and G is the event "moving to node $\boxed{4}$ from node $\boxed{3}$". Clearly, if the tourist is at node $\boxed{3}$ then only event F or G can occur. Relation (10) holds because, when the tourist is at node $\boxed{3}$, she choses which train line to take at random so she travels to nodes $\boxed{2}$ or $\boxed{4}$ with equal probability $1/2$.

Using similar arguments, it can be deduced that

$$p_4 = \frac{1}{3}p_1 + \frac{1}{3}p_2 + \frac{1}{3}p_3, \tag{11}$$

where the factors $1/3$ appear because when the tourist is at node $\boxed{4}$ she travels to nodes $\boxed{1}$, $\boxed{2}$ and $\boxed{3}$ with equal probability.

Now, if (8) is used, the two equations (10) and (11), become

$$\begin{aligned} 2p_3 - p_4 &= 0, \\ -p_3 + 3p_4 &= 1. \end{aligned} \tag{12}$$

These are easily solved to find

$$p_3 = \frac{1}{5}, \qquad p_4 = \frac{2}{5}. \tag{13}$$

This is are the exact results, and they are consistent with the approximations given by the numerical simulations performed earlier.

3 Electric circuits and random walks

Let us explore in more detail the intriguing connections, evident from Figure 3, between the electric circuit problem and the random journey on the mini-Tube. The connection between the vector of voltages $\mathbf{x}$ and the vector $\mathbf{f}$ containing the net current out of each node is

$$\mathbf{f} = \mathbf{A}^T\mathbf{A}\mathbf{x}. \tag{14}$$

At any node $\boxed{i}$ at which Kirchhoff's current law holds we look at the *row* of $\mathbf{A}^T\mathbf{A}$ corresponding to node $\boxed{i}$ and set that component equal to zero. Then the net current out of node $\boxed{i}$ is zero: this is the KCL condition.

We also know that the Laplacian $\mathbf{K}$ has the decomposition

$$\mathbf{K} = \mathbf{A}^T\mathbf{A} = \mathbf{D} - \mathbf{W}, \tag{15}$$

where $\mathbf{D}$ is the degree matrix and $\mathbf{W}$ is the adjacency matrix. This means that the *row* of $\mathbf{A}^T\mathbf{A}$ corresponding to node $\boxed{i}$ is

$$\text{degree(node } i) \times x_i - \sum_{\text{j}\in\text{adjacent nodes}} x_j = 0, \tag{16}$$

where degree(node i) is the number of edges connected to node $\boxed{i}$; the sum is over all the nodes connected (or "adjacent") to node $\boxed{i}$. This equation can be written

$$x_i = \frac{1}{\text{degree(node } i)} \sum_{\text{j}\in\text{adjacent nodes}} x_j. \tag{17}$$

But, if you write down the two equations coming from (17) for nodes $\boxed{3}$ and $\boxed{4}$ in the electric circuit problem in Figure 3, they are found to be exactly the same as the equations satisfied by the probabilities in the problem of a random journey on the mini-Tube.

This is why we find that the voltages for the circuit problem in Figure 3 coincide with the probabilities for the random journey on the mini-Tube. This connection will now be explored in more detail.

Harmonic potentials: Another interpretation of (17) is that the potential x_i at node $\boxed{i}$ is the *average* of the potentials at all the other adjacent nodes to which it is connected by a conductor, or edge. We call any potential having this property a *harmonic potential*.

4 Method of relaxation

Before exploring the connection between the electric circuit problem and the random journey problem in more detail, we pause to explore implications of the observation that, in the electric circuit problem, the potentials at nodes where KCL holds are given by the average of the potentials at all adjacent nodes. This suggests a novel numerical algorithm for the computation of the probabilities, or voltages, called the *method of relaxation*. It is an *iterative method* whereby we initially assign the initial potentials as follows:

$$x_1 = 1, \qquad x_2 = 0, \qquad x_3 = 0, \qquad x_4 = 0 \tag{18}$$

This initial choice, represented by the first diagram in Figure 4, does not satisfy (17) but it is a simple choice, and it satisfies the conditions $x_1 = 1, x_2 = 0$ at the chosen boundary nodes. The idea is to cycle around the nodes and update their voltages so that they satisfy (17).

Consider first node $\boxed{4}$. Currently its potential is $x_4 = 0$ but we now *update it with the average of the potentials at adjacent nodes*. In this case, we change

$$x_4 = 0 \mapsto \frac{1}{3}(1 + 0 + 0) = \frac{1}{3}. \tag{19}$$

The potential at $x_3 = 0$ remains unchanged. This updated set of potentials is represented by the second diagram in Figure 4,

Next we turn to node $\boxed{3}$. We update its potential with the average of the potentials at its adjacent nodes:

$$x_3 = 0 \mapsto \frac{1}{2}\left(\frac{1}{3} + 0\right) = \frac{1}{6}. \tag{20}$$

The potential at $x_4 = \frac{1}{3}$ remains unchanged. This updated set of potentials is represented by the third diagram in Figure 4.

We have updated the two unknown potentials x_3 and x_4. Now we repeat – or "iterate" – this procedure in the hope that it might converge to the required solution.

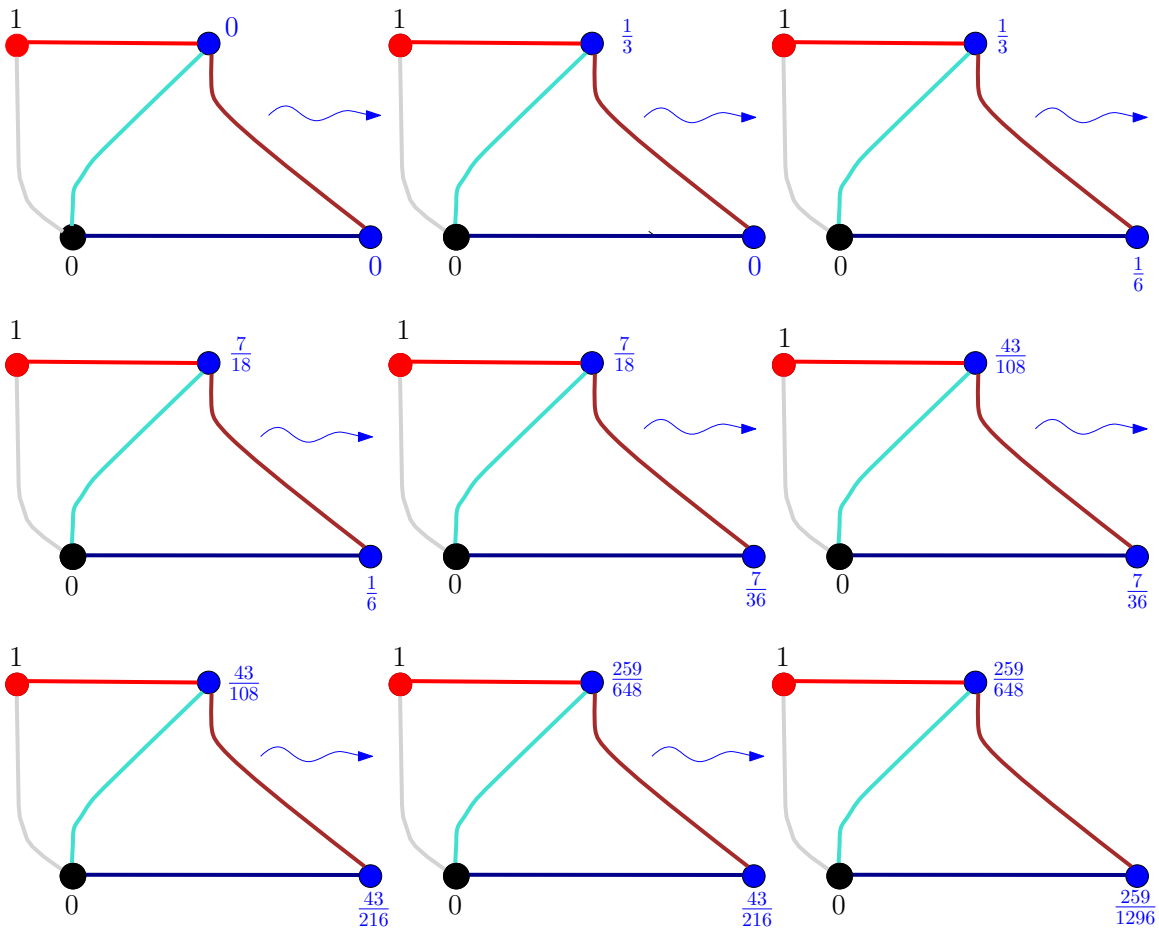

Figure 4: Sequence of updated node potentials in the method of relaxation.

In the next iteration the potential x_4 of $\boxed{4}$ is again updated with the average of the potentials at adjacent nodes:

$$x_4 = 0 \mapsto \frac{1}{3}\left(1 + 0 + \frac{1}{6}\right) = \frac{7}{18}. \tag{21}$$

The potential at $x_3 = \frac{1}{6}$ remains unchanged. This updated set of potentials is represented by the fourth diagram in Figure 4,

Next the potential x_3 of $\boxed{3}$ is updated with the average of the potentials at adjacent nodes:

$$x_3 = 0 \mapsto \frac{1}{2}\left(0 + \frac{7}{18}\right) = \frac{7}{36}. \tag{22}$$

The potential at $x_4 = \frac{7}{18}$ remains unchanged. This updated set of potentials is represented by the fifth diagram in Figure 4.

Figure 4 shows the results of two more cycles through this procedure leading to

$$x_3 = \frac{259}{1296} \approx 0.39969, \qquad x_4 = \frac{259}{638} \approx 0.19984. \tag{23}$$

These are already very close to the known values $2/5 = 0.4$ and $1/5 = 0.2$. They are sufficiently close that it is reasonable to stop the iteration.

This example demonstrates that the "method of relaxation" appears to converge to the solution in this case. Of course, it remains to prove that the method works in general.

Interestingly, if the two problems are equivalent, we now have three different ways to arrive at the solution to the electric circuit problem, or conversely, the problem of the random journey on the mini-Tube:

(1) We can use linear algebra methods based on Schur complements to find the unknown potentials at the nodes.

(2) We can perform a simulated random walk multiple times and take the average of the hit rates.

(3) We can use the iterative method of relaxation just described to compute the potentials numerically.

In applied mathematics, it is always valuable to have multiple ways to solve a problem, not least because some methods can be more computationally efficient than others. Moreover, noticing theoretical analogies of this kind allows us to use theoretical techniques and ideas arising from one problem area to study solutions of very different problems where those ideas might not be so obviously relevant.

Since harmonic potentials appear to underlie both the eletric circuit and random journey problems, we now explore them in more detail.

5 Harmonic potentials

Any set of potentials $\{x_i\}$ satisfying the conditions

$$x_i = \frac{1}{\text{degree}(\text{node } i)} \sum_{\text{j}\in\text{adjacent nodes}} x_j \tag{24}$$

where the potential x_i at node $\boxed{\text{i}}$ is the *average* of the potentials at all the other adjacent nodes to which it is connected by a conductor, or edge, is called a *harmonic potential*.

Given a connected graph it is natural to separate nodes into *boundary nodes* denoted and *interior nodes*. At the boundary nodes, the potential is specified. At the interior nodes, KCL holds and the potential is harmonic there.

The maximum principle for harmonic potentials: The *maximum principle* states the the maximum possible value of the potential in this connected graph occurs at a boundary node.

To prove this, let M be the maximum node potential in the graph. If M is attained at a boundary node then the result is already established. Therefore we suppose the potential at some interior node, x_i, is equal to M, i.e., $x_i = M$. The same must then be true of all its adjacent nodes because the value at x_i is the average of the adjacent node potentials so none can be lower than M (and none can be higher than M, by definition). If all the nodes adjacent to x_i are in the interior of D then, by the same argument, all *their* neighbours must have potential M also. Continuing in this way we must *eventually* find a neighbour that is a boundary node because the graph is connected thereby implying the result: the maximum node potential M is attained on the boundary.

The minimum principle for harmonic potentials: The *minimum principle* states the the minimum possible value of the potential in this connected graph occurs at the boundary nodes. The proof is a straightforward adaptation of that for the maximum principle.

The uniqueness principle for harmonic potentials: How do we know for sure that the solution of the electric circuit problem and the solution of the random journey on the mini-Tube are the same? The proof is an application of the *uniquenss principle for harmonic potentials*.

Let x_i and p_i be harmonic potentials having the same boundary values at the boundary nodes of a connected graph. Define the new potential

$$d_i = x_i - p_i. \tag{25}$$

By assumption this potential, which being the difference of two harmonic potentials is also harmonic, is zero at the boundary points. But by the maximum and minimum principle the set of node potentials $\{d_i\}$ attains both its maximum and minimum values at the boundary nodes. But these are both zero meaning that d_i is zero on all the nodes of the graph. Therefore

$$x_i = p_i \tag{26}$$

at all nodes of the graph.

6 Random walks and electric circuits

We have seen how an electric circuit in which all conductors have equal unit conductance can be associated with a harmonic potential. It has also been seen how the mathematical problem for a 2-point source-sink circuit problem, with a + node and a - node, is analogous to the mathematical problem associated with a random walk

on a graph and the probability of whether a random walker reaches the + node of the graph before the - node.

This is a special case of a more general mathematical analogy between the two problems which we now explain. Two definitions are needed for this.

Let p_i be the *hitting probability* of the + node for a random walk starting at node $\boxed{i}$.

The *hopping probability* p_{ij} is the probability of hopping from node $\boxed{i}$ to an adjacent node $\boxed{j}$. Figure (5) shows the hopping probabilities for the random walk in the mini-Tube. Note that

$$p_{43} = \frac{1}{3}, \qquad p_{34} = \frac{1}{2}, \tag{27}$$

so the hopping probability between two nodes can be different in the two directions.

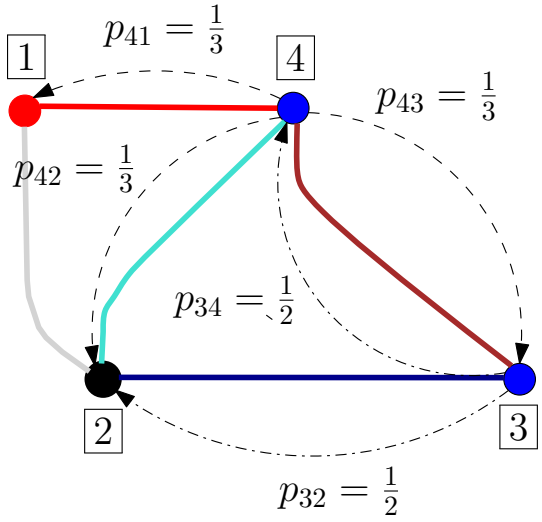

Figure 5: Hopping probabilities for the random walk in the mini-Tube.

A mathematical analogy: Let two boundary nodes, a $+$ node and a $-$ node, in an electric circuit be fixed at voltage 1 and 0, respectively. Let c_{ij} be the conductance of the conductor between node $\boxed{i}$ and node $\boxed{j}$. The voltage x_i at any interior node $\boxed{i}$ is the same as the hitting probability p_i for a random walker, starting at node $\boxed{i}$ to reach the $+$ node before the $-$ node when the hopping probability p_{ij} from node $\boxed{i}$ to node $\boxed{j}$ is

$$p_{ij} = \frac{c_{ij}}{c_i}, \qquad c_i = \sum_{j \in \text{adjacent}} c_{ij}, \tag{28}$$

where the sum denotes a sum over all nodes adjacent to node $\boxed{i}$.

Escape probability and effective conductance: Given this analogy it is natural to ask if there is a probabilistic interpretation of the effective conductance of the

circuit. By Ohm's Law, and the definition of effective conductance,

$$C_{\text{eff}} = -\sum_{j\in\text{adjacent}} c_{+j}(x_j - x_+) = \sum_{j\in\text{adjacent}} c_{+j}(x_+ - x_j). \tag{29}$$

Here we use the notation c_{+j} to denote the conductance of the edge joining to $+$ node to node j and x_+ is the potential at the $+$ node. As usual C_{eff} is the current entering the circuit at the + node and leaving it at the - node. This can be rewritten as

$$C_{\text{eff}} = x_+ c_+ - \sum_{j\in\text{adjacent}} c_{+j}x_j, \tag{30}$$

where

$$c_+ = \sum_{j\in\text{adjacent}} c_{+j}. \tag{31}$$

Equation (30) can also be written

$$C_{\text{eff}} = x_+ c_+ - \sum_{j\in\text{adjacent}} c_{+j}x_j\left(\frac{c_+}{c_+}\right) = c_+\left(x_+ - \sum_{j\in\text{adjacent}} \frac{c_{+j}}{c_+}x_j\right). \tag{32}$$

But

$$\frac{c_{+j}}{c_+} = p_{+j}, \tag{33}$$

where p_{+j} is the hopping probability from the $+$ node to any adjacent nodes hence

$$C_{\text{eff}} = c_+\left(x_+ - \sum_{j\in\text{adjacent}} p_{+j}x_j\right). \tag{34}$$

Now $x_+ = 1$ and

$$\sum_{j\in\text{adjacent}} p_{+j}x_j \tag{35}$$

is precisely the probability of starting at the $+$ node, leaving on a random walk, and then returning to the $+$ node before hitting the $-$ node. That is, it is the probability of returning to the $+$ node after leaving it. The quantity

$$1 - \sum_{j\in\text{adjacent}} p_{+j}x_j \tag{36}$$

is therefore the probability of *not returning*, that is, *escaping*, from the $+$ node after leaving it. Therefore (30) says

$$C_{\text{eff}} = c_+ p_{\text{esc}} \tag{37}$$

or

$$p_{\text{esc}} = \frac{C_{\text{eff}}}{c_+}. \tag{38}$$

From its definition (31), c_+ is the sum of the conductances of all the edges adjacent to the $+$ node.

For the circuit example,

$$C_{\text{eff}} = \frac{8}{5} \tag{39}$$

and since node $\boxed{1}$ has two conductors of unit conductance attached to it,

$$c_+ = 2. \tag{40}$$

Formula (38) therefore implies that

$$p_{\text{esc}} = \frac{C_{\text{eff}}}{c_+} = \frac{4}{5} \tag{41}$$

which is consistent with the result of the earlier numerical simulation.

7 Dirichlet's principle

The energy dissipation associated with current in a two-point source-sink circuit with a $+$ node at unit voltage and a $-$ node grounded at zero voltage is defined as the quantity

$$\mathcal{E}(\mathbf{x}) = \mathbf{x}^T \mathbf{K} \mathbf{x}, \tag{42}$$

where $\mathbf{K}$ is the Laplacian matrix of the graph representing the circuit and $\mathbf{x}$ is *any* potential defined at the nodes.

Dirichlet's principle says that the particular node potential $\mathbf{x}_*$ with boundary potentials $x_+ = 1$ and $x_- = 0$ and with KCL holding at all the other interior nodes minimizes the energy dissipation:

$$\mathcal{E}(\mathbf{x}_*) = \min_{\mathbf{x}} \mathcal{E}(\mathbf{x}). \tag{43}$$

In other words, if one minimizes the energy dissipation viewed as a function of the voltages at the nodes then it is the voltages satisfying KCL at all the interior nodes that minimizes the energy dissipation.

To prove Dirichlet's principle, consider first the fundamental relation between voltage vector $\mathbf{x}$ and the vector $\mathbf{f}$ of the net currents out of each node:

$$\mathbf{K}\mathbf{x} = \mathbf{f}. \tag{44}$$

Assume, without loss of generality, that the ordering is chosen to that the first element of $\mathbf{x}$ corresponds to the $+$ node and the second element to the $-$ node.

Hence, *assuming KCL holds at the interior nodes*, we know that the system to solve is

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \\ \hat{\mathbf{x}} \end{pmatrix}, \qquad \mathbf{f} = \begin{pmatrix} f \\ -f \\ \mathbf{0} \end{pmatrix}, \tag{45}$$

where f is the net current into the $+$ node, i.e., the effective conductance. Now introduce the sub-block decomposition

$$\mathbf{K} = \begin{pmatrix} \mathbf{P} & \mathbf{Q}^T \\ \mathbf{Q} & \mathbf{R} \end{pmatrix}, \tag{46}$$

where $\mathbf{P}$ is a 2-by-2 sub-block and $\mathbf{R}$ is a $(n-2)$-by-$(n-2)$ sub-block. The linear system (44) can be written

$$\begin{pmatrix} \mathbf{P} & \mathbf{Q}^T \\ \mathbf{Q} & \mathbf{R} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \hat{\mathbf{x}} \end{pmatrix} = \begin{pmatrix} f \\ -f \\ \mathbf{0} \end{pmatrix}. \tag{47}$$

The equation for $\hat{\mathbf{x}}$ is

$$\mathbf{Q}\mathbf{e} + \mathbf{R}\hat{\mathbf{x}} = 0, \qquad \mathbf{e} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{48}$$

From a previous exercise, we know that $\mathbf{R}$ is positive definite and therefore invertible. Therefore,

$$\hat{\mathbf{x}} = -\mathbf{R}^{-1}\mathbf{Q}\mathbf{e}. \tag{49}$$

This is the expression for the voltages of the interior nodes *when KCL is imposed there*.

Let us turn now to the energy dissipation. It can be written using the same sub-block decomposition as

$$\begin{aligned} \mathcal{E}(\mathbf{x}) = \mathbf{x}^T\mathbf{K}\mathbf{x} &= (\mathbf{e}^T, \hat{\mathbf{x}}^T) \begin{pmatrix} \mathbf{P} & \mathbf{Q}^T \\ \mathbf{Q} & \mathbf{R} \end{pmatrix} \begin{pmatrix} \mathbf{e} \\ \hat{\mathbf{x}} \end{pmatrix} \\ &= \mathbf{e}^T\mathbf{P}\mathbf{e} + \mathbf{e}^T\mathbf{Q}^T\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{Q}\mathbf{e} + \hat{\mathbf{x}}^T\mathbf{R}\hat{\mathbf{x}}. \end{aligned} \tag{50}$$

Taking a transpose of a scalar quantity leaves it invariant therefore

$$\mathbf{e}^T\mathbf{Q}^T\hat{\mathbf{x}} = (\mathbf{e}^T\mathbf{Q}^T\hat{\mathbf{x}})^T = \hat{\mathbf{x}}^T\mathbf{Q}\mathbf{e}. \tag{51}$$

This means that (50) can be written as

$$\mathcal{E}(\mathbf{x}) = \mathbf{e}^T\mathbf{P}\mathbf{e} + \underbrace{2\hat{\mathbf{x}}^T\mathbf{Q}\mathbf{e} + \hat{\mathbf{x}}^T\mathbf{R}\hat{\mathbf{x}}}_{A}, \tag{52}$$

where we have isolated a term denoted by A for special treatment. It is useful now

to consider the quantity

$$(\hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}(\hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e}) = \hat{\mathbf{x}}^T\mathbf{R}\hat{\mathbf{x}} + (\mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}\hat{\mathbf{x}} + \hat{\mathbf{x}}^T\mathbf{R}(\mathbf{R}^{-1}\mathbf{Q}\mathbf{e}) + (\mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}(\mathbf{R}^{-1}\mathbf{Q}\mathbf{e}). \tag{53}$$

Using again the result on the transpose of a scalar quantity then

$$(\mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}\hat{\mathbf{x}} = ((\mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}\hat{\mathbf{x}})^T = \hat{\mathbf{x}}^T\mathbf{R}^T(\mathbf{R}^{-1}\mathbf{Q}\mathbf{e}) = \hat{\mathbf{x}}^T\mathbf{Q}\mathbf{e}, \tag{54}$$

where, in the last equality, we have used the fact the sub-block $\mathbf{R}$ is a symmetric matrix: $\mathbf{R}^T = \mathbf{R}$; it inherits this property from $\mathbf{K}$ of which it is a sub-block. Hence (53) is

$$(\hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}(\hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e}) = \underbrace{\hat{\mathbf{x}}^T\mathbf{R}\hat{\mathbf{x}} + 2\hat{\mathbf{x}}^T\mathbf{Q}\mathbf{e}}_{A} + (\mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{Q}\mathbf{e}, \tag{55}$$

where we spot the same term A appearing in (52). On substituting for A from (55) in (52) we find a new expression for the energy dissipation:

$$\mathcal{E}(\mathbf{x}) = \underbrace{\mathbf{e}^T\mathbf{P}\mathbf{e} - (\mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{Q}\mathbf{e}}_{B} + (\hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e})^T\mathbf{R}(\hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e}). \tag{56}$$

We have isolated another term, called B, in this expression. If we now introduce

$$\mathbf{X} = \hat{\mathbf{x}} + \mathbf{R}^{-1}\mathbf{Q}\mathbf{e} \tag{57}$$

then the energy dissipation (56) can be written in the revealing form

$$\mathcal{E}(\mathbf{x}) = \mathbf{X}^T\mathbf{R}\mathbf{X} + B. \tag{58}$$

The term B is independent of $\hat{\mathbf{x}}$ so it does not play a role in the minimization, while $\mathbf{R}$ is known to be positive definite. It follows that $\mathcal{E}(\mathbf{x})$ is minimized when

$$\mathbf{X} = 0, \qquad \Longleftrightarrow \qquad \hat{\mathbf{x}} = -\mathbf{R}^{-1}\mathbf{Q}\mathbf{e}. \tag{59}$$

This is precisely the same $\hat{\mathbf{x}}$ as given in (49) which we showed is relevant when KCL holds at all the interior nodes. We have therefore proved Dirichlet's principle.

8 Thomson's principle

The energy dissipation in a circuit can also be written as

$$\mathcal{E}(\mathbf{x}) = \mathbf{x}^T\mathbf{K}\mathbf{x} = \mathbf{x}^T\mathbf{A}^T\mathbf{C}\mathbf{A}\mathbf{x} = \underbrace{(-\mathbf{A}\mathbf{x})}_{-\mathbf{e}^T}^T\underbrace{(-\mathbf{C}\mathbf{A}\mathbf{x})}_{\mathbf{w}}, \tag{60}$$

where $\mathbf{e}$ is the vector of potential differences across the edges and $\mathbf{w}$ is the vector of edge currents. Hence,

$$\mathcal{E}(\mathbf{x}) = -\mathbf{e}^T\mathbf{w}. \tag{61}$$

From Ohm's law, if w_k is the current in edge k and e_k is the potential difference across it then

$$w_k = -c_k e_k, \tag{62}$$

where c_k denotes the conductance of the conductor that forms edge k. The expression (61) can therefore be written as

$$\mathcal{E}(\mathbf{x}) = \sum_{\text{edge k}} \left(\frac{w_k}{c_k}\right) w_k := \tilde{\mathcal{E}}(\mathbf{w}). \tag{63}$$

This sum is now over all the edges of the graph. In (63) we have introduced the new notation $\tilde{\mathcal{E}}(\mathbf{w})$ because we are now thinking of the dissipation *not* as a function of $\mathbf{x}$, the voltages at the nodes, but as a function of $\mathbf{w}$, the currents in the edges. This is how *Thomson's principle* differs from Dirichlet's principle.

We can now state *Thomson's principle*. It says that the current $\mathbf{w}_*$ derived from a voltage potential that satisfies KCL at all interior nodes minimizes the energy dissipation $\tilde{\mathcal{E}}(\mathbf{w})$ over all possible currents $\mathbf{w}$ defined in the circuit – including ones not necessarily derivable from a voltage potential. That is,

$$\tilde{\mathcal{E}}(\mathbf{w}_*) = \min_{\mathbf{w}} \tilde{\mathcal{E}}(\mathbf{w}). \tag{64}$$

Thomson's principle and Dirichlet's principle can be viewed as *dual* statements of the same idea: that the current in an electric circuit minimizes the energy dissipation. Thomson thought about the energy dissipation as a function of the edge currents; Dirichlet thought about it as a function of the node voltages.

The proof of Thomson's principle is left as an exercise.

9 Tellegen's theorem

Given a connected graph with m-by-n incidence matrix $\mathbf{A}$ then if $\mathbf{e} \in \mathbb{R}^m$ is *any* vector of potential differences and $\mathbf{w}' \in \mathbb{R}^m$ is *any* vector of edge variables satisfying KCL – or, more generally, the zero divergence condition – at all the nodes, then

$$\mathbf{e}^T\mathbf{w}' = 0. \tag{65}$$

This is *Tellegen's theorem*.

We have used $\mathbf{w}'$ instead of just $\mathbf{w}$ for good reason: we are used to seeing the connection, $\mathbf{w} = -\mathbf{C}\mathbf{e}$, which is just Ohm's law relating the currents $\mathbf{w}$ in the edges to the potential differences $\mathbf{e}$ across them. However, Tellegen's theorem is true *without* this additional relationship. Indeed, it holds even when the vector of

potential differences $\mathbf{e}$ and the current vector $\mathbf{w}'$ have no connection to each other.

The proof of Tellegen's theorem is easy. Since $\mathbf{e}$ is a vector of potential differences then

$$\mathbf{e} = \mathbf{A}\mathbf{x} \tag{66}$$

for some vector $\mathbf{x}$ of potentials. It follows that

$$\mathbf{e}^T\mathbf{w}' = (\mathbf{A}\mathbf{x})^T\mathbf{w}' = \mathbf{x}^T\mathbf{A}^T\mathbf{w}'. \tag{67}$$

But

$$\mathbf{A}^T\mathbf{w}' = 0 \tag{68}$$

since $\mathbf{w}'$ satisfies KCL at all the nodes, so its divergence at the nodes is zero. The result (65) then follows.

Modified form of Tellegen's theorem: Suppose, on some connected graph with incidence matrix $\mathbf{A}$, a vector $\mathbf{w}' \in \mathbb{R}^m$ satisfies

$$-\mathbf{A}^T\mathbf{w}' = \mathbf{f}' \tag{69}$$

for some $\mathbf{f}' \in \mathbb{R}^n$. Then if $\mathbf{e}$ is any vector of potential differences, i.e., $\mathbf{e} = \mathbf{A}\mathbf{x}$ for some vector of potentials $\mathbf{x} \in \mathbb{R}^n$, then

$$\mathbf{e}^T\mathbf{w}' = -\mathbf{x}^T\mathbf{f}'. \tag{70}$$

Relation (70) is interesting: the left hand side is a dot product of two vectors in $\mathbb{R}^m$ while the right hand side is a dot product of two vectors in $\mathbb{R}^n$. This form of Tellegen's theorem says that these two dot products in the different vector spaces is the same. The special case $\mathbf{f}' = 0$ corresponds to the earlier statement (65).

Again, the proof of (70) is easy:

$$\mathbf{e}^T\mathbf{w}' = (\mathbf{A}\mathbf{x})^T\mathbf{w}' = \mathbf{x}^T\mathbf{A}^T\mathbf{w}' = -\mathbf{x}^T\mathbf{f}', \tag{71}$$

which is precisely (70).

Effective conductance and energy dissipation: Consider a 2-point source-sink electrical circuit with a $+$ node set to unit voltage and a $-$ node grounded. Let $\mathbf{w}$ be the vector of currents in the circuit and let $\mathbf{e}$ be the vector of potential differences, $\mathbf{e} = \mathbf{A}\mathbf{x}$. Then, by Tellegen's theorem,

$$\mathbf{e}^T\mathbf{w} = -\mathbf{x}^T\mathbf{f}. \tag{72}$$

But this implies

$$\mathcal{E}(\mathbf{x}) = -\mathbf{e}^T\mathbf{w} = \mathbf{x}^T\mathbf{f} = x_+C_{\text{eff}} - x_-C_{\text{eff}} \tag{73}$$

where we have used the fact that the components of $\mathbf{f}$ are zero at all interior nodes, since KCL holds there, while the components corresponding to the $+$ node and $-$

node are C_{eff} and $-C_{\text{eff}}$ respectively. Since $x_+ = 1$ and $x_- = 0$ we conclude that

$$\mathcal{E}(\mathbf{x}_*) = C_{\text{eff}}. \tag{74}$$

Remarkably, the energy dissipation at its minimum value – which we know by Dirichlet's principle occurs when KCL is satisfied at all interior nodes – is given by the effective conductance of the 2-point source/sink circuit!