

THE MODULUS AND EPIDEMIC PROCESSES ON GRAPHS

by

MAX GOERING

B.S., Kansas State University, 2015

---

A THESIS

submitted in partial fulfillment of the  
requirements for the degree

MASTER OF SCIENCE

Department of Mathematics  
College of Arts and Sciences

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

2015

Approved by:

Major Professor  
Pietro Poggi-Corradini

# Copyright

MAX GOERING

2015

# Abstract

This thesis contains three chapters split into two parts. In the first chapter, the discrete  $p$ -modulus of families of walks is introduced and discussed from various perspectives. Initially, we prove many properties by mimicking the theory from the continuous case and use Arne Beurling's criterion for extremality to build insight and intuition regarding the modulus. After building an intuitive understanding of the  $p$ -modulus, we proceed to switch perspectives to that of convex analysis. From here, uniqueness and existence of extremal densities is shown and a better understanding of Beurling's criterion is developed before describing an algorithm that approximates the value of the  $p$ -modulus arbitrarily well.

In the second chapter, an exclusively edge-based approach to the discrete transboundary modulus is described. Then an interesting application is discussed with some preliminary numerical results.

The final chapter describes four different takes of the Susceptible-Infected (SI) epidemic model on graphs and shows them to be equivalent. After developing a deep understanding of the SI model, the epidemic hitting time is compared to a variety of different graph centralities to indicate successful alternative methods in identifying important agents in epidemic spreading. Numerical results from simulations on many real-world graphs are presented. They indicate the effective resistance, which coincides with the 2-modulus for connecting families, is the most closely correlated indicator of importance to that of the epidemic hitting time. In large part, this is suspected to be due to the global nature of both the effective resistance and the epidemic hitting time. Thanks to the equivalence between the epidemic hitting time and the expected distance on an randomly exponentially weighted graph, we uncover a deeper connection- the effective resistance is also a lower bound for the epidemic hitting time, showing an even deeper connection.

# Table of Contents

List of Figures	vi
List of Tables	viii
Acknowledgements	viii
Dedication	x
<b>1 The Modulus</b>	<b>1</b>
1.1 Walks on Finite Graphs . . . . .	1
1.1.1 Preliminaries . . . . .	1
1.1.2 Beurling’s Criterion and Extremal Densities . . . . .	6
1.1.3 Properties of the $p$ -Modulus . . . . .	10
1.1.4 Capacity . . . . .	16
1.2 Discrete Modulus via Convex Optimization. . . . .	19
1.2.1 Existence of Essential Subfamilies . . . . .	20
1.2.2 Karush-Kuhn-Tucker conditions and the Modulus as a Convex Program . . . . .	24
1.2.3 Computing the Modulus . . . . .	29
<b>2 Transboundary Modulus</b>	<b>36</b>
2.1 More graph theory . . . . .	36
2.2 Transboundary modulus . . . . .	38
2.2.1 Preliminaries . . . . .	38

2.2.2	Virtual Graphs . . . . .	39
2.3	Quadrangulations of the sphere . . . . .	44
2.3.1	Sheffield's Bijection . . . . .	44
2.3.2	Maps and words with one loop . . . . .	47
2.3.3	Maps and words with more than one loop . . . . .	50
<b>3</b>	<b>Graph centralities and epidemic measures</b>	<b>53</b>
3.1	Dijkstra's Algorithm . . . . .	54
3.2	The four fold path to the SI model . . . . .	56
3.2.1	The SI model . . . . .	56
3.2.2	Node-event based model . . . . .	63
3.2.3	Edge-event based model . . . . .	64
3.2.4	Variable-lengths model . . . . .	64
3.2.5	Dijkstra on the Fly . . . . .	67
3.3	Metrics and Centralities . . . . .	68
3.3.1	Metrics . . . . .	68
3.3.2	Centralities . . . . .	69
3.4	Real-World Networks . . . . .	71
3.5	Numerical Methods and Results . . . . .	73
3.5.1	Properties of the Epidemic Hitting Time . . . . .	74
	<b>Bibliography</b>	<b>79</b>
	<b>A Probability</b>	<b>83</b>
A.1	The Exponential Distribution . . . . .	83

# List of Figures

1.1	Karate Club Network: A network of friendships between the 34 members of a karate club at a US university, just before the club split. The network is as described by Wayne Zarchary <sup>1</sup> . . . . .	2
1.2	Three simple graphs . . . . .	4
1.3	Walks vs. Curves: The values next to the edges define a density $\rho$ . . . . .	5
1.4	Left: House Graph. Right: House graph with the values assigned to each edge by the extremal density of the connecting family $\Gamma(\{1\}, \{2\})$ for $1 < p < \infty$ . . . . .	9
1.5	On the left is an example of a circuit in series. The effective resistance is computed as $R_{\text{eff}} = R_1 + R_2$ . On the right is an example of a circuit in parallel. Here, the effective resistance is computed as $\frac{1}{R_{\text{eff}}} = \frac{1}{R_1} + \frac{1}{R_2}$ . Consequently the effective capacity, satisfies $\frac{1}{C_{\text{eff}}} = \frac{1}{C_1} + \frac{1}{C_2}$ for the circuit in series on the left and it satisfies $C_{\text{eff}} = C_1 + C_2$ for the circuit in parallel on the right. . . . .	14
1.6	Node 2 is depicted as blue and node $N$ is depicted as red. Here $N = 10$ . Despite how quickly $ \Gamma^* $ as the size of the clique grows, the number of paths necessary to consider by the algorithm is relatively very small. . . . .	29
2.2	Despite only having three vertices and three sides, this is an example of quadrangle in a graph. . . . .	37
2.3	From $(G, C)$ to $\tilde{G}$ . . . . .	40
2.4	The canonical bijection from a planar map to a quadrangulation in four steps. . . . .	45

2.6 Steps two through five of creating the planar map  $M$  from the word in (2.3.3).  
It's worth noting the creation of the planar map is only unique up to graph isomorphisms. Also, it is particularly hard to draw the map  $M$  with a planar embedding that uses only straight edges. The graph layout shown would either lose the pattern of straight edges, or lose its planar embedding within the next few steps. . . . . 51

3.1 A table of values of the Spearman Rank coefficient between each centrality and the epidemic hitting time centrality. We note that in every network except the Karate club network, the Effective Resistance outperforms every other centrality. We also remark that in<sup>2</sup>, there was a similar observation of the Karate club network not behaving like the other networks, likely due its being a small network with a few dominating nodes. . . . . 74

3.2 A table of values of the total variation between each centrality and the epidemic hitting time centrality. We note that in every network, the Effective Resistance outperforms every other centrality. This is in large part due to the weight of the peripheral nodes. The effective resistance and epidemic hitting time assign a large enough weight to these peripheral nodes, while the other centralities underestimate their importance. . . . . 75

3.3 Each figure shows the importance of the nodes in the Lesmis network, based off of a different centrality. We observe that not only do the effective resistance and epidemic hitting time centralities visually appear to match up quite well, they are also the only two centralities that assign relative importance to the peripheral nodes. See<sup>3</sup>. . . . . 76

# List of Tables

2.1 Summary of quantitative local observations and their corresponding global geometric properties.  $\tilde{T}$  and  $T$  are the sets of primal diagonal edge-sets before and after (respectively) merging the conformal loop ensemble into a single loop. Similarly,  $\tilde{T}'$  and  $T'$  are the dual diagonal edge-sets before and after merging the conformal loop ensemble into a single loop. . . . . 52



# Acknowledgments

Firstly, I would like to thank my Masters advisor, Professor Pietro Poggi-Corradini. Without your guidance, patience, and direction none of this would have been possible. It was reassuring that you would guide me away from the “rabbit hole” when I began veering off course or into topics that were too deep to learn in the time frame we had allotted, and it continues to be exciting to work with someone who has so many clever and unique ideas to pursue, even as we are face-to-face with deadlines.

I’d also like to thank my committee members...

...Professor Nathan Albin, you have helped guide me with research projects since my very first year as a mathematics major. Sometimes I felt that I would never learn to make my own simulations, but as you promised, with enough practice and enough worthwhile projects, I have hope to eventually become “an expert”.

...Professor Hrant Hakobyan, for always being willing (and seemingly always able) to provide valuable insight to approaching a new problem. Thank you for the time you spent teaching me both structured and reading courses alike.

...Professor Faryad Darabi Sahneh, for the time you spent introducing me to so many aspects of applied graph theory. You were always prepared with a brief explanation and would follow up with providing me a more thorough resource. You provided a great model for how to learn new topics.

There are many professors for whom I am grateful, that have volunteered their own time to help me grow in mathematical maturity...

...Professor Virginia Naibo, for the many intense and interesting courses I have taken from you. You have introduced me to many aspects of mathematics that I hope to pursue further in subsequent years. I was fortunate to have you as my undergraduate advisor.

Thank you for all the advice on courses, REUs, and math conferences. You helped me get on the right track early, and that has had and will continue to have unforetold compounding benefits.

...Professor Ivan Blank, for the excursion through geometric measure theory and partial differential equations. Your generosity to start with a student whose goals at the time were simply, "I like multivariable calculus" and offer your time to teach what grew into truly deep mathematics has undoubtedly helped me discover the beauty of mathematics.

...Professor Dan Vokok, for volunteering time from your schedule to teach me both complex analysis and functional analysis as independent studies. Your introduction to complex analysis opened up so many more course options that I have enjoyed pursuing.

...Professor Sarah Reznikoff, for helping me with the less attractive side of academia. Your assistance with fellowships, REUs, and graduate school applications has saved me so many headaches.

Finally, I would like to thank my family....

...Eric and Ryan, you have both been great role models since we have all grown up. You have both shown me that if I act like a responsible adult when necessary, it opens up lots of time to have fun and still act like one of the three little Goering brothers.

...Mom and Dad, you not only had to put up with me for the first 18 years of my life, but you inexplicably chose to continue to put up with my shenanigans. Without your patience and support, I am not certain I would have even attended college, on top of that be gearing up to be a lifelong academic.

# Dedication

My thesis is dedicated to the memory of my grandfather, Professor Raymond Flory. May your life always be an inspiration.

# Chapter 1

## The Modulus

The modulus of families of walks on graphs (or just modulus) is a generalization of the intuitively familiar concept of capacity from an electrical circuit. The generalization allows us to consider concepts like capacitance and hence effective resistance on multigraphs, directed graphs, and even graphs with self-loops. However the power of the modulus does not end there. The modulus of a family of walks can be thought of as an indicator of the richness of short paths in the family. In fact, the modulus has been shown to be equivalent to the effective resistance, the max-flow/min-cut problem<sup>4</sup>. Since the modulus is equivalent to the effective resistance, it also shares many relations to random walks. Moreover, it can even be used in parabolic vs. hyperbolic type problems on graphs<sup>5</sup>.

### 1.1 Walks on Finite Graphs

#### 1.1.1 Preliminaries

##### Basic Graph Theory

A **simple graph** or **simple network** is an ordered pair  $G = (V, E)$  where  $V$ , the

**vertex-set**, is a collection of objects treated as nodes and  $E$ , the **edge-set**, is a collection of objects treated as links between elements in  $V$ . An element of the edge-set, denoted  $e = \{v_1, v_2\} \in E$  for  $v_1, v_2 \in V$  means that there is an undirected link between nodes  $v_1$  and  $v_2$ . Moreover, in this case, we say that  $v_1$  and  $v_2$  are **neighbors**, denoted  $v_1 \sim v_2$ . Further, a simple graph has no **self-loops**, i.e.,  $\{v, v\} \notin E$  for all  $v \in V$ , and edges between pairs of nodes all have multiplicity one. The above is a complete characterization of simple graphs.

Depending upon the application, the meaning of nodes and edges can vary greatly. For example, see Figure 1.1, Zachary’s Karate Club, for a popular example of a graph representing a social network. Therein, each vertex represents a person and each edge between two vertices represents a friendship.

A graph is called a **finite graph** if the vertex and edge sets are both finite. For a simple graph, it is sufficient that the vertex set is finite, since if  $N$  denotes the number of nodes, then there are at most  $\binom{N}{2}$  edges. It is common to represent a graph by a matrix  $A = (a_{ij})_{1 \leq i, j \leq N}$ , called the **adjacency** matrix, where  $a_{ij} = 1$  if  $v_i \sim v_j$  and zero otherwise.

## Densities and the Modulus

A mapping  $\rho : E \rightarrow [0, \infty]$  is called a **density**. A **walk** (with  $n$  hops) on a graph is a string,  $W := v_0 e_1 v_1 e_2 v_2 \cdots e_n v_n$ , of alternating vertices and edges, so that  $e_k = \{v_{k-1}, v_k\}$ , for each  $k \in \{1, 2, \dots, n\}$ . For a given density  $\rho$ , the corresponding  $\rho$ -length of

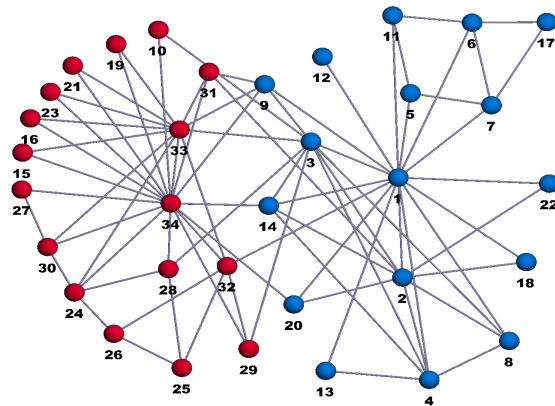


Figure 1.1: Karate Club Network: A network of friendships between the 34 members of a karate club at a US university, just before the club split. The network is as described by Wayne Zarchary<sup>1</sup>

a walk,  $W$ , denoted  $\ell_\rho(W)$  is defined by

$\ell_\rho(W) := \sum_{k=1}^N \rho(e_k)$ . If  $\Gamma$  is a collection of walks on a graph  $G$ , we say that  $\rho$  is **admissible** for  $\Gamma$  if  $\ell_\rho(W) \geq 1$  whenever  $W \in \Gamma$ . Further, we denote the set of admissible densities of a family of walks  $\Gamma$  as  $\text{Adm}(\Gamma)$ . We define the  $p$ -**energy** of a density by  $\mathcal{E}_p(\rho) := \sum_{e \in E} \rho(e)^p$  for  $1 \leq p < \infty$ . Finally, the  $p$ -**modulus** of a family of walks  $\Gamma$  is denoted  $\text{Mod}_p(\Gamma)$  and defined

$$\text{Mod}_p(\Gamma) := \inf_{\rho \in \text{Adm}(\Gamma)} \mathcal{E}_p(\rho).$$

An admissible density  $\rho_0$  that achieves this infimum, is called an **extremal density**.

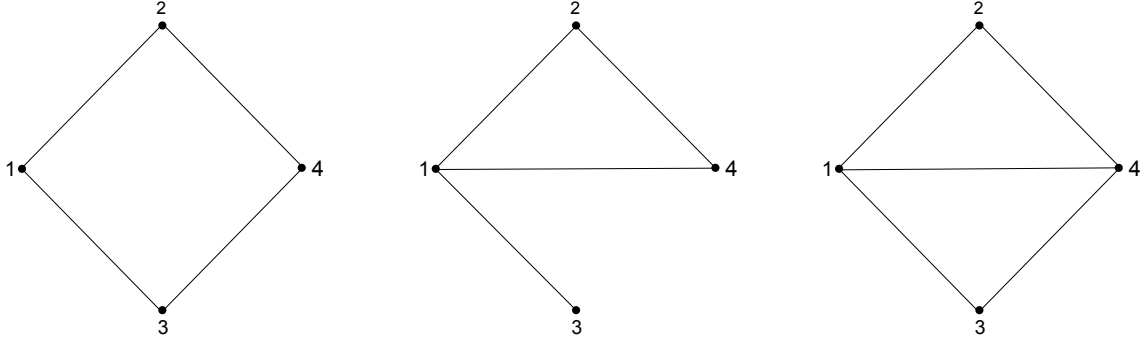
For a walk  $\gamma \in \Gamma$ , we say the **trace of**  $\gamma$  is the subgraph of vertices and edges of  $\gamma$ . That is, the trace of a walk  $\gamma = v_0 e_1 v_1 e_2 \cdots e_n v_n$  is  $J(\gamma) = \left( \{v_j\}_{j=0}^n, \{e_j\}_{j=1}^n \right) = (V(\gamma), E(\gamma))$ . A common type of family of walks is the **connecting family of walks**. Formally, if  $G = (V, E)$  and if  $A, B \subset V$  such that  $A \cap B = \emptyset$ , then the connecting family is defined by

$$\Gamma_G(A, B) := \left\{ W = v_0 e_1 \cdots e_n v_n \mid v_0 \in A \text{ and } v_n \in B, \text{ for some } \{e_j\}_{j=1}^n \subset E, \{v_j\}_{j=0}^n \subset V \right\}.$$

When the underlying graph is understood, we will suppress the dependence of our notation on  $G$ . Furthermore, a subset of vertices  $C \subset V$  is called a **cut** for  $\Gamma_G(A, B)$  if  $\forall \gamma \in \Gamma_G(A, B), V(\gamma) \cap C \neq \emptyset$ .

I have stated that the modulus of families of walks is a good measure of the wealth of short paths. However, this does not seem evident from the definition. To shed light on this matter, we consider several basic graphs and compare the modulus to other intuitive measures of the ‘‘richness of shortest paths.’’ Two such intuitive measures of the wealth of shortest paths are (1) simply count the number of shortest **simple paths**, or walks that repeat no vertex, and (2) to count the length of the shortest path.

**Example 1.1.1.** Consider the three graphs  $G, H$ , and  $K$  depicted in Figure 1.2 and the



$G$  has two simple paths, a shortest path of length two from 1 to 4, and  $\text{Mod}_2(\Gamma_G) = 1$ .

$H$  has two simple paths, a shortest path of length one from 1 to 4, and  $\text{Mod}_2(\Gamma_H) = \frac{3}{2}$ .

$K$  has three simple paths, a shortest path of length one from 1 to 4, and  $\text{Mod}_2(\Gamma_K) = 2$ .

Figure 1.2: Three simple graphs

connecting family of walks  $\Gamma(\{1\}, \{4\})$  for each graph. I claim that:

$$\text{Mod}_2 \Gamma_G(\{1\}, \{4\}) = 1 \quad \text{Mod}_2 \Gamma_H(\{1\}, \{4\}) = \frac{3}{2} \quad \text{Mod}_2 \Gamma_K(\{1\}, \{4\}) = 2.$$

Showing these are the correct values of the modulus is not easy directly from the definition. However, after the next section, namely Theorem 1.1.3, the reader should view verifying these values as a simple exercise. The values of the modulus are desirable, as we would not want to consider any two of these three graphs as having an equivalently rich family of walks from vertex 1 to vertex 4. Intuitively, we would want  $\Gamma_K(\{1\}, \{4\})$  to be strictly richer than  $\Gamma_H(\{1\}, \{4\})$  which is strictly richer than  $\Gamma_G(\{1\}, \{4\})$ . However, we see that the number of simple paths overlaps for  $G$  and  $H$ , and the length of the shortest path overlaps for  $H$  and  $K$ . Hence, only the modulus differentiates between all three graphs.

At best, we still only have a little evidence that the modulus should be a measure of the richness of shortest paths. The following demonstrates how the modulus is directly related to the richness of shortest paths.

**Proposition 1.1.1. (Alternative Definition).** *Let  $\Gamma$  be a non-empty family of non-trivial*

(non-constant) walks. Given a density  $\rho : E \rightarrow [0, \infty]$ , define  $L_\Gamma(\rho) := \inf_{\gamma \in \Gamma} \ell_\rho(\gamma)$ , and let  $S(\Gamma) := \{\rho : L_\Gamma(\rho) > 0\}$ . Then,

$$\text{Mod}_p \Gamma = \inf_{\rho \in S(\Gamma)} \frac{\mathcal{E}_p(\rho)}{L_\Gamma(\rho)^p}.$$

*Proof.* ( $\geq$ ) We note that if  $\rho \in \text{Adm}(\Gamma)$ , then  $L_\Gamma(\rho) \geq 1$ . Therefore,

$$\inf_{\rho \in S(\Gamma)} \frac{\mathcal{E}_p(\rho)}{L_\Gamma(\rho)^p} \leq \inf_{\rho \in \text{Adm}(\Gamma)} \frac{\mathcal{E}_p(\rho)}{L_\Gamma(\rho)^p} \leq \inf_{\rho \in \text{Adm}(\Gamma)} \mathcal{E}_p(\rho) = \text{Mod}_p \Gamma.$$

( $\leq$ ) On the other hand, if  $\tilde{\rho} \in S(\Gamma)$  is an arbitrary density, define  $\rho := \frac{\tilde{\rho}}{L_\Gamma(\tilde{\rho})}$  so that,  $L_\Gamma(\rho) = 1$ . Then  $\rho \in \text{Adm}(\Gamma)$  and

$$\text{Mod}_p(\Gamma) \leq \mathcal{E}_p(\rho) = \frac{1}{L_\Gamma(\tilde{\rho})^p} \sum_{e \in E} \tilde{\rho}(e)^p = \frac{\mathcal{E}_p(\tilde{\rho})}{L_\Gamma(\tilde{\rho})^p}.$$

Taking the infimum over  $\tilde{\rho} \in S(\Gamma)$  completes the proof.  $\square$

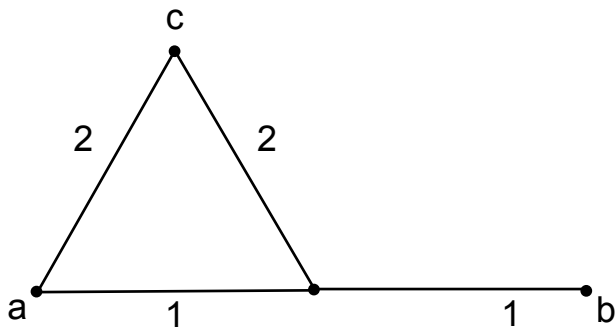


Figure 1.3: Walks vs. Curves: The values next to the edges define a density  $\rho$ .

**Remark 1.1.1.** Another common family of walks is the **via-family of walks**:

$$\Gamma_G(A, B; C) = \{W = v_0 e_1 \cdots e_n v_n \mid v_0 \in A, v_n \in B, \text{ and } v_j \in C \text{ for some } 0 < j < n\}.$$



This quickly leads to the question, what is the shortest walk from  $a$  to  $b$  that traverses  $c$  in Figure 1.3? By first taking the shortest walk from  $a$  to  $c$  and then the shortest walk from  $c$  to  $b$ , one can easily observe that the shortest walk has  $\rho$ -length equal to 5. On the other hand, a **curve** is any connected subgraph of  $G$ . So, the shortest curve between  $a$  and  $b$  that visits  $c$  has length equal to 4. It turns out that finding a shortest walk is a fairly easy problem to handle using, for instance Dijkstra’s algorithm, which runs in polynomial time  $O(N^2)$ . Meanwhile, finding a shortest curve through 3 vertices is an instance of the *Graphical Steiner Minimal Tree Problem*, which is NP-complete. This is the main reason why in the discrete case it is more convenient to work with families of walks than families of curves.

### 1.1.2 Beurling’s Criterion and Extremal Densities

In this section we will discuss the historically relevant and intuitively helpful adaptation of the modulus from the continuous setting. However, in the discrete case, one can gain a “complete” understanding of the modulus from the point of view of convex analysis. Hence, we will begin by borrowing a fact that will be developed in Section 1.2 when we switch our attention to techniques from convex analysis.

**Fact 1.1.2.** *For  $1 < p < \infty$  and any family of walks  $\Gamma$  that contains no constant walks, there exists a unique extremal density  $\rho_p$ . That is, for each  $1 < p < \infty$  there exists  $\rho_p$  such that*

$$\text{Mod}_p \Gamma = \mathcal{E}_p(\rho_p).$$

In Example 1.1.1, I made a claim regarding the value of the 2-modulus for a few simple connecting families of walks but did not attempt to show that those values were indeed the value of the modulus. Since the modulus is defined as an infimum, it is easy to get upper-bounds on the value of the modulus. Typically, it is difficult to attain non-trivial (for instance the modulus is always non-negative) lower bounds. However, Arne Beurling’s

famous criteria for extremality is a very useful sufficient condition to test for extremality of a density.

**Theorem 1.1.3. (Beurling's Extremality Criterion).** *For a fixed  $1 < p < \infty$ , a density  $\rho \in \text{Adm } \Gamma$  is extremal if there exists  $\tilde{\Gamma} \subset \Gamma$  with  $\ell_\rho(\gamma) = 1$  for all  $\gamma \in \tilde{\Gamma}$  such that*

$$\sum_{e \in E} h(e) \rho^{p-1}(e) \geq 0 \text{ whenever } h : E \rightarrow \mathbb{R} \text{ with } \ell_h(\gamma) \geq 0 \text{ for all } \gamma \in \tilde{\Gamma}. \quad (1.1.1)$$

*Proof.* **(Beurling's Criterion)**

Let  $\rho, \tilde{\Gamma}$  be as in the hypothesis of (1.1.1) and let  $\sigma \in \text{Adm } \Gamma$ . Define  $h := \sigma - \rho$ . Then  $\ell_h(\gamma) \geq 0$  for all  $\gamma \in \tilde{\Gamma}$ . By (1.1.1) we have

$$\sum_{e \in E} \sigma(e) \rho(e)^{p-1} - \sum_{e \in E} \rho(e)^p = \sum_{e \in E} h(e) \rho^{p-1}(e) \geq 0,$$

so that

$$0 \leq \sum_{e \in E} \rho(e)^p \leq \sum_{e \in E} \sigma(e) \rho(e)^{p-1} \leq \left( \sum_{e \in E} \sigma(e)^p \right)^{\frac{1}{p}} \left( \sum_{e \in E} \rho(e)^{(p-1)q} \right)^{\frac{1}{q}}, \text{ where } \frac{1}{p} + \frac{1}{q} = 1.$$

The second inequality is an application of Hölder's Inequality. Since  $q = \frac{p}{p-1}$ , the above reads

$$\sum_{e \in E} \rho(e)^p \leq \left( \sum_{e \in E} \sigma(e)^p \right)^{\frac{1}{p}} \left( \sum_{e \in E} \rho(e)^p \right)^{1 - \frac{1}{p}}.$$

Consequently,

$$\left( \sum_{e \in E} \rho(e)^p \right)^{\frac{1}{p}} \leq \left( \sum_{e \in E} \sigma(e)^p \right)^{\frac{1}{p}}.$$

Taking the  $p$ th power of both sides and recalling that  $\sigma$  is an arbitrary admissible density shows that  $\rho$  is indeed extremal.  $\square$

A subfamily  $\tilde{\Gamma}$  that satisfies the hypothesis in Beurling's Extremality Criterion is called a **Beurling subfamily**. For connecting families it is always sufficient, but not necessary to look at the collection of all simple walks in the family of walks  $\Gamma$ . For a given density  $\rho$  and family of walks  $\Gamma$ , we define

$$\Gamma_0(\rho) := \{\gamma \in \Gamma : \ell_\rho(\gamma) = 1\}. \quad (1.1.2)$$

The converse to Beurling's Criterion, which will be proved in [Section 1.2, Theorem 1.2.4] states

**Theorem 1.1.4.** *If  $\rho$  is extremal then  $\Gamma_0(\rho)$  is always a Beurling subfamily.*

Even in the face of a converse to Beurling's criterion, the true power of Beurling subfamilies is the following.

**Theorem 1.1.5.** *If  $\tilde{\Gamma}$  is a Beurling subfamily of  $\Gamma$ , then for each  $1 < p < \infty$  it follows  $\text{Mod}_p \tilde{\Gamma} = \text{Mod}_p \Gamma$ .*

*Proof.* Since  $\tilde{\Gamma}$  is a Beurling subfamily, letting  $\rho_0 \in \text{Adm}(\Gamma)$  be the extremal density for  $\Gamma$ , we see that  $\tilde{\Gamma} \subset \Gamma_0(\rho_0)$  and by hypothesis (1.1.1) holds for  $\rho_0$  with  $\tilde{\Gamma}$ .

By monotonicity,  $\text{Mod}_p(\tilde{\Gamma}) \leq \text{Mod}_p(\Gamma)$ . On the other hand, let  $\rho \in \text{Adm}(\tilde{\Gamma})$ . Define  $h : E \rightarrow \mathbb{R}$  by  $h = \rho - \rho_0$ . Then, for every  $\gamma \in \tilde{\Gamma}$  we have  $\ell_h(\gamma) = \ell_\rho(\gamma) - \ell_{\rho_0}(\gamma) = 0$  and

$$\sum_{e \in E} h(e) \rho_0^{p-1}(e) \geq 0,$$

follows as a consequence of (1.1.1). Expanding out  $h$ , yields

$$\mathcal{E}_p(\rho_0) \leq \sum_{e \in E} \rho(e) \rho_0^{p-1}(e) \leq \left( \sum_{e \in E} \rho(e)^p \right)^{\frac{1}{p}} \left( \sum_{e \in E} \rho_0(e)^p \right)^{1-\frac{1}{p}}.$$

The latter inequality follows by applying Hölder. Dividing over and taking the  $p$ th power achieves the desired result.  $\square$

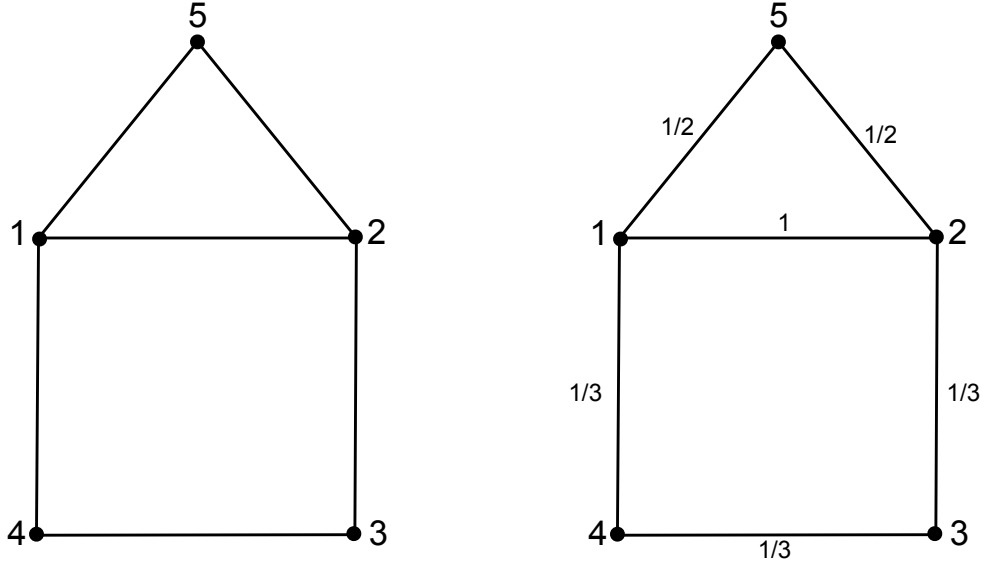


Figure 1.4: Left: House Graph. Right: House graph with the values assigned to each edge by the extremal density of the connecting family  $\Gamma(\{1\}, \{2\})$  for  $1 < p < \infty$ .

**Example 1.1.2.** Consider the House graph  $G$  as in Figure 1.4 and the connecting family of walks  $\Gamma = \Gamma_G(\{1\}, \{2\})$ . If we restrict ourselves to the simple walks in  $\Gamma$ , denoted  $\Gamma_s$ , then it's straightforward to verify that  $\Gamma_s$  only contains the three paths  $P_1 = (1, 2)$ ,  $P_2 = (1, 5, 2)$ ,  $P_3 = (1, 4, 3, 2)$ . For clarity, if  $e = \{x, y\}$ , we write  $h(e) = h(x, y)$ . Letting  $\Gamma_s$  play the role of the Beurling subfamily, we can verify that  $\rho_0$  defined as in the right half of Figure 1.4 is extremal. Indeed, for  $h : E \rightarrow \mathbb{R}$  suppose  $h$  satisfies

$$\ell_h(\gamma) \geq 0 \text{ for all } \gamma \in \Gamma_s. \tag{1.1.3}$$

For each  $k = 1, 2, 3$ , let  $\gamma_k$  be the walk that traverses the vertices of  $P_k$  in the same order. Since the walks  $\gamma_1, \gamma_2$ , and  $\gamma_3$  partition the edges of  $G$  and traverse each edge exactly once,

$$\sum_{e \in E} h(e) \rho_0^{p-1}(e) = \sum_{e \in \gamma_1} h(e) \rho_0^{p-1}(e) + \sum_{e \in \gamma_2} h(e) \rho_0^{p-1}(e) + \sum_{e \in \gamma_3} h(e) \rho_0^{p-1}(e).$$

Moreover, since  $\rho_0$  is constant on each walk  $\gamma_j$ , the above can be reduced to

$$\begin{aligned} \sum_{e \in E} \rho_0^{p-1}(e)h(e) &= \sum_{e \in \gamma_1} h(e) + \left(\frac{1}{2}\right)^{p-1} \sum_{e \in \gamma_2} h(e) + \left(\frac{1}{3}\right)^{p-1} \sum_{e \in \gamma_3} h(e) \\ &= \ell_h(\gamma_1) + 2^{1-p} \ell_h(\gamma_2) + 3^{1-p} \ell_h(\gamma_3) \geq 0, \end{aligned}$$

due to (1.1.3). Therefore,  $\rho_0$  and  $\Gamma_s$  satisfy the hypothesis of Theorem 1.1.3 for each  $p > 1$ . Consequently,  $\rho_0$  is the extremal metric for the  $p$ -modulus on the House graph  $G$ . It is a coincidence, caused by the lack of edges shared by any walks in  $\Gamma_s$ , that the same density  $\rho_0$  is extremal for all  $p$ .

**Remark 1.1.2.** At this point, the reader is encouraged to follow the method of Example 1.1.2 to verify the value of the 2-modulus proposed in Example 1.1.1 for each graph.

### 1.1.3 Properties of the $p$ -Modulus

Here we build the basic properties of the  $p$ -modulus of families of walks and discuss suitable examples to lend intuition to how these properties are used.

**Remark 1.1.3.** If  $\Gamma$  is a family of walks on a finite graph  $G$  that contains a constant walk, i.e., a walk with zero hops. Then  $\text{Mod}_p \Gamma = \infty$ .

Indeed, if  $\gamma_0 \in \Gamma$  is a constant walk, then  $\ell_\rho(\gamma_0) = 0$  for every  $\rho$ -density. Consequently,  $\text{Adm}(\Gamma) = \emptyset$  and  $\inf_{\rho \in \emptyset} \mathcal{E}_p(\rho) = \infty$ .

**Proposition 1.1.6. (The modulus is an outer measure).** *Assume that  $\Gamma_j$  is a family of walks in a finite graph  $G$  for each  $j \in \mathbb{N}$ . Then,*

1. **(Empty Family).** *If  $\Gamma_1 = \emptyset$ , the empty family, then  $\text{Mod}_p \Gamma_1 = 0$ .*
2. **(Monotonicity).** *If  $\Gamma_1 \subset \Gamma_2$ , then  $\text{Mod}_p(\Gamma_1) \leq \text{Mod}_p(\Gamma_2)$ .*
3. **(Countable Subadditivity).**  $\text{Mod}_p \left( \bigcup_{j=1}^{\infty} \Gamma_j \right) \leq \sum_{j=1}^{\infty} \text{Mod}_p \Gamma_j$ .

*Proof.* For each  $j \in \mathbb{N}$  let  $\Gamma_j$  be a family of walks in a finite graph  $G$ .

1. If  $\Gamma_1 = \emptyset$ , then every  $\rho$  is admissible, including  $\rho_0 \equiv 0$ . Hence  $0 \leq \text{Mod}_p \Gamma \leq \mathcal{E}_p(\rho_0) = 0$ .

2. If  $\Gamma_1 \subset \Gamma_2$  then  $\rho \in \text{Adm}(\Gamma_2)$  implies that  $\rho \in \text{Adm}(\Gamma_1)$ , so that  $\text{Adm}(\Gamma_2) \subset \text{Adm}(\Gamma_1)$ .

Therefore,

$$\text{Mod}_p \Gamma_1 = \inf_{\rho \in \text{Adm} \Gamma_1} \mathcal{E}_p(\rho) \leq \inf_{\rho \in \text{Adm} \Gamma_2} \mathcal{E}_p(\rho) = \text{Mod}_p \Gamma_2.$$

3. Let  $\Gamma := \bigcup_{j=1}^{\infty} \Gamma_j$  and fix  $\epsilon > 0$ . For each  $j$ , choose  $\rho_j \in \text{Adm}(\Gamma_j)$  such that

$$\mathcal{E}_p(\rho_j) \leq \text{Mod}_p \Gamma_j + \frac{\epsilon}{2^j}.$$

Define  $\rho := \left( \sum_{j=1}^{\infty} \rho_j^p \right)^{\frac{1}{p}}$ . For any  $\gamma \in \Gamma$ , there exists  $k \in \mathbb{N}$  so that  $\gamma \in \Gamma_k$ . Since  $\rho \geq \rho_k$  we have that  $\ell_\rho(\gamma) \geq 1$ . Hence,  $\rho \in \text{Adm}(\Gamma)$ . Moreover,

$$\begin{aligned} \text{Mod}_p \Gamma &\leq \mathcal{E}_p(\rho) = \sum_{e \in E} \rho(e)^p = \sum_{e \in E} \sum_{j=1}^{\infty} \rho_j(e)^p = \sum_{j=1}^{\infty} \sum_{e \in E} \rho_j(e)^p \\ &= \sum_{j=1}^{\infty} \mathcal{E}_p(\rho_j) \leq \epsilon + \sum_{j=1}^{\infty} \text{Mod}_p \Gamma_j. \end{aligned}$$

We can interchange the order of summation without concern by Tonelli's theorem. Hence, taking  $\epsilon$  to zero attains the desired result.  $\square$

**Remark 1.1.4.** Exclusively using monotonicity, we observe that in Example 1.1.1 we have  $\text{Mod}_2(G) \leq \text{Mod}_2(K)$  and  $\text{Mod}_2(H) \leq \text{Mod}_2(K)$ . Infact, this relationship will hold for  $p \neq 2$  as well.

**Definition 1.1.7.** Given two walks  $\gamma_1$  and  $\gamma_2$  on a graph  $G = (V, E)$  we say that  $\gamma_1 \preceq \gamma_2$ , (in words, either  $\gamma_1$  is *subordinate* to  $\gamma_2$  or equivalently  $\gamma_2$  *dominates*  $\gamma_1$ ) if  $\ell_\rho(\gamma_1) \leq \ell_\rho(\gamma_2)$  for every density  $\rho$ .

We note that in particular it is not required that  $\rho$  is admissible, as the ordering  $\preceq$  is independent of the family of walks.

**Proposition 1.1.8.** *For  $1 \leq p < \infty$ , let  $G = (V, E)$  be a simple finite graph, and let  $\Gamma, \Gamma_1$ , and  $\Gamma_2$  be families of walks.*

- **(Shorter Walks).** *Suppose  $\Gamma_1$  and  $\Gamma_2$  are two families of walks. Further, suppose for every  $\gamma \in \Gamma_2$  there exists  $\sigma \in \Gamma_1$  such that  $\sigma \preceq \gamma$  then  $\text{Mod}_p \Gamma_2 \leq \text{Mod}_p \Gamma_1$ .*
- **(Symmetry).** *Suppose  $T : V \rightarrow V$  is a bijection and a graph isomorphism, that is,  $\{v_1, v_2\} \in E \iff \{T(v_1), T(v_2)\} \in E$  and  $T \circ T = \text{Id}$ . Then for any  $T$ -invariant family of walks  $\Gamma$ , the extremal density for  $\Gamma$  is in the closure of the  $T$ -invariant admissible densities, i.e.,*

$$\text{Mod}_p \Gamma = \inf_{\rho \in A_T(\Gamma)} \mathcal{E}_p(\rho)$$

$$\text{where } A_T(\Gamma) := \left\{ \rho \in \text{Adm}(\Gamma) \mid \rho \circ T = \rho \right\}.$$

*Proof.*

- **(Shorter Walks).** Let  $\rho \in \text{Adm} \Gamma_1$ . Then by hypothesis, for each  $\gamma \in \Gamma_2$ , there exists  $\sigma \in \Gamma_1$  so that  $\sigma \preceq \gamma$ . Therefore,  $1 \leq \ell_\rho(\sigma) \leq \ell_\rho(\gamma)$  and consequently  $\rho \in \text{Adm} \Gamma_2$ . Hence,  $\text{Adm} \Gamma_1 \subset \text{Adm} \Gamma_2$  which yields the desired result:

$$\text{Mod}_p \Gamma_2 = \inf_{\rho \in \text{Adm} \Gamma_2} \mathcal{E}_p(\rho) \leq \inf_{\rho \in \text{Adm} \Gamma_1} \mathcal{E}_p(\rho) = \text{Mod}_p \Gamma_1.$$

- **(Symmetry Rule).** Let  $\rho \in \text{Adm} \Gamma$ . Define  $\rho_1 = \rho \circ T$  so that  $\ell_{\rho_1}(\gamma) = \ell_\rho(T \circ \gamma)$ . Since  $T(\Gamma) = \Gamma$ , this shows  $\rho_1 \in \text{Adm}(\Gamma)$ . Moreover, since  $T$  is an involution, it is in particular a bijection. Hence,  $\mathcal{E}_p(\rho_1) = \mathcal{E}_p(\rho)$ . Next, define  $\rho_2 = \frac{\rho_1 + \rho}{2}$ . It is easily verified that  $\rho_2 \in \text{Adm}(\Gamma)$ <sup>1</sup> and  $\rho_2 = \frac{\rho_1 + \rho}{2} = \left( \frac{\rho + \rho_1}{2} \right) \circ T = \rho_2 \circ T$ . Therefore,

---

<sup>1</sup>In fact, we will see that every convex combination of admissible densities is admissible.

$\rho_2 \in A_T(\Gamma)$ . By the discrete version of Jensen's inequality,

$$\left(\frac{\rho_1(e) + \rho(e)}{2}\right)^p \leq \frac{1}{2}\rho_1(e)^p + \frac{1}{2}\rho(e)^p \quad \forall e \in E, \forall 1 \leq p < \infty,$$

so that,

$$\mathcal{E}_p(\rho_2) = \sum_{e \in E} \left(\frac{\rho_1(e) + \rho(e)}{2}\right)^p \leq \sum_{e \in E} \frac{1}{2}\rho_1(e)^p + \frac{1}{2}\rho(e)^p = \frac{1}{2}(\mathcal{E}_p(\rho_1) + \mathcal{E}_p(\rho)) = \mathcal{E}_p(\rho).$$

Taking the infimum over  $\rho \in \text{Adm } \Gamma$  in the above yields

$$\inf_{\rho \in A_T(\Gamma)} \mathcal{E}_p(\rho) \leq \text{Mod}_p \Gamma.$$

Finally, since  $A_T(\Gamma) \subset \text{Adm } \Gamma$ , the reverse inequality follows.

□

The hypothesis of the symmetry rule may seem complicated at first blush. However, a particularly useful informal interpretation of the symmetry rule is that for any graph  $G$  and family of walks  $\Gamma$ , if there are two connected subgraphs  $H_1 = (V_1, E_1)$  and  $H_2 = (V_2, E_2)$  such that  $H_1$  and  $H_2$  are disjoint and switching  $H_1$  with  $H_2$  does not change  $G$  or  $\Gamma$ , then there exists a bijection  $T : E_1 \rightarrow E_2$  so that all extremal densities  $\rho_p$  for  $p \geq 1$  satisfy  $\rho_p(e) = \rho_p(T(e))$  for all  $e \in E_1$ .

The last result we adapt from the continuous case is reminiscent of the parallel and serial rules for circuits. We will see in the next section that the particular case when  $p = 2$ , not only are there comparisons to circuits, but these rules are precisely the familiar rules for calculating the effective capacitance when circuits are combined in series or in parallel.

**Proposition 1.1.9.**

- **(Parallel Rule).** *Let  $G = (V, E)$  and  $A_j, B_j \subset V$  for  $j = 1, 2$ . Suppose  $\Gamma = \Gamma_1(A_1, B_1) \cup \Gamma_2(A_2, B_2)$  are such that  $E(\gamma_1) \cap E(\gamma_2) = \emptyset$  for all  $\gamma_1 \in \Gamma_1$  and  $\gamma_2 \in \Gamma_2$ .*



Then

$$\text{Mod}_p \Gamma = \text{Mod}_p \Gamma_1 + \text{Mod}_p \Gamma_2.$$

- **(Serial Rule).** Let  $C$  be a cut for  $\Gamma := \Gamma(A_1, A_2)$ . For  $j = 1, 2$  define  $\Gamma_j := \Gamma(A_j, C)$ .

Then for  $1 < p < \infty$  and  $\frac{1}{p} + \frac{1}{q} = 1$ ,

$$\frac{1}{\text{Mod}_p \Gamma} \geq \left\{ \frac{(\text{Mod}_p \Gamma_1)^{\frac{q}{p}} + (\text{Mod}_p \Gamma_2)^{\frac{q}{p}}}{[\text{Mod}_p \Gamma_1 \text{Mod}_p \Gamma_2]^{\frac{q}{p}}} \right\}^{\frac{p}{q}} = \left\{ \left( \frac{1}{\text{Mod}_p \Gamma_1} \right)^{\frac{q}{p}} + \left( \frac{1}{\text{Mod}_p \Gamma_2} \right)^{\frac{q}{p}} \right\}^{\frac{p}{q}}$$

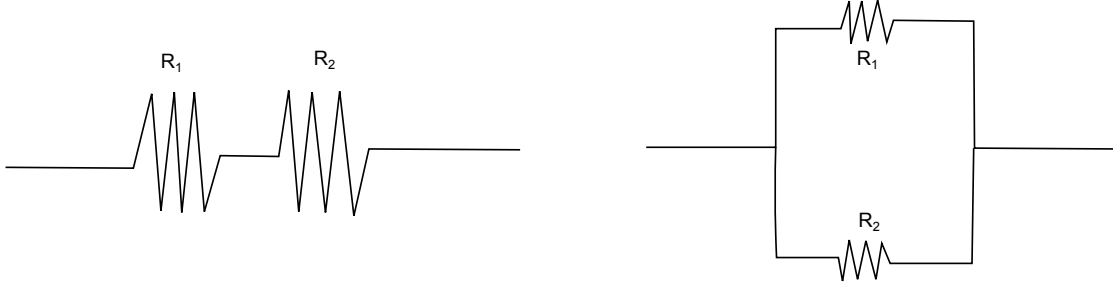


Figure 1.5: On the left is an example of a circuit in series. The effective resistance is computed as  $R_{\text{eff}} = R_1 + R_2$ . On the right is an example of a circuit in parallel. Here, the effective resistance is computed as  $\frac{1}{R_{\text{eff}}} = \frac{1}{R_1} + \frac{1}{R_2}$ . Consequently the effective capacity, satisfies  $\frac{1}{C_{\text{eff}}} = \frac{1}{C_1} + \frac{1}{C_2}$  for the circuit in series on the left and it satisfies  $C_{\text{eff}} = C_1 + C_2$  for the circuit in parallel on the right.

*Proof.*

- **(Parallel Rule).** ( $\leq$ ). This follows from subadditivity of the modulus.

( $\geq$ ). Let  $E_j = \cup_{\gamma \in \Gamma_j} E(\gamma)$  for  $j = 1, 2$ . Then by hypothesis,  $E_1 \cap E_2 = \emptyset$ . For  $\rho \in \text{Adm}(\Gamma)$  define  $\rho_j(e) = \rho(e)\mathbf{1}_{E_j}(e)$ . Then  $\mathcal{E}_p(\rho) \geq \mathcal{E}_p(\rho_1) + \mathcal{E}_p(\rho_2)$ , so that

$$\inf_{\rho_j \in \text{Adm} \Gamma_j} \mathcal{E}_p(\rho_1) + \mathcal{E}_p(\rho_2) \leq \inf_{\rho \in \text{Adm}} \mathcal{E}_p(\rho) = \text{Mod}_p \Gamma.$$

- (Serial Rule). Without loss of generality, assume  $G = (V, E)$  is connected and  $\Gamma_j \neq \emptyset$ . For  $j = 1, 2$ , let  $E_j = \bigcup_{\gamma \in \Gamma_j} E(\gamma)$  and for  $\tilde{\rho}_j \in \text{Adm}(\Gamma_j)$  define  $\rho_j = \tilde{\rho}_j \mathbb{1}_{E_j}$ . Then  $\rho_j \in A(\Gamma_j)$ .

Let  $\rho = a\rho_1 + b\rho_2$  for some  $a+b = 1$  to be determined later. Then for  $\gamma \in \Gamma$ , there exists  $\gamma_j \in \Gamma_j$  such that  $\gamma_j \preceq \gamma$  for  $j = 1, 2$ . Hence,  $1 = a + b \leq a\ell_{\rho_1}(\gamma_1) + b\ell_{\rho_2}(\gamma_2) = \ell_{\rho}(\gamma)$  and consequently  $\rho \in \text{Adm } \Gamma$ .

We observe,

$$\mathcal{E}_p(\rho) = a^p \mathcal{E}_p(\rho_1) + b^p \mathcal{E}_p(\rho_2) =: a^p x + b^p y =: f(a, b).$$

Using Lagrange multipliers, we can minimize  $f$  constrained to  $a + b = 1$ . In doing so, one attains the system of equations,

$$\begin{cases} \lambda p a^{p-1} x = 1 \\ \lambda p b^{p-1} y = 1 \\ a + b - 1 = 0. \end{cases}$$

Since  $a, b, p$ , and  $\text{Mod}_p \Gamma_j$  are non-zero, we solve for  $\lambda$  in the first equation and find  $\lambda = \frac{1}{p a^{p-1} x}$ . Substituting into the second yields,

$$\frac{b^{p-1} y}{a^{p-1} x} = 1.$$

Using  $a = 1 - b$  we attain,

$$\frac{b}{1-b} = \left( \frac{x}{y} \right)^{\frac{1}{p-1}}.$$

Hence,

$$b = \frac{\left( \frac{x}{y} \right)^{\frac{1}{p-1}}}{1 + \left( \frac{x}{y} \right)^{\frac{1}{p-1}}} = \frac{x^{\frac{1}{p-1}}}{y^{\frac{1}{p-1}} + x^{\frac{1}{p-1}}}.$$

Back substitution shows

$$a = \frac{y^{\frac{1}{p-1}}}{y^{\frac{1}{p-1}} + x^{\frac{1}{p-1}}}.$$

With this choice of  $a$  and  $b$  it follows that,

$$\mathcal{E}_p(\rho) = \left( \frac{y^{\frac{1}{p-1}}}{y^{\frac{1}{p-1}} + x^{\frac{1}{p-1}}} \right)^p x + \left( \frac{x^{\frac{1}{p-1}}}{y^{\frac{1}{p-1}} + x^{\frac{1}{p-1}}} \right)^p y = xy \left[ \frac{x^{\frac{1}{p-1}} + y^{\frac{1}{p-1}}}{\left( x^{\frac{1}{p-1}} + y^{\frac{1}{p-1}} \right)^p} \right] = \frac{xy}{\left( x^{\frac{q}{p}} + y^{\frac{q}{p}} \right)^{p-1}},$$

where  $\frac{1}{p} + \frac{1}{q} = 1$ . The penultimate equality follows simply because  $\frac{p}{p-1} = 1 + \frac{1}{p-1}$ . As a consequence of the preceding equation,

$$\text{Mod}_p \Gamma \leq \mathcal{E}_p(\rho) = \frac{xy}{\left( x^{\frac{q}{p}} + y^{\frac{q}{p}} \right)^{p-1}}.$$

By back substituting for  $x$  and  $y$ , simplifying, and taking the infimum over  $\rho_j \in \text{Adm } \Gamma_j$  one attains the generalized serial rule,

$$\frac{1}{\text{Mod}_p \Gamma} \geq \left\{ \frac{(\text{Mod}_p \Gamma_1)^{\frac{q}{p}} + (\text{Mod}_p \Gamma_2)^{\frac{q}{p}}}{[\text{Mod}_p \Gamma_1 \text{Mod}_p \Gamma_2]^{\frac{q}{p}}} \right\}^{\frac{p}{q}} = \left\{ \left( \frac{1}{\text{Mod}_p \Gamma_1} \right)^{\frac{q}{p}} + \left( \frac{1}{\text{Mod}_p \Gamma_2} \right)^{\frac{q}{p}} \right\}^{\frac{p}{q}}.$$

□

### 1.1.4 Capacity

Here, we briefly demonstrate the connection between the 2-modulus and the capacity. For  $G = (V, E)$  and  $A, B$  disjoint subsets of  $V$ , we consider the **potential function**  $u : V \rightarrow \mathbb{R}$  that has the property  $u|_A \leq 0$  and  $u|_B \geq 1$ . The **gradient** of a potential function,

$\rho_u : E \rightarrow [0, \infty)$  is a density defined on an edge  $e = \{v_1, v_2\}$  by  $\rho_u(e) = |u(v_1) - u(v_2)|$ . The ordered pair  $(A, B)$  is called a **condenser** and its **capacity** is defined as

$$\text{Cap}(A, B) := \inf_{\substack{u|_A \leq 0 \\ u|_B \geq 1}} \mathcal{E}_2(\rho_u).$$

The function  $U : V \rightarrow \mathbb{R}$  that attains the infimum is called the **capacitary** function. A capacitary function always exists and is unique. This can be seen by introducing the degree matrix  $D = (d_{ij})$  where  $d_{ii}$  is the degree of the  $i$ th node and  $d_{ij} = 0$  whenever  $i \neq j$ . Then, the combinatorial Laplacian,  $L := D - A$  is symmetric. Hence,

$$\mathcal{E}_2(\rho_u) = \frac{1}{2} \sum_{e \in E} \rho_u(e)^2 = U^T L U,$$

is a quadratic form and minimization can be handled by Lagrange multipliers. Also  $U$  is harmonic, i.e.,  $LU = 0$  on  $V \setminus (A \cup B)$ . Therefore, uniqueness can be derived from the maximum principle of harmonic functions. For more details regarding the capacity and effective resistance on graphs, see<sup>6</sup>.

**Proposition 1.1.10.** *It is always true that*

$$\text{Cap}(A, B) = \text{Mod}_2(\Gamma(A, B)).$$

*Proof.* If  $U$  is the capacitary function for  $(A, B)$  and  $v_0 e_1 v_1 e_2 \cdots e_n v_n = \gamma \in \Gamma(A, B)$ , then

$$\begin{aligned} 1 \leq |U(b) - U(a)| &= \left| \sum_{k=1}^n U(x_k) - U(x_{k-1}) \right| \\ &\leq \sum_{k=1}^n |U(x_k) - U(x_{k-1})| = \sum_{k=1}^n \rho_U(e_k) = \ell_{\rho_U}(\gamma). \end{aligned}$$

Hence  $\rho_U \in \text{Adm}(\Gamma(A, B))$ , so that

$$\text{Mod}_2 \Gamma(A, B) \leq \mathcal{E}_2(\rho_U) = \text{Cap}(A, B).$$

On the other hand, if  $\rho \in \text{Adm}(\Gamma(A, B))$  for  $U$  defined by

$$U(v) = \inf\{\ell_\rho(\gamma) : \gamma = v_0 e_1 v_1 \cdots e_n v \text{ and } v_0 \in A\},$$

we observe  $U \equiv 0$  on the set  $A$ , and  $U \geq 1$  on the set  $B$ . Moreover, if  $e = \{v, w\}$ ,  $U(v) \leq U(w) + \rho(e)$  and  $U(w) \leq U(v) + \rho(e)$  so that  $|U(w) - U(v)| \leq \rho(e)$ . Hence,  $\rho_U \leq \rho$ . If  $U \equiv 1$  on  $B$  then this shows that  $\text{Cap}(A, B) \leq \text{Mod}_2 \Gamma(A, B)$ . So, we consider the case where there exists  $v_j \in B$  such that  $U(v_j) > 1$ . Here, define  $F(v) = \min\{U(v), 1\}$  for each  $v \in V$ . Then for  $e = \{v, w\}$  there are three cases (by switching rolls of  $v$  and  $w$  if necessary).

Case 1: If  $U(v)$  and  $U(w) < 1$ . Then  $\rho_F(e) = \rho_U(e)$ .

Case 2: If  $U(v) \geq 1$  and  $U(w) < 1$ . Then,  $\rho_F(e) \leq \rho_U(e)$  since  $|1 - U(w)| \leq |U(v) - U(w)|$ .

Case 3: If  $1 \leq \min\{U(v), U(w)\}$ . Then,  $0 = \rho_F(e) \leq \rho_U(e)$ .

Thus  $\rho_F(e) \leq \rho_U(e)$  for all  $e \in E$ . Consequently,  $\text{Cap}(A, B) \leq \mathcal{E}(\rho_F) \leq \mathcal{E}(\rho_U) \leq \mathcal{E}(\rho)$ .

Taking the infimum over  $\rho \in \text{Adm}(\Gamma(A, B))$  completes the proof.  $\square$

As a special case, if  $A = \{a\}$  and  $B = \{b\}$  for some  $a, b \in V$  with  $a \neq b$ , we write  $\text{Cap}(a, b) = \text{Mod}_2 \Gamma(a, b)$ , which says the effective conductance, or the multiplicative inverse of the effective resistance, between  $a$  and  $b$  is the *connecting families* 2-modulus (see 1.1.1) between  $\{a\}$  and  $\{b\}$ . Effective resistance can be computed by Equation ER. For more details on the modulus and effective conductance see<sup>6</sup> and for more details on effective resistance see<sup>7</sup>.

When comparing the Parallel and Serial rules to their respective results from physics, keep in mind that  $(\text{Mod}_2 \Gamma_G(A, B))^{-1}$  is interpreted as the effective resistance in the graph  $G$  from  $A$  to  $B$ . (See Figure 1.5.) We now review Proposition 1.1.9 in the case where  $p = 2$ .

**Corollary 1.1.11.**

- **(Parallel Rule).** Let  $G = (V, E)$  and  $A_j, B_j \subset V$  for  $j = 1, 2$ . Suppose  $\Gamma = \Gamma_1(A_1, B_1) \cup \Gamma_2(A_2, B_2)$  are such that  $V(\gamma_1) \cap V(\gamma_2) = \emptyset$  for all  $\gamma_1 \in \Gamma_1$  and  $\gamma_2 \in \Gamma_2$ . Moreover, let  $p = 2$ . Then,

$$\text{Mod}_2 \Gamma = \text{Mod}_2 \Gamma_1 + \text{Mod}_2 \Gamma_2. \quad (1.1.4)$$

- **(Serial Rule).** Let  $C$  be a cut for  $\Gamma := \Gamma(A_1, A_2)$ . For  $j = 1, 2$  define  $\Gamma_j = \Gamma(A_j, C)$ . Then,

$$\frac{1}{\text{Mod}_2 \Gamma} \geq \frac{1}{\text{Mod}_2 \Gamma_1} + \frac{1}{\text{Mod}_2 \Gamma_2}.$$

## 1.2 Discrete Modulus via Convex Optimization.

Here, we see that for any family of walks  $\Gamma$  in a finite graph, there exists a finite subfamily  $\Gamma^* \subset \Gamma$  with the property that  $\text{Adm}(\Gamma^*) = \text{Adm}(\Gamma)$ , and hence  $\text{Mod}_p \Gamma = \text{Mod}_p \Gamma^*$  for all  $1 \leq p$ . This will be very important when we consider the modulus as a convex optimization problem in Section 1.2.2. We call such a subfamily  $\Gamma^*$  an **essential subfamily**.

The subfamily  $\Gamma^*$  is by no means unique, since if  $\text{Adm}(\Gamma^*) = \text{Adm}(\Gamma)$  then for any  $\gamma \in \Gamma \setminus \Gamma^*$  we have  $\text{Adm}(\Gamma^* \cup \{\gamma\}) = \text{Adm}(\Gamma)$ . But, as  $\Gamma$  is frequently an infinite family of walks, knowing the existence of an essential subfamily is immediately beneficial. In fact, in Section 1.2.3 we discuss how existence of an essential subfamily guarantees that an algorithm will terminate in finite time when it is within a fixed  $\epsilon > 0$  of the actual value of the  $p$ -modulus. Moreover, the algorithm that will be described computes the corresponding approximation to the extremal density.

### 1.2.1 Existence of Essential Subfamilies

The construction of  $\Gamma^*$  can be understood through the partial order  $\preceq$  on walks defined in Definition 1.1.7. Now, we define the nonnegative **edge multiplicity of a walk** by

$$m : \Gamma \times E \rightarrow \mathbb{N}_0 \quad m(\gamma, e) = \text{the number of times } \gamma \text{ crosses edge } e.$$

Choosing an enumeration of the edges, the edge multiplicity can be used to define a mapping from  $\Gamma$  into  $\mathbb{N}_0^{|E|}$  by associating to each walk its vector of multiplicities, i.e.,

$$\gamma \mapsto x_\gamma := (m(\gamma, e_1), m(\gamma, e_2), \dots, m(\gamma, e_{|E|})). \quad (1.2.1)$$

Considering walks as vectors in  $\mathbb{R}^{|E|}$ , we have an equivalent definition of the partial order  $\preceq$  in terms of the vector representation of walks. If  $x$  and  $y$  are the vector representations of  $\gamma_x$  and  $\gamma_y$ , then

$$x \preceq y \quad \text{if and only if} \quad x_i \leq y_i \quad \forall i \in \{1, 2, \dots, |E|\}. \quad (1.2.2)$$

The corresponding inequality is strict if in addition  $x \neq y$ . It's straightforward to check the asserted equivalence

$$\gamma_x \preceq \gamma_y \iff x \preceq y, \quad (1.2.3)$$

by defining for each  $i \in \{1, 2, \dots, |E|\}$  the  $\rho$ -density  $\rho_i(e_j) = \delta_{ij}$  since  $\ell_{\rho_k}(\gamma) = m(\gamma, e_k)$ . Hence,  $\gamma_x \preceq \gamma_y$  implies  $x \preceq y$ . On the other hand, for any density  $\rho$  it follows

$$\ell_\rho(\gamma) = \sum_{k=1}^{|E|} m(\gamma, e_k) \rho(e_k). \quad (1.2.4)$$

As a consequence,  $x \preceq y$  implies  $\gamma_x \preceq \gamma_y$ .

To show that an essential subfamily always exists, one must first build-up some machin-

ery. To this end,

**Lemma 1.2.1.** *Every non-empty  $X \subset \mathbb{N}_0^n$  has a minimal element with respect to the partial ordering defined in (1.2.2).*

*Proof.* Define the mapping  $h : \mathbb{N}_0^n \rightarrow \mathbb{N}_0$  by<sup>2</sup>

$$h(x) := \sum_{j=1}^n x_j.$$

By the well-ordering of the natural numbers, the set  $\{h(x) : x \in X\}$  has a smallest value. Let  $x \in X$  be such that  $h(x)$  equals this smallest value. Then  $x$  must be a minimal element. Indeed, suppose that  $y \in X$  with  $y \preceq x$ . Then  $y_i \leq x_i$  for all  $i = 1, 2, \dots, n$ . But  $h(x) \leq h(y)$  implies that no  $y_i$  can be strictly less than the corresponding  $x_i$  and consequently  $y = x$ .

□

**Theorem 1.2.2.** *Let  $X \subset \mathbb{N}_0^n$ . Then there exists a finite subset  $X^* \subset X$  such that*

$$X = \cup_{x^* \in X^*} \{x \in X : x^* \preceq x\}.$$

*Proof.*

We proceed by induction on  $n$ . If  $n = 1$ , the theorem follows since  $\mathbb{N}_0$  is well-ordered;  $X^*$  can be taken as the minimum of  $X$ .

Now supposed the result holds for dimensions  $n = 1, 2, \dots, k - 1$ . Then we need to show the result holds for dimension  $n = k$ . To this end, let  $X \subset \mathbb{N}_0^k$ . If  $X$  is empty, there is no work to be done. Otherwise, let  $x^0$  be a minimal element of  $X$ , as guaranteed by Lemma 1.2.1. Define  $X^0$  to be the set of elements in  $X$  that do not dominate  $x^0$ , i.e.,

$$X^0 := \{x \in X : x^0 \not\preceq x\}.$$

---

<sup>2</sup>When  $x = x_\gamma$ , the function  $h$  counts the number of edges traversed (with multiplicity) by the walk  $\gamma$ .



For each subset  $I \subset \mathcal{I} := \{1, 2, \dots, n\}$ , define

$$X_I^0 := \{x \in X^0 : x_i < x_i^0 \ \forall i \in I, \text{ and } x_i \geq x_i^0, \ \forall i \notin I\}.$$

That is  $X_I^0$  is the subset of  $X$  where the domination fails exclusively on the components in  $I$ . Moreover,

$$X^0 = \bigcup_{I \subset \mathcal{I}} X_I^0.$$

Since for every  $I \subset \mathcal{I}$ ,  $X_I^0 \subset X^0$ , and  $x \in X^0$  implies  $x^0 \not\preceq x$ , it follows that  $X_\emptyset^0 = \emptyset$ , since  $x \in X_\emptyset^0$  must dominate  $x^0$  in every coordinate, contradicting  $x \in X^0$ .

On the other extreme,  $X_{\mathcal{I}}^0 = \emptyset$ , since  $x \in X_{\mathcal{I}}^0$  means  $x \prec x^0$  contradicting the minimality of  $x^0$ .

Thus, we restrict our attention to  $I \subset \mathcal{I}$  such that  $1 \leq |I| \leq k-1$  and  $X_I^0$  is non-empty. Let  $\ell := |I|$ . Then we have seen that  $1 \leq \ell \leq k-1$ . Hence, by reordering, any  $x \in X_I^0$  can be rewritten as  $x = (x', x'')$  where  $x' \prec (x^0)'$  and  $(x^0)'' \preceq x''$ , and  $x'$  is made up of the first  $\ell$  entries of  $x$  and  $x''$  makes up the remaining  $k-\ell$  entries.

To simplify the notation, we consider projections  $p_I, q_I$  where  $p_I(x) = x'$  and  $q_I(x) = x''$ . Since  $\mathbb{N}_0^n$  is finite dimensional and in each dimension there are only finitely many values in  $\mathbb{N}_0$  smaller than  $x_i^0$  for each component  $i = 1, 2, \dots, \ell$ , then there can only be finitely many images of projections  $p_I(x_{I,j})$  so that  $p_{I,j}(x_{I,j}) \prec (x^0)'$  for some  $x_{I,j} \in X_I^0$  and  $j = 1, 2, \dots, m_I$ . Hence, we let  $\{z_{I,j}\}_{j=1}^{m_I} \subset \mathbb{N}_0^\ell$  be an enumeration of vectors satisfying  $z_{I,j} = p_I(x_{I,j})$  for  $x_{I,j}$  as described. Then for each  $I \subset \mathcal{I}$  and  $j = 1, 2, \dots, m_I$  we define the set

$$X_{I,j}^0 := \left\{ q_I(x) \in \mathbb{N}_0^{k-\ell} : (z_j, q_I(x_{I,j})) \in X_I^0 \right\}.$$

Since  $1 \leq k-\ell \leq k-1$ , the inductive hypothesis applies to  $X_{I,j}^0$  so that there exists some  $Y^*$ , an essential subset of  $X_{I,j}^0$ . Consequently, the set

$$X_{I,j}^* := \{(z_{I,j}, y) : y \in Y^*\}$$

is finite for each  $j = 1, 2, \dots, m_I$  and has the property that every  $x \in X_I^0$  with  $x = (z_j, x'')$  dominates some element of  $X_{I,j}^*$ .

Being careful to properly re-index appropriately for each  $I \subset \mathcal{I}$ , we can now define

$$X_I^* := \bigcup_{j=1}^{m_I} X_{I,j}^*, \quad \text{and} \quad X^* := \{x_0\} \cup \left\{ \bigcup_{I \subset \mathcal{I}} X_I^* \right\},$$

where  $X_I^* = \emptyset$  whenever  $X_I^0 = \emptyset$ . Since each  $X_I^*$  is a finite union of finite sets, and  $X^*$  is a finite union of the sets  $X_I^*$ , we have that  $X^*$  is finite. Moreover, given any  $x \in X$ , either  $x_0 \preceq x$  or  $x \in X_{I,i}^0$  for some  $I$  and  $j$  and hence dominates at least one  $x^* \in X_{I,j}^*$ . Thus,  $X^*$  has the desired properties and is a finite subset of  $X$ .  $\square$

Now we have the necessary tools to prove the main result.

**Theorem 1.2.3.** *Any family of walks  $\Gamma$  admits an essential subfamily.*

*Proof.* The mapping of  $\Gamma$  into  $\mathbb{N}_0^{|E|}$  defined in (1.2.1) preserves partial orders as described in (1.2.3). Therefore, let  $X = \{x_\gamma : \gamma \in \Gamma\}$  and  $X^* \subset X$  be the finite essential subset guaranteed by Theorem 1.2.2. We can construct  $\Gamma^*$  by adding one<sup>3</sup>  $\gamma$  to  $\Gamma^*$  for each  $x_\gamma \in X^*$ . Then  $\Gamma^*$  is finite. Moreover, since the mapping preserves partial orders, every  $\gamma \in \Gamma$  dominates some element of  $\Gamma^*$ . Thus, if  $\rho \in \text{Adm}(\Gamma^*)$  and  $\gamma \in \Gamma$  then by the design of  $\Gamma^*$  there exists  $\gamma^* \in \Gamma^*$  so that

$$\ell_\rho(\gamma) \geq \ell_\rho(\gamma^*) \geq 1.$$

Hence,  $\text{Adm}(\Gamma^*) \subset \text{Adm}(\Gamma)$ . On the other hand,  $\Gamma^* \subset \Gamma$ , implies  $\text{Adm}(\Gamma^*) \supset \text{Adm}(\Gamma)$ , attaining both inclusions.  $\square$

---

<sup>3</sup>More than one walk  $\gamma$  can correspond to each  $x_\gamma$ . For example, consider  $\Gamma$  to be the set of all walks that are loops with 5 hops. Then for each  $x_\gamma$  the loop can start in 5 different nodes and have 2 different directions, so that there are 10 unique  $\gamma \in \Gamma$  for each  $x_\gamma$  in this particular family of walks.

## 1.2.2 Karush-Kuhn-Tucker conditions and the Modulus as a Convex Program

We say that  $v$  is a **linear combination** of  $\{v_k\}_{k \in \mathcal{I}}$ , where  $\mathcal{I}$  is an arbitrary index set if there exists a finite  $I \subset \mathcal{I}$  such that

$$v = \sum_{k \in I} a_k v_k = a_{k_1} v_{k_1} + a_{k_2} v_{k_2} + \dots + a_{k_m} v_{k_m}. \quad (1.2.5)$$

For vector spaces  $X, Y$ , a function  $f : X \rightarrow Y$  is said to be linear if it preserves linear combinations, i.e.,

$$f\left(\sum_{j=1}^m a_j v_j\right) = \sum_{j=1}^m a_j f(v_j) \quad \forall a_j \in \mathbb{R}, \forall v_j \in X.$$

A convex combination is a special case of a linear combination. We say that  $v$  is a **convex combination** of  $\{v_k\}_{k \in I}$  if  $v$  can be expressed as in 1.2.5 with the additional constraint that  $\sum_{k \in I} a_k = 1$  and each  $a_k \geq 0$ . Any set that is closed under convex combinations is called a **convex set**. For example,  $\text{Adm}(\Gamma)$  is a convex set, since  $\rho := a_1 \rho_1 + \dots + a_m \rho_m \in \text{Adm}(\Gamma)$  whenever  $a_1 + \dots + a_m = 1$ , as demonstrated below.

$$\rho(\gamma) = \sum_{k=1}^m a_k \rho_k(\gamma) \geq \sum_{k=1}^m a_k = 1, \quad \forall \gamma \in \Gamma,$$

where the inequality follows since  $\rho_k \in \text{Adm}(\Gamma)$  for all  $k$ .

A function  $f : X \rightarrow \mathbb{R}$  is called **convex** if for any  $v = a_1 v_1 + \dots + a_m v_m$ , a convex combination of elements of  $X$ , it follows that

$$f(v) \leq a_1 f(v_1) + \dots + a_m f(v_m). \quad (1.2.6)$$

For example, the mappings  $x \rightarrow x^p$  are convex for  $p \geq 1$ .

An **affine transformation** is a special case of a convex combination where equality holds in (1.2.6). Less formally, an affine transformation is a translation of a linear transformation. In particular, the Euclidean dot product against a fixed vector  $x$ , denoted  $\langle \cdot, x \rangle : \mathbb{R}^m \rightarrow \mathbb{R}$ , is a linear transformation, so that  $1 - \langle \cdot, x \rangle : \mathbb{R}^m \rightarrow \mathbb{R}$  is an affine transformation, and consequently a convex function.

A **standard convex optimization problem** is an optimization problem where one minimizes a convex functional over a convex set, subject to finitely many convex equality and inequality constraints. We will show that in the discrete setting, computing the modulus is a standard convex optimization problem.

In the preceding section, we have already discussed that after enumerating the edges, each density  $\rho$  can be thought of as a vector in  $\mathbb{R}^{|E|}$  by  $\rho = (\rho(e_1), \rho(e_2), \dots, \rho(e_n))$ . Similarly, by (1.2.1), every walk  $\gamma$  corresponds to a vector  $x_\gamma \in \mathbb{N}_0^{|E|} \subset \mathbb{R}^{|E|}$ . In (1.2.4) it was demonstrated that the  $\rho$ -length is simply the dot product in  $\mathbb{R}^{|E|}$  of a vector with a  $\rho$ -density. With these observations and the examples of convexity and convex functions given above, we can write the modulus as a standard convex optimization problem. A first (very reasonable) attempt, would look like

$$\begin{aligned} & \text{minimize} && \mathcal{E}_p(\rho) \\ & \text{subject to} && \rho \in \text{Adm}(\Gamma). \end{aligned}$$

Expanding the meaning of this, we have (after enumerating the edges) the equivalent convex program,

$$\begin{aligned} & \text{minimize} && \sum_{k=1}^{|E|} \rho(e_k)^p \\ & \text{subject to} && 1 - \sum_{k=1}^{|E|} \rho(e_k) m(\gamma, e_k) \leq 0 \quad \forall \gamma \in \Gamma. \end{aligned}$$

We see that the functional  $\mathcal{E}_p(\rho)$  is convex, as it is a linear combination of convex functions. Further, one can observe that the above looks like a convex program with zero equality constraints. However, the astute reader will have observed the difficulty that there are  $|\Gamma|$  (and hence, potentially infinitely many) inequality constraints. Therefore, we have not yet written the modulus as a standard convex optimization problem. Moreover, Theorem 1.2.3, guarantees the existence of an essential subfamily  $\Gamma^*$  so that we instead consider,

$$\begin{aligned} & \text{minimize} && \mathcal{E}_p(\rho) \\ & \text{subject to} && \rho \in \text{Adm}(\Gamma^*), \end{aligned}$$

which is a standard convex optimization problem with finitely many inequality constraints. After enumerating  $\Gamma^*$  we can expand this as an “ordinary convex program” (language from 8)

$$\begin{aligned} & \text{minimize} && \mathcal{E}_p(\rho) && (1.2.7) \\ & \text{subject to} && 1 - \ell_\rho(\gamma_k) \leq 0 && k = 1, 2, \dots, m, \end{aligned}$$

where  $m := |\Gamma^*|$ .

How does writing the modulus as an ordinary convex program help? Since convex analysis is a large field of mathematics, this provides a whole new set of tools to use to inspect the modulus. In particular, we define the Lagrangian, which should be reminiscent of Lagrange multipliers.

$$L(\rho, \lambda) := \mathcal{E}_p(\rho) + \sum_{k=1}^m \lambda_k (1 - \ell_\rho(\gamma_k)). \quad (1.2.8)$$

An optimization problem is said to have the property of **strong duality** if the Lagrangian has a saddle point, i.e., if there exists some  $(\rho^*, \lambda^*)$  so that

$$L(\rho^*, \lambda) \leq L(\rho^*, \lambda^*) \leq L(\rho, \lambda^*) \quad \forall \rho : E \rightarrow \mathbb{R}, \quad \forall \lambda \in \mathbb{R}_+^m.$$

For convex problems, Slater's condition (see<sup>8</sup> Theorem 28.2) states that if the optimal value is bounded below, then the existence of a strictly feasible point is sufficient to ensure strong duality. A **strictly feasible** point, is a point where every inequality constraint holds with strict inequality. Since the modulus is bounded below by zero, so long as  $\Gamma$  does not have any walks with zero hops, it is sufficient to consider the particular  $\rho \equiv 2$  to show that there is indeed a strictly feasible point.

Since the ordinary problem is convex, sufficiently smooth (in fact infinitely smooth), and exhibits strong duality, the Karush-Kuhn-Tucker (KKT) conditions provide necessary and sufficient conditions for optimality. The KKT conditions ensure the existence of an optimal  $\rho^* : E \rightarrow \mathbb{R}$  and dual optimal  $\lambda^* \in \mathbb{R}_+^m$  (<sup>8</sup> Theorem 28.3) satisfying

$$\lambda_k^* \geq 0, \quad 1 - \ell_{\rho^*}(\gamma_k) \leq 0 \quad \text{for } k = 1, 2, \dots, m \quad (1.2.9)$$

$$\lambda_k^*(1 - \ell_{\rho^*}(\gamma_k)) = 0 \quad \text{for } k = 1, 2, \dots, m \quad (1.2.10)$$

$$\nabla_{\rho} L(\rho^*, \lambda^*) = 0. \quad (1.2.11)$$

In addition to being sufficient conditions that an optimal  $\rho^*$  exists, in the case where  $p > 1$  and the functional  $\mathcal{E}_p(\rho)$  is *strictly* convex, the KKT conditions also guarantee *uniqueness* of the solution. The density  $\rho^*$  is the minimizer of (1.2.7). Moreover, the property exhibited in (1.2.10), called **complementary slackness** provides a meaningful interpretation of  $\tilde{\Gamma}$  from Beurling's criterion in terms of the values of the Lagrange multipliers  $\lambda^*$  in Beurling's Criterion.

**Theorem 1.2.4.** *Assume  $\Gamma$  is a finite (and enumerated) family of walks. Let  $\rho^*$  and  $\lambda^*$  be a saddle point for the Lagrangian  $L$ . Then*

$$\tilde{\Gamma} := \{\gamma_k : \lambda_k^* > 0\} \quad (1.2.12)$$

*can be taken to be the subfamily  $\tilde{\Gamma}$  in Beurling's Criterion (Theorem 1.1.3). The converse*

of Beurling's Criterion (Theorem 1.1.4) follows as a consequence.

*Proof.* First, if  $\tilde{\Gamma}$  is as in (1.2.12) and  $\Gamma_0(\rho^*)$  as in (1.1.2), then complementary slackness guarantees  $\tilde{\Gamma} \subset \Gamma_0(\rho^*)$ .

Next, suppose  $h : E \rightarrow \mathbb{R}$  satisfies  $\ell_h(\gamma) \geq 0$  for all  $\gamma \in \tilde{\Gamma}$ , and that  $\mu > 0$ .

The saddle-point property ensures that

$$\begin{aligned}
0 &\leq L(\rho^* + \mu h, \lambda^*) - L(\rho^*, \lambda^*) \\
&= \sum_{e \in E} (\rho^*(e) + \mu h(e))^p - \sum_{e \in E} \rho^*(e)^p + \sum_{\substack{k=1 \\ \lambda_k^* > 0}}^m \lambda_k^* (1 - \ell_{\rho^* + \mu h}(\gamma_k)) \\
&= \sum_{e \in E} (\rho^*(e) + \mu h(e))^p - \sum_{e \in E} \rho^*(e)^p + \sum_{\substack{k=1 \\ \lambda_k^* > 0}}^m \lambda_k^* (1 - \ell_{\rho^*}(\gamma_k) - \ell_{\mu h}(\gamma_k)) \\
&= \sum_{e \in E} (\rho^*(e) + \mu h(e))^p - \sum_{e \in E} \rho^*(e)^p - \mu \sum_{\substack{k=1 \\ \lambda_k^* > 0}}^m \lambda_k^* \ell_h(\gamma_k) \\
&= \mu \left[ \frac{\sum_{e \in E} (\rho^*(e) + \mu h(e))^p - \sum_{e \in E} \rho^*(e)^p}{\mu} - \sum_{\substack{k=1 \\ \lambda_k^* > 0}}^m \lambda_k^* \ell_h(\gamma_k) \right].
\end{aligned}$$

The fourth line follows since  $\ell_{\rho^*}(\gamma_k) = 1$  for all  $\gamma_k \in \tilde{\Gamma}$ .

Moreover, since  $\mu > 0$ , we have

$$\sum_{\substack{k=1 \\ \lambda_k^* > 0}}^m \lambda_k^* \ell_h(\gamma_k) \leq \frac{\sum_{e \in E} (\rho^*(e) + \mu h(e))^p - \sum_{e \in E} \rho^*(e)^p}{\mu} \quad \forall \mu > 0.$$

Taking the limit as  $\mu \downarrow 0$  achieves,

$$\sum_{\substack{k=1 \\ \lambda_k^* > 0}}^m \lambda_k^* \ell_h(\gamma_k) \leq \nabla [(\rho^*(\vec{e}))^p] \cdot h(\vec{e}) = p \sum_{e \in E} \rho(e)^{p-1} h(e).$$

By the hypothesis on  $h$ , and the fact that  $0 < 1 < p$  we have shown that (1.1.1) holds.

□

### 1.2.3 Computing the Modulus

There are many computational libraries that exist to solve convex optimization problems. So, it would not be unreasonable to believe that after seeing the discrete version of the modulus can always be written as in (1.2.7), that all computational difficulties have been handled.<sup>4</sup> However, it can be very difficult to find a small  $\Gamma^*$ , so the number of inequality constraints, that is the cardinality of  $\Gamma^*$ , is reasonable for computation.

To expand upon this idea we introduce the notion of a choked graph in Figure 1.6. It is not hard to imagine a generalization of this graph, to a complete graph on  $N - 1$  nodes. After enumerating this complete graph, we add an  $N$ th node, and only connect it to node 1. Then, consider the family of walks from node 2 to node  $N$ , denoted  $\Gamma(2, N)$  and the family  $\Gamma_s := \{\gamma : 2 \mapsto N \mid \gamma \text{ is simple}\}$  as a candidate for  $\Gamma^*$ .

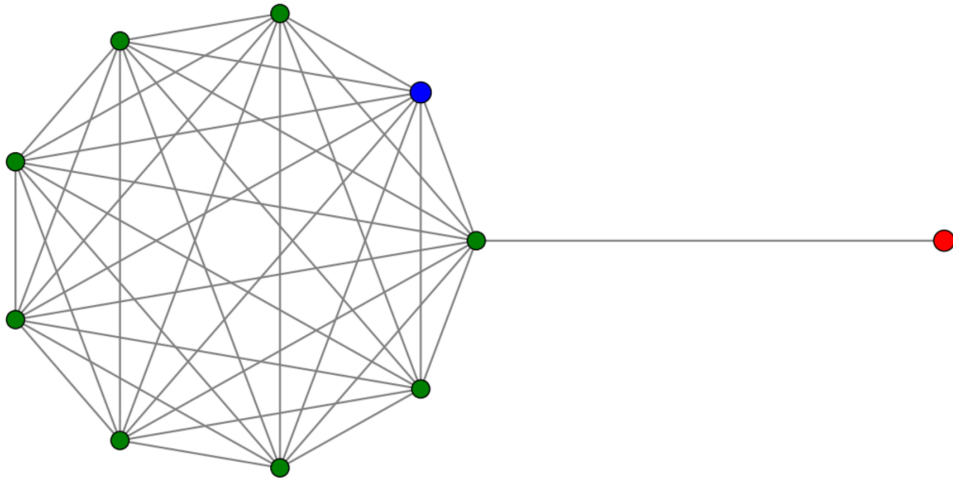


Figure 1.6: Node 2 is depicted as blue and node  $N$  is depicted as red. Here  $N = 10$ . Despite how quickly  $|\Gamma^*|$  as the size of the clique grows, the number of paths necessary to consider by the algorithm is relatively very small.

Since the proper subset of  $\Gamma_s$  that is defined as the subset of walks that each have

---

<sup>4</sup>Especially since it is easy to see that the family  $\Gamma_s = \{\text{all simple walks in the family } \Gamma\}$  is a superset of the necessary family of constraints  $\Gamma^*$ .



precisely  $N - 1$  hops, has exactly  $(N - 3)!$  walks in it, it's clear that  $|\Gamma_s| > (N - 3)!$ .

However, the edge connecting node 1 and  $N$  represents a “choke point” or bottleneck, where every path must traverse this edge. Therefore, it is natural to expect that there may exist a relatively small family  $\Gamma' \subset \Gamma$  such that  $\text{Mod}_p(\Gamma') \approx \text{Mod}_p(\Gamma)$ .

In this section, an algorithm for approximating the modulus of a family of curves  $\Gamma$  is presented. This algorithm performs especially well in cases where such a small  $\Gamma'$  exists. In the following, we assume that there exists an algorithm, denoted  $\text{shortest}(\rho)$ , which produces a  $\rho$ -shortest walk that exists in  $\Gamma$ . That is, given any  $\rho : E \rightarrow [0, \infty)$

$$\gamma^* = \text{shortest}(\rho) \quad \implies \quad \forall \gamma \in \Gamma : \ell_\rho(\gamma^*) \leq \ell_\rho(\gamma).$$

For connecting families of walks, there are many well-known and rapid algorithms that satisfy the job of  $\text{shortest}(\rho)$ , for example, Dijkstra's algorithm, see<sup>9</sup> is discussed in Section 3.1. Pseudocode for the modulus approximation algorithm is given in Algorithm 1.

---

**Algorithm 1** Approximates  $\text{Mod}_p(\Gamma)$  with an error tolerance of  $0 < \epsilon_{\text{tol}} < 1$ .

---

```

 $\rho \leftarrow 0$ 
 $\Gamma' \leftarrow \emptyset$ 
loop
   $\gamma \leftarrow \text{shortest}(\rho)$ 
  if  $\ell_\rho(\gamma)^p \geq 1 - \epsilon_{\text{tol}}$  then
    stop
  end if
   $\Gamma' \leftarrow \Gamma' \cup \{\gamma\}$ 
   $\rho \leftarrow \text{argmin}\{\mathcal{E}_p(\rho) : \rho \in \text{Adm}(\Gamma')\}$ 
end loop

```

---

During each iteration through the loop, a shortest walk is chosen from  $\Gamma$  using the provided  $\text{shortest}$  algorithm. If the stopping criterion is not met, the new walk is added to  $\Gamma'$  and the convex optimization problem described in (1.2.7) is solved using a standard convex programming package.

In an attempt to simplify notation for the next theorem and its proof, we observe that

$\mathcal{E}_p(\rho) = \|\rho\|_p^p$ . Moreover, we will have continue to let  $(p, q)$  be Hölder exponents, that is,  $\frac{1}{p} + \frac{1}{q} = 1$ .

**Theorem 1.2.5.** *Let  $\Gamma$  be a family of walks on a finite graph and suppose that  $\rho^*$  is the extremal density for  $\text{Mod}_p(\Gamma)$  with  $1 < p < \infty$ . Fix an error tolerance  $0 < \epsilon_{tol} < 1$ . Then, Algorithm 1 will terminate in finite time, and will output a family  $\Gamma' \subset \Gamma$  and a density  $\rho$ .*

Moreover,  $\Gamma'$  has the property that

$$0 \leq \frac{\text{Mod}_p(\Gamma) - \text{Mod}_p(\Gamma')}{\text{Mod}_p(\Gamma)} \leq \epsilon_{tol}, \quad (1.2.13)$$

while  $\rho$  satisfies

$$\frac{\|\rho^* - \rho\|_p}{\|\rho^*\|_p} \leq \begin{cases} 2^{\frac{1}{q}} \epsilon_{tol}^{\frac{1}{p}} & p \geq 2 \\ 2^{\frac{1}{p}} \left(\frac{q}{p} \epsilon_{tol}\right)^{\frac{1}{q}} & 1 < p < 2. \end{cases} \quad (1.2.14)$$

*Proof.* In Theorem 1.2.3, we saw that the size of  $\Gamma'$  is bounded above by  $|\Gamma_s|$ , some finite number. Therefore, since the size of  $\Gamma'$  is monotonically increasing with integer step-size within the loop, the algorithm will definitely terminate in finite time. Despite the upper bound that guarantees termination of the algorithm, we will see in some examples that it actually terminates in relatively very few iterations.

To see that (1.2.13) holds, we observe that the identity  $\mathcal{E}_p(\rho) = \text{Mod}_p(\Gamma')$  is a loop invariant. Moreover, the loop can only terminate if  $\ell := \ell_\rho(\gamma) > 0$ . However, when  $\ell > 0$ , then the choice of  $\gamma$  in the loop implies that  $\frac{1}{\ell}\rho \in \text{Adm}(\Gamma)$ . Consequently,

$$\text{Mod}_p(\Gamma) \leq \mathcal{E}_p\left(\frac{\rho}{\ell}\right) = \frac{1}{\ell^p} \mathcal{E}_p(\rho) = \frac{1}{\ell^p} \text{Mod}_p(\Gamma').$$

When the loop terminates,  $\ell^p \geq 1 - \epsilon_{tol}$ , which implies

$$\text{Mod}_p(\Gamma) \leq (1 - \epsilon_{tol})^{-1} \text{Mod}_p(\Gamma').$$

Hence,

$$1 - \epsilon_{\text{tol}} \leq \frac{\text{Mod}_p(\Gamma')}{\text{Mod}_p(\Gamma)}$$

so that

$$\frac{\text{Mod}_p(\Gamma) - \text{Mod}_p(\Gamma')}{\text{Mod}_p(\Gamma)} = 1 - \frac{\text{Mod}_p(\Gamma')}{\text{Mod}_p(\Gamma)} \leq 1 - (1 - \epsilon_{\text{tol}}) = \epsilon_{\text{tol}},$$

where the lower bound in (1.2.13) follows by monotonicity.

In order to prove (1.2.14), consider the Clarkson inequality for  $p \geq 2$ :

$$\mathcal{E}_p(\rho + \rho^*) + \mathcal{E}_p(\rho - \rho^*) \leq 2^{p-1}(\mathcal{E}_p(\rho) + \mathcal{E}_p(\rho^*)). \quad (1.2.15)$$

Since  $\rho^* \in \text{Adm}(\Gamma)$  (because it is extremal) and  $(1 - \epsilon_{\text{tol}})^{\frac{-1}{p}} \rho \in \text{Adm}(\Gamma)$  by the stopping condition of Algorithm 1, then

$$\frac{\rho + \rho^*}{1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}}} \in \text{Adm}(\Gamma),$$

since

$$\ell_{\rho^* + \rho}(\gamma) = \ell_{\rho^*}(\gamma) + \ell_{\rho}(\gamma) \geq 1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}} \quad \forall \gamma \in \Gamma.$$

Hence,

$$\mathcal{E}_p(\rho + \rho^*) \geq \left(1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}}\right)^p \text{Mod}_p(\Gamma). \quad (1.2.16)$$

Moreover, since  $\rho^*$  is extremal and  $\Gamma' \subset \Gamma$ , we know that

$$\mathcal{E}_p(\rho) = \text{Mod}_p(\Gamma') \leq \text{Mod}_p(\Gamma) = \mathcal{E}_p(\rho^*). \quad (1.2.17)$$

Substituting (1.2.16) and (1.2.17) into (1.2.15) (in the second and third lines respectively)

yields

$$\begin{aligned}
\mathcal{E}_p(\rho - \rho^*) &\leq 2^{p-1} (\mathcal{E}_p(\rho) + \mathcal{E}_p(\rho^*)) - \mathcal{E}_p(\rho + \rho^*) \\
&\leq 2^{p-1} [\mathcal{E}_p(\rho) + \mathcal{E}_p(\rho^*)] - \left(1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}}\right)^p \text{Mod}_p(\Gamma) \\
&\leq 2^{p-1} (2 \text{Mod}_p(\Gamma)) - \left(1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}}\right)^p \text{Mod}_p(\Gamma) \\
&= 2^p \text{Mod}_p(\Gamma) \left[1 - \left(\frac{1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}}}{2}\right)^p\right] \\
&\leq 2^p \text{Mod}_p(\Gamma) \left[1 - \frac{1 + (1 - \epsilon_{\text{tol}})}{2}\right] \\
&= 2^{p-1} \text{Mod}_p(\Gamma) \epsilon_{\text{tol}}.
\end{aligned}$$

The final inequality follows by applying Jensen's inequality to the convex function  $p \mapsto x^p$ .

Dividing by  $\text{Mod}_p(\Gamma)$  yields the desired result,

$$\frac{\|\rho^* - \rho\|_p^p}{\|\rho^*\|_p^p} \leq 2^{p-1} \epsilon_{\text{tol}}.$$

It remains to show the case when  $1 < p < 2$ . In this case, the Clarkson inequality reads:

$$\|\rho + \rho^*\|_p^q + \|\rho - \rho^*\|_p^q \leq 2 \left( \|\rho\|_p^p + \|\rho^*\|_p^p \right)^{\frac{q}{p}}. \quad (1.2.18)$$

Starting with the same procedure as in the  $p \geq 2$  case, we substitute (1.2.16) and (1.2.17)

into (1.2.18) to achieve,

$$\begin{aligned}
\|\rho - \rho^*\|_p^q &\leq 2 \left( \|\rho\|_p^p + \|\rho^*\|_p^p \right)^{\frac{q}{p}} - \|\rho + \rho^*\|_p^q \\
&\leq 2 \left( \|\rho\|_p^p + \|\rho^*\|_p^p \right)^{\frac{q}{p}} - \left( 1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}} \right)^q \|\rho^*\|_p^q \\
&\leq 2^{1+\frac{q}{p}} \left( \|\rho^*\|_p^p \right)^{\frac{q}{p}} - \left( 1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}} \right)^q \|\rho^*\|_p^q \\
&= \left[ 2^{1+\frac{q}{p}} - \left( 1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}} \right)^q \right] \|\rho^*\|_p^q \\
&= 2^q \left[ 1 - \left( \frac{1 + (1 - \epsilon_{\text{tol}})^{\frac{1}{p}}}{2} \right)^q \right] \|\rho^*\|_p^q \\
&\leq 2^q \left[ 1 - \left( \frac{1 + (1 - \epsilon_{\text{tol}})^{\frac{q}{p}}}{2} \right) \right] \|\rho^*\|_p^q,
\end{aligned}$$

where the penultimate equality used  $\frac{1}{q} + \frac{1}{p} = 1$  and the final inequality once again used discrete Jensen's inequality. Moreover, picking up in the last chain of inequalities/equalities, computing the common denominator inside the square brackets yields,

$$\begin{aligned}
\|\rho - \rho^*\|_p^q &\leq 2^{q-1} \left[ 1 - (1 - \epsilon_{\text{tol}})^{\frac{q}{p}} \right] \|\rho^*\|_p^q \\
&\leq 2^{q-1} \left[ 1 - \left( 1 - \frac{q}{p} \epsilon_{\text{tol}} \right) \right] \|\rho^*\|_p^q \\
&= \frac{2^{\frac{q}{p}} q}{p} \epsilon_{\text{tol}} \|\rho^*\|_p^q = 2^{\frac{q}{p}} \frac{q \epsilon_{\text{tol}}}{p} \|\rho^*\|_p^q,
\end{aligned}$$

where the second inequality follows from the tangent line approximation to  $x \rightarrow (1 - x)^{\frac{q}{p}}$  at the point  $x = 0$ . (Since  $\frac{q}{p} > 1$ , the tangent line is a lower bound, and consequently, subtracting a smaller number, makes the right-hand side larger.) This final bound is equivalent to (1.2.14) when  $1 < p < 2$ .  $\square$

While it is great to know that the modulus and extremal density can be approximated arbitrarily well in finite time, there is much more our understanding can gain by viewing

the modulus as a convex program. For some of the much broader theory and intuition, see<sup>4</sup>.

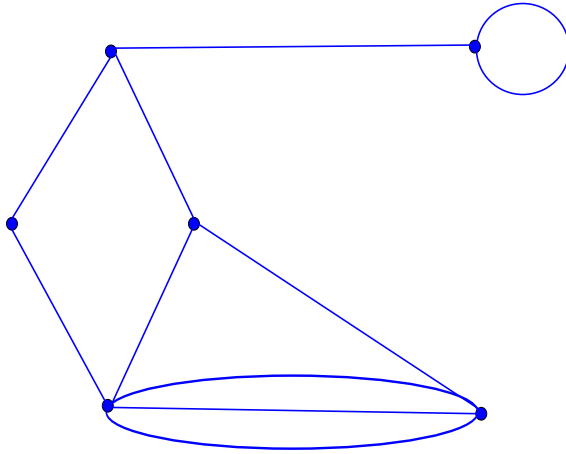
# Chapter 2

## Transboundary Modulus

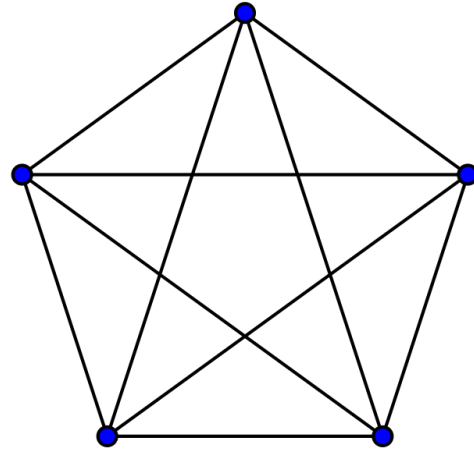
### 2.1 More graph theory

A **multigraph** is an ordered pair  $G = (V, E)$  where  $V$  is still a vertex-set, but  $E$  is an edge multiset. A **multiset** is a generalization of the idea of a set, where multiple instances of an element can occur. In terms of sets,  $\{a\} = \{a, a\}$ , however as a multiset  $\{a\} \subsetneq \{a, a\}$ . In order to be able to identify different edges in a multigraph, we consider edges in a multigraph as ordered triples  $\{v_1, v_2, j\} \in E$  for  $v_1, v_2 \in V$  and  $j \in \mathbb{N}$ . Multigraphs can be directed or undirected. An example of a directed multigraph would be the network of flights between cities in the United States, where nodes are airports, and edges represent specific flights between two airports.

A graph is called **planar** or a **planar map** if it can be embedded into the plane in such a way that no edges cross (see Figure 2.1a). Kuratowski's theorem, stated below and proved in<sup>10</sup>, is a sufficient and necessary condition to check if a graph has a planar embedding based on isomorphisms of smoothed subgraphs. A vertex of degree two, say  $w$ , can be **smoothed**. That is, if  $e_1 = \{w, v\}$  and  $e_2 = \{u, w\}$ , we can remove the vertex  $w$  and replace it with the edge  $e = \{u, v\}$ . While Kuratowski's theorem is stated for graphs (and not multigraphs), it is easy to see that the only limitation to planar embeddings of a multigraph comes from its



(a) A planar embedding of a multigraph with self-loops.



(b) The complete graph on 5 vertices is a non-planar map.

underlying graph (when you include the same edges, but all with multiplicity one). This can be seen by drawing the edges of higher multiplicity thinner and closer together as needed.

**Theorem 2.1.1. (Kuratowski).** *A graph  $G$  is planar if and only if it has no subgraph  $G'$  that can be smoothed into either  $K_5$  or  $Bipartite(3, 3)$ .*

In a graph, a **quadrangle** is a collection of vertices and edges, so that starting at any one of the (up to) 4 vertices, you can follow the edges in a fixed order and end up where you started in exactly 4 hops. In this sense, every quadrilateral is a quadrangle. However, some quadrangles can be quite surprising (see Figure 2.2).

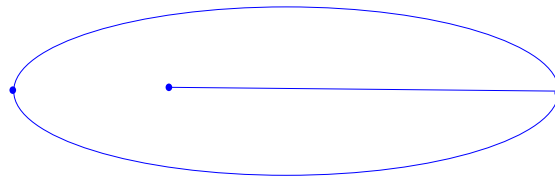


Figure 2.2: Despite only having three vertices and three sides, this is an example of quadrangle in a graph.

A **tree** is a graph  $G = (V, E)$  that has no closed loops. It is immediate to see that no strict, undirected multigraph is a tree, as any edge with multiplicity greater than one, immediately creates a loop.



## 2.2 Transboundary modulus

Random-cluster models are important in graph theory as well as many applications. We will briefly discuss one such application to Quantum Liouville Gravity (QLG) surfaces, particularly an approach of Scott Sheffield, Jason Miller, and Bertrand Duplantier<sup>11</sup>. One key property of a QLG surface is conformal invariance. Hence, we introduce a discrete analog of Oded Schramm's conformally invariant transboundary modulus as a proposed tool to better understand the random planar maps that arise in this particular random-cluster model.

### 2.2.1 Preliminaries

For a graph  $G = (V, E)$  we say that a subset of edges is **connected** if the minimal subgraph induced by the edge-set is connected. The transboundary modulus is defined on an ordered pair  $(G, C)$  where  $G = (V, E)$  is a (multi)graph and  $C = \{C_1, \dots, C_m\}$  is a set of clusters. A **cluster** is connected subset of  $E$ , i.e.,  $C_j = \{e_1^j, e_2^j, \dots, e_{m_j}^j\}$ . It is further necessary that the clusters  $C_j$  are pairwise-disjoint subsets of  $E$ . The set  $\partial C_j \subset V$  is called the boundary of  $C_j$  and is defined as the set of vertices that are an endpoint to edges in  $C_j$  as well as endpoints to edges not in  $C_j$ . That is,

$$\partial C_j := \{v \in V : \text{there exists some } (v, u) \in C_j \text{ and some } (v, w) \in E \setminus C_j\}.$$

Given such a multigraph  $G = (V, E)$  and clusters  $\{C_j\}$  we define the set

$$E_0 = E \setminus \bigcup C_j.$$

Then a walk on  $(G, C)$  is a string of vertices and edges,

$$\gamma = v_0 e_1 v_1 \cdots e_n v_n.$$

The walk  $\gamma$  is said to **enter the cluster**  $C_j$  whenever  $e_k \in C_j$  but  $e_{k-1} \notin C_j$ . For a family of walks  $\Gamma$  we can define the function  $\mathcal{N}_C : E_0 \cup C \times \Gamma \rightarrow \mathbb{N}_0$  for each  $\gamma \in \Gamma$  by

$$\mathcal{N}_C(\omega, \gamma) = \begin{cases} \text{The number of times } \gamma \text{ crosses edge } e & \text{if } \omega = e \in E_0, \\ 1 & \text{if } \omega = C_j \in C \text{ and } E(\gamma) \cap C_j \neq \emptyset \\ 0 & \text{if } \omega = C_j \in C \text{ and } E(\gamma) \cap C_j = \emptyset. \end{cases}$$

A  $\rho$ -density on  $(G, C)$  is a function  $\rho : E_0 \cup C \rightarrow [0, \infty)$ . The  $\rho$ -length of a walk on  $(G, C)$  is defined as

$$\Delta_\rho(\gamma) = \sum_{e \in E_0} \mathcal{N}(e, \gamma) \rho(e) + \sum_{C_j \in C} \mathcal{N}(C_j, \gamma) \rho(C_j).$$

A  $\rho$ -density on  $(G, C)$  is said to be **admissible** for a family of walks  $\Gamma$  if  $\Delta_\rho(\gamma) \geq 1$  for every  $\gamma \in \Gamma$ . The set of all admissible densities for  $\Gamma$  on  $(G, C)$  is denoted  $\text{Adm}(\Gamma; (G, C))$ . Moreover, the  $(p, q)$ -**energy of a density  $\rho$  on  $(G, C)$**  is defined as

$$\mathcal{E}_{p,q}(\rho) = \sum_{e \in E_0} \rho(e)^p + \sum_{C_j \in C} \rho(C_j)^q.$$

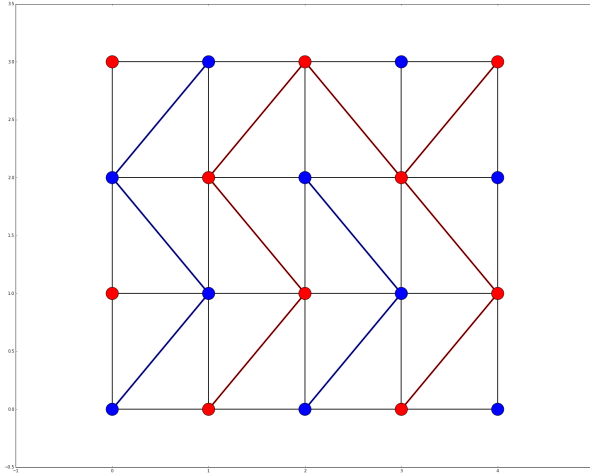
Finally, the  $(p, q)$ -**transboundary modulus of  $\Gamma$  on  $(G, C)$**  is defined as

$$\text{Mod}_{p,q}(\Gamma, C) = \inf \left\{ \mathcal{E}_{p,q}(\rho) \mid \rho \in \text{Adm}(\Gamma; (G, C)) \right\}.$$

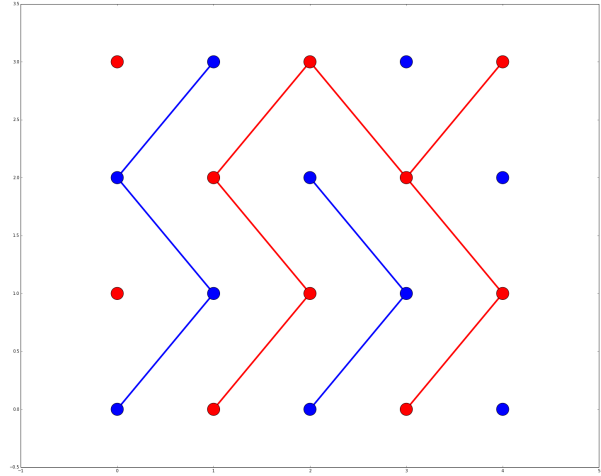
## 2.2.2 Virtual Graphs

The definition of the transboundary modulus for a graph  $G = (V, E)$  over the set of clusters  $C$  is a natural way of thinking about the transboundary modulus as an adaptation from the continuum case. However, as with the  $p$ -modulus, the discrete setting allows for a simpler approach to the transboundary modulus. This approach requires creating a so called **virtual graph**  $\tilde{G} = (\tilde{V}, \tilde{E})$  induced by  $(G, C)$ .

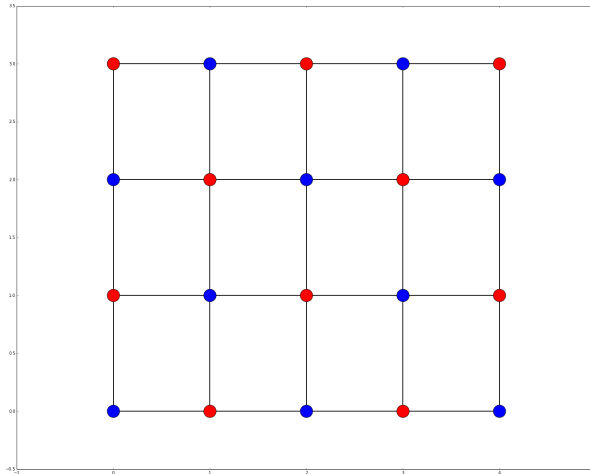
To create the virtual graph  $\tilde{G} = (\tilde{V}, \tilde{E})$  we begin with the graph  $G(E_0)$ , that is, the



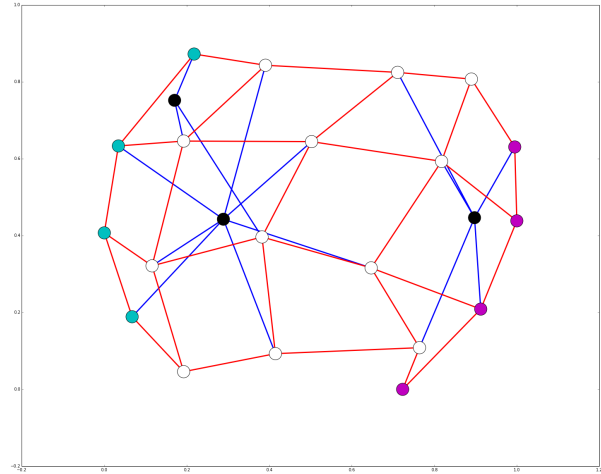
The  $G = (V, E)$  with clusters  $C$  in blue and red.



Only the edges in  $C$  are shown.



Only the edges in  $E_0$  are shown.



The virtual graph  $\tilde{G}$ .

Figure 2.3: From  $(G, C)$  to  $\tilde{G}$ .

subgraph of  $G$  induced by the edges in  $E_0$ . We define the set  $V_v = \{n_j\}_{j=1}^m$  where  $n_j$  is the virtual node corresponding to  $C_j$ . Then,  $\tilde{V} := V \cup \{V_v\}$ . Moreover, for  $j = 1, 2, \dots, m$ , create the edge-sets  $E_j = \{e_1^j, e_2^j, \dots, e_{j_n}^j\}$ , which contain the virtual-edges connecting  $n_j$  to each node in  $\partial C_j$ . Then  $\tilde{E} = \bigcup_{j=0}^m E_j$ . In this manner, you can always create the virtual graph  $\tilde{G}$  corresponding to  $(G, C)$ . See Figure 2.3.

A  $\rho$ -density on the virtual graph  $\tilde{G} = (\tilde{V}, \tilde{E})$  is a map  $\rho : \tilde{E} \rightarrow [0, \infty)$ . A walk on the

virtual graph  $\tilde{G}$  is a string of alternating nodes and edges,

$$W = v_0 e_1 v_1 \cdots e_n v_n,$$

where  $\{e_k\} \subset \tilde{E}$  and  $\{v_k\} \subset \tilde{V}$ . To define the  $\rho$ -length on a virtual graph  $\tilde{G}$ , we first define the function  $\eta : E_0 \cup V_v \times \Gamma \rightarrow \mathbb{N}_0$  for each  $\gamma \in \Gamma$  by

$$\eta(\omega, \gamma) := \begin{cases} \text{the number of times } \gamma \text{ traverses } \omega & \text{if } \omega \in E_0 \\ 1 & \text{if } \omega \in \tilde{V} \cap V(\gamma) \\ 0 & \text{if } \omega \in \tilde{V} \setminus V(\gamma). \end{cases}$$

Then, the  $\rho$ -length of a virtual walk  $\gamma$  is defined as

$$L(\gamma) = \sum_{e \in E_0} \eta(e, \gamma) \rho(e) + \sum_{j=1}^m \sum_{e \in E_j} \frac{\rho(e)}{|E_j|} \eta(n_j, \gamma). \quad (2.2.1)$$

A density is called **admissible for  $\tilde{\Gamma}$  on  $\tilde{G}$**  if it satisfies

$$L(\gamma) \geq 1 \quad \forall \gamma \in \tilde{\Gamma} \quad \text{and} \quad \rho(e_k^j) = \rho(e_\ell^j) \quad \forall e_\ell^j, e_k^j \in E_j \quad \forall j = 1, 2, \dots, m.$$

The reason it is natural to require that all edges in a set  $E_j$  have the same weight is because this ensures that when visiting the virtual node  $n_j$ , has the same  $\rho$ -cost regardless of where you came from or where you left. In particular, the cost to visit node  $n_j$  equals  $\rho(e_1^j)$ , for all  $E_j$ . Hence, the  $(p, q)$ -energy on a virtual graph is defined as

$$\mathcal{E}_{p,q}(\rho) = \sum_{e \in E_0} \rho(e)^p + \sum_{j=1}^m \rho(e_1^j)^q. \quad (2.2.2)$$

Moreover, for a family of walks  $\tilde{\Gamma}$  on  $\tilde{G}$ ,

$$\text{Mod}_{p,q}(\tilde{\Gamma}; \tilde{G}) = \inf \{ \mathcal{E}_{p,q}(\rho) : \rho \in \text{Adm}(\tilde{\Gamma}, \tilde{G}) \}.$$

We now describe a mapping  $\sigma$  that takes walks on  $(G, C)$  into walks on the virtual graph  $\tilde{G}$  in a canonical way. For a walk  $\gamma$  on  $(G, C)$ , let  $\tilde{\gamma}$  be the walk satisfying:

- i)  $\mathcal{N}(e, \gamma) = \eta(e, \tilde{\gamma})$  for all  $e \in E_0$
- ii)  $\mathcal{N}(C_j, \gamma) = \eta(n_j, \tilde{\gamma})$  for all  $j = 1, 2, \dots, m$ ,
- iii) the order in which  $\gamma$  visits vertices while not in a cluster  $C_j$  is the same as the order  $\tilde{\gamma}$  visits vertices of  $\tilde{G}$ , and
- iv) the order in which  $\gamma$  enters clusters  $C_j$  is the same as the order  $\tilde{\gamma}$  visits the corresponding virtual nodes  $n_j$ .

**Remark 2.2.1.** While the mapping  $\sigma$  is not one-to-one because it ignores the number of hops  $\gamma$  takes while within a cluster  $C_j$ , it is well-defined since the (i)-(iv) define the vertices, edges, and their orders that the walk  $\tilde{\gamma}$  can take on  $\tilde{G}$ .

Furthermore, for any density  $\rho$  on  $(G, C)$  there is a canonical transformation  $\sigma'$  to a density  $\tilde{\rho}$  on  $\tilde{G}$  defined by

$$\tilde{\rho}(e) := \rho(e) \quad \forall e \in E_0 \quad \text{and} \quad \tilde{\rho}(e_k^j) := \rho(C_j) \quad \text{for each } j = 1, 2, \dots, m, \text{ and all } 1 \leq k \leq j_n.$$

**Proposition 2.2.1.** *If  $\sigma, \sigma'$  are as described, then*

$$\Delta_\rho(\gamma) = L_{\sigma'(\rho)}(\sigma(\gamma)) \quad \text{and} \quad \mathcal{E}_{p,q}(\rho) = \mathcal{E}_{p,q}(\sigma'(\rho)).$$

*Proof.* Starting from the definition of  $\Delta_\rho$  we attain,

$$\begin{aligned} \Delta_\rho(\gamma) &= \sum_{e \in E_0} \mathcal{N}(e, \gamma) \rho(e) + \sum_{C_j \in C} \mathcal{N}(C_j, \gamma) \rho(C_j) \\ &= \sum_{e \in E_0} \eta(e, \sigma(\gamma)) \rho(e) + \sum_{j=1}^m \eta(n_j, \sigma(\gamma)) \rho(C_j) \\ &= \sum_{e \in E_0} \eta(e, \sigma(\gamma)) \sigma'(\rho(e)) + \sum_{j=1}^m \eta(n_j, \sigma(\gamma)) \sigma'(\rho(e_1^j)). \end{aligned}$$

The second and third equalities used the definition of  $\sigma$  and  $\sigma'$  respectively. Moreover,  $\sigma'(\rho(e_k^j)) = \sigma'(\rho(e_\ell^j))$  for all  $1 \leq j \leq m$  and  $1 \leq k, \ell \leq n_j$ , that is  $\sigma'(\rho(\cdot))$  is constant on  $E_j$ . Consequently,

$$\eta(n_j, \sigma(\gamma))\sigma'(\rho(e_1^j)) = \sum_{e \in E_j} \frac{\sigma'(\rho(e))}{|E_j|} \eta(n_j, \sigma(\gamma))$$

As a result,

$$\sum_{j=1}^m \eta(n_j, \gamma)\sigma'(\rho(e_j^1)) = \sum_{j=1}^m \sum_{e \in E_j} \frac{\sigma'(\rho(e))}{|E_j|} \eta(n_j, \sigma(\gamma))$$

Hence,

$$\Delta_\rho(\gamma) = \sum_{e \in E_0} \eta(e, \sigma(\gamma))\sigma'(\rho(e)) + \sum_{j=1}^m \sum_{e \in E_j} \frac{\sigma'(\rho(e))}{|E_j|} \eta(e, \sigma(\gamma)) = L_{\sigma'(\rho)}(\sigma(\gamma)).$$

Now, using the definition of the  $(G, C)$  energy of a density, the definition of  $\sigma'$ , and then the definition of the energy of a density of a virtual graph, we attain

$$\mathcal{E}_{p,q}(\rho) = \sum_{e \in E_0} \rho(e)^p + \sum_{C_j \in C} \rho(C_j)^q = \sum_{e \in E_0} \sigma'(\rho(e))^p + \sum_{j=1}^m \sigma'(\rho(e_1^j))^q = \mathcal{E}_{p,q}(\sigma'(\rho)).$$

□

An immediate consequence is the following corollary.

**Corollary 2.2.2.** *If  $\Gamma$  is a family of walks on  $(G, C)$ , then*

$$\sigma'(\text{Adm}(\Gamma; (G, C))) = \text{Adm}(\sigma(\Gamma); \tilde{G}) \quad \text{and} \quad \text{Mod}_{p,q}(\Gamma; (G, C)) = \text{Mod}_{p,q}(\sigma(\Gamma); \tilde{G}).$$

This result guarantees the validity of using virtual networks to compute the transboundary modulus. The advantages of a virtual network approach to the transboundary modulus are numerous. The greatest benefit is that by compressing each cluster down to a single node, the function `shortest()` as used in Algorithm 1 is much easier to write and much

quicker computationally in the case of a virtual graph as opposed to a graph with clusters. Moreover, the algorithm defined in Algorithm 1 is guaranteed to terminate in finite time while computing the transboundary modulus on a virtual node. While the style of proof that achieved the error bounds in Theorem 1.2.5 can be followed to find similar bounds for virtual networks, the bounds themselves are no longer the same.

## 2.3 Quadrangulations of the sphere

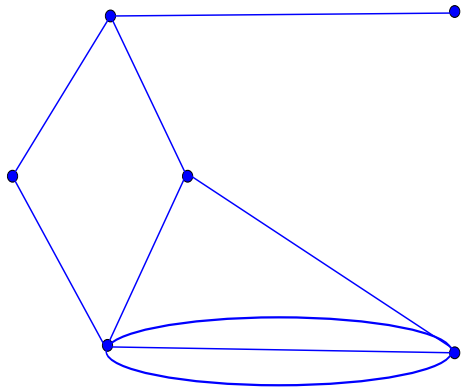
### 2.3.1 Sheffield’s Bijection

There is a well-known and canonical bijection between planar maps and quadrangulations. If you begin with a planar map  $M = (V, E)$ , we will call the graph  $M$  the primal map, and we call its vertices and edges are called **primal-vertices** and **primal-edges**. To create the quadrangulation of  $M$  denoted  $Q(M)$ , first add a **dual-vertex** to each face of the planar map. As we think of planar maps as being on the surface of a sphere, the “outside” face also should have a dual-vertex added. Then, for each dual-vertex, create one dual-edge connecting said vertex to each primal-vertex in the polygon surrounding its face. Removing your primal-edges, you have created the quadrangulation  $Q(M)$  (see Figure 2.4).

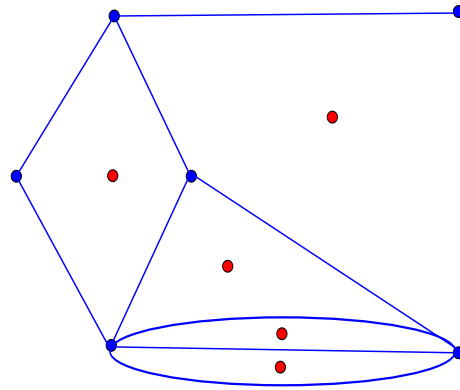
In<sup>12</sup>, Scott Sheffield introduced a bijection between rooted planar maps decorated with a conformal loop ensemble (CLE) and words on a certain semi-group  $\Theta$ . This bijection is closely related to the Schaeffer bijection<sup>13</sup> between trees and rooted planar maps.

The Sheffield bijection actually creates a quadrangulation. As it happens, the “outside” of the map, is always a quadrangle, so it is natural to imagine these quadrangulations as being embedded on a sphere. Sheffield’s bijection can be used to show that these planar maps with CLEs converge to Liouville Quantum Gravity (LQG) with a CLE<sup>11</sup> in the limiting case where the “size” of the faces of the quadrangles converge to zero.

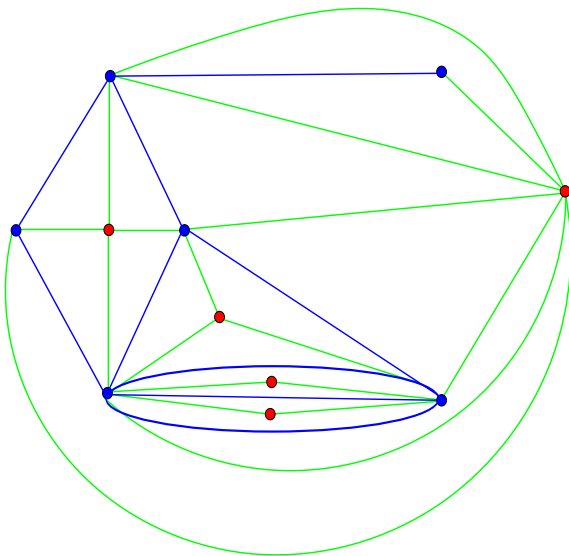
A mnemonic to remember the semi-group  $\Theta$  is to think of a peculiar burger shop that makes hamburgers,  $\textcircled{\text{H}}$ , and cheeseburgers,  $\textcircled{\text{C}}$ . Customers can order hamburgers  $\boxed{\text{H}}$ , cheese-



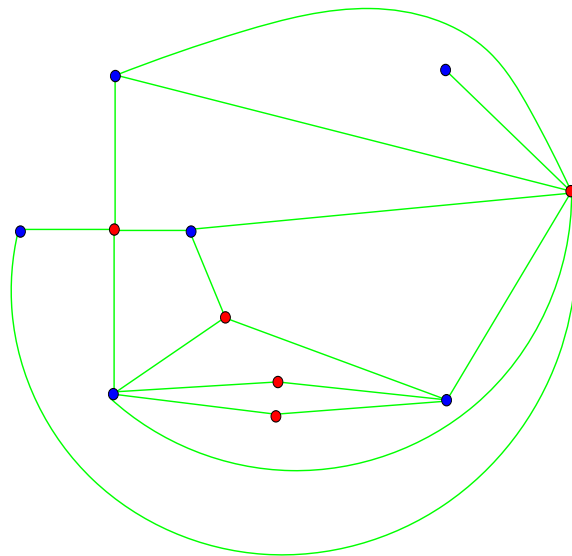
(a) The primal map  $M$



(b) Dual vertices are added in red.



(c) Quadrilateral edges are added in green.



(d) The quadrangulation  $Q(M)$ .

Figure 2.4: The canonical bijection from a planar map to a quadrangulation in four steps.



burgers  $\boxed{C}$ , or the most recent (freshest) burger made  $\boxed{F}$ . The owner of the shop has a Last-In-First-Out (LIFO) business model, and a single burger chute where all burgers made are placed. If a customer orders a cheeseburger, they search through the chute until the first cheeseburger they find and eat it to fulfill their order. However, if there is no cheeseburger in the chute, the customer leaves (angrily). There is nothing special about cheeseburgers, i.e., the shop follows the same procedure when a customer orders a hamburger. When a customer orders a “freshest” burger,  $\boxed{F}$ , they take whatever burger is on the top of the chute, eat it, and their order is fulfilled.

This story is formalized through the associative semi-group  $\Theta$  with elements  $\{\textcircled{C}, \textcircled{H}, \boxed{C}, \boxed{H}, \boxed{F}\}$ . A string of symbols in  $\Theta$  is called a **word** and is read from left to right. The symbols in a word can be “reduced” in the following two ways:

$$\textcircled{C}\boxed{H} = \boxed{H}\textcircled{C} \quad \text{and} \quad \textcircled{H}\boxed{C} = \boxed{C}\textcircled{H} \quad (2.3.1)$$

$$\textcircled{C}\boxed{C} = \emptyset = \textcircled{C}\boxed{F} \quad \text{and} \quad \textcircled{H}\boxed{H} = \emptyset = \textcircled{H}\boxed{F}. \quad (2.3.2)$$

The equations (2.3.1) and (2.3.2) are called the commutivity and order-fulfillment relations respectively.

Recalling the mnemonic, a word can be thought of keeping track of a burger shop’s activity for a day. For example, the word  $W = \boxed{C}\textcircled{C}\textcircled{H}\textcircled{C}\boxed{H}\boxed{C}$ , would mean that the shop had a customer order a cheeseburger and leave emptyhanded, then the shop made a cheeseburger, a hamburger, and another cheeseburger before two customers ordered a hamburger and then a cheeseburger. This means, that at the end of the day, keeping track of unfulfilled orders and unordered burgers the shop left one cheeseburger order unfulfilled and had one uneaten cheeseburger. Using first the commutivity relation then the order fulfillment relations, we discover

$$W = \boxed{C}\textcircled{C}\textcircled{H}\textcircled{C}\boxed{H}\boxed{C} \sim \boxed{C}\textcircled{C}\textcircled{H}\textcircled{H}\textcircled{C}\boxed{C} \sim \boxed{C}\textcircled{C} =: \overline{W}.$$

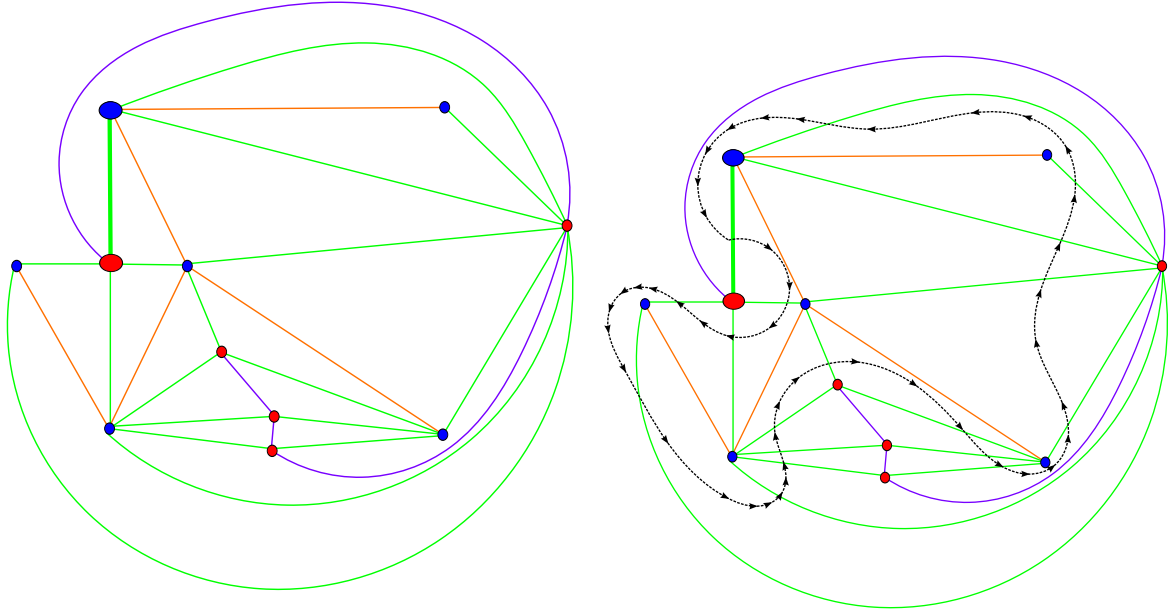
The  $\sim$  symbol denotes equivalence under the relations (2.3.1) and (2.3.2). We will continue to use a bar to denote that a word is in reduced form. The fact that this is a convention should be a hint to the fact that the reduced form of a word always exists and is unique, if we rigorously define what is meant by “reduced form”. A word  $W$  is said to be in **reduced form** if there are no burger orders following the production of any burger. In particular, a typical reduced word is a string of burger orders followed by a string of burgers. However, it is possible that there are no unfulfilled orders and/or no un-ordered burgers. A proof of this claim is made in <sup>12</sup> and is based off of induction on the length of the word.

The Sheffield bijection is based upon the canonical quadrangulation achieved in Figure 2.4 and described in the preceding section. The key to the Sheffield bijection is that in Figure 2.4c, the primal and quadrilateral edges actually create a triangulation, with each quadrilateral face containing a primal diagonal. The Sheffield bijection allows us to choose either primal or dual diagonal edges. A dual edge, is an edge connecting two dual vertices. It is possible to replace each primal diagonal with the corresponding dual diagonal. In this manner a word  $W$  chooses a set of diagonals and a particular quadrangulation of the sphere, hence creating a triangulation.

### 2.3.2 Maps and words with one loop

We use an example to describe how the Sheffield bijection works. Herein, we only describe the simple case where the diagonals are chosen to create a primal tree and a dual tree.

In Figure 2.5a, we show such a choice of diagonal edges, along with a **root** of the map, i.e., a special edge that tells us where to begin. To create the word  $W$  corresponding to the map  $M$ , we want to trace a path that crosses each quadrilateral edge once and never crosses a diagonal edge. We let primal-diagonal edges correspond to cheeseburgers and dual-diagonal edges correspond to hamburgers in the sense that the *first* time that this path enters a quadrilateral with a primal/dual edge, a cheeseburger/hamburger (respectively) is *produced*. The second time that this path enters a quadrilateral with a primal/dual edge, a



(a) The orange (primal) diagonals correspond to cheeseburger orders/consumption and the purple (dual) diagonals correspond to hamburger orders/consumption. The bold edge and larger vertices denote the root of the map. (b) The unique path that crosses each quadrilateral edge once and never crosses a diagonal edge.

cheeseburger/hamburger (respectively) is *ordered*. In Figure 2.5b, the unique loop is drawn, and we can trace out the corresponding word to be:

$$W = \textcircled{C}\textcircled{C}\textcircled{C}\textcircled{H}\textcircled{C}\textcircled{H}\textcircled{H}\textcircled{H}\textcircled{C}\textcircled{C}\textcircled{H}\textcircled{H}\textcircled{H}\textcircled{C}\textcircled{C}\textcircled{C}\textcircled{C}\textcircled{H}. \quad (2.3.3)$$

Upon inspection, it is straightforward using the reduction relations to show that  $\overline{W} = \emptyset$ . This is no coincidence. Every representation of the empty word creates a rooted mating of two trees, and every rooted mating of two trees creates a unique representation of the empty word.

It is more difficult to describe how to create a planar map from a word  $W$ . Figure 2.6 shows the first few steps in re-creating the planar map in Figure 2.5a from its corresponding word. Here, I describe the process. When creating the graph, we will follow the procedure that the vertices  $p_k/d_k$  are the  $k$ th primal/dual vertices which for  $k > 0$  coincides with the

$k$ th cheeseburger/hamburger respectively.<sup>1</sup> For  $k \geq 1$ , the edges  $q_k/c_k/h_k$  denote the  $k$ th quadrilateral/primal/dual edges respectively. The  $k$ th primal/dual edges correspond to the  $k$ th cheeseburger/hamburger respectively. Meanwhile, typically<sup>2</sup> the  $k$ th quadrilateral edge corresponds to the  $k$ th letter in the word, the  $k$ th primal/dual edges correspond to the production or consumption of a cheeseburger/hamburger respectively. For clarity in creating the map, we will also keep track of our “current inventory”, where we have an evolving cheeseburger/hamburger queue  $P_t/D_t$  that contains the primal vertices corresponding to the cheeseburgers/hamburgers produced but not ordered at time  $t$ .

To create the the map  $M$ , begin with the map that only includes our root. So, we let  $M_0 = (\{p_0, d_0\}, \{q_0\})$  where edge  $q_0$  connects  $p_0$  to  $d_0$ . At this point our cheeseburger/hamburger queues are  $P_0 = [p_0]$  and  $D_0 = [d_0]$ . Since the first symbol in the word from (2.3.3) is a  $\textcircled{C}$ , we define  $M_1$  by adding a vertex  $p_1$  to  $M_0$ . We connect  $p_1$  to  $p_0$  (the most recently produced, unconsumed cheeseburger) with primal edge  $c_1$  and connect  $p_1$  to  $d_0$  (the most recently produced, unconsumed hamburger) with edge  $q_1$ . We then have  $P_1 = [p_0, p_1]$  and  $D_1 = [d_0]$ . Since the next symbol in  $W$  is another  $\textcircled{C}$ ,  $M_2$  is created by adding a new vertex  $p_2$  connected to  $p_1$  and  $d_0$  with edges  $c_2$  and  $q_2$  respectively. Now,  $P_2 = [p_0, p_1, p_2]$  and  $D_2 = [d_0]$ . The third symbol in  $W$  is again  $\textcircled{C}$  so that  $M_3$  is created by adding vertex  $p_3$  connecting it to vertices  $p_2$  and  $d_0$  with edges  $c_3$  and  $q_3$  respectively. Now,  $P_3 = [p_0, p_1, p_2, p_3]$  and  $D_3 = [d_0]$ . Since the 4th symbol in  $W$  is  $\textcircled{H}$ ,  $M_4$  is created by connecting vertex  $d_1$  to nodes  $p_3$  and  $d_0$  through edges  $q_4$  and  $h_1$  respectively. Now,  $P_4 = [p_0, p_1, p_2, p_3]$  and  $D_4 = [d_0, d_1]$ . The 5th symbol in  $W$  is  $\boxed{C}$ , so we consume the most recently produced cheeseburger. Looking at  $P_4$ , we see that the most recently produced (unordered) cheeseburger corresponds to node  $p_3$ , so  $M_5$  is created by adding a quadrilateral edge  $q_5$  between nodes  $p_2$  (the most recently produced, uneaten cheeseburger preceding  $p_3$ ) and nodes  $d_1$  (the most recently produced, uneaten hamburger). This process, proceeds

---

<sup>1</sup>The endpoints of the rooted edge,  $(p_0, d_0)$  do not correspond to hamburgers and cheeseburgers in the word  $W$ , as they do not have a primal or dual edge connected to them as they are created.

<sup>2</sup>It is a little inconvenient that the the 0th quadrilateral edge,  $q_0$ , corresponds to the final letter in the word, replacing the fact that no action is taken when the last letter is observed.

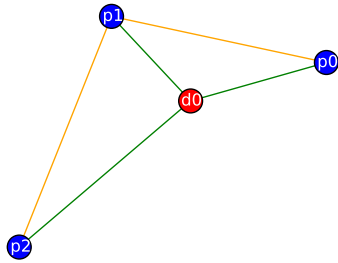
similarly until the final symbol in the word. After treating the penultimate symbol in the described manner, we do nothing for the final symbol since the rooted edge  $q_0$  corresponds to the action that would otherwise be determined by the final symbol.

A Python algorithm that creates a multigraph from a given representation of the null-word, labels, colors, and identifies both the nodes and edges in the described method above can be found at <https://www.math.ksu.edu/~maxximus/Sbijection>. Despite only discussing the special case of the bijection where the word reduces to the empty-word and uses no symbols  $\boxed{F}$ , the algorithm can handle a “freshest” order. To handle the case where your word does use the symbol  $\boxed{F}$ , one analyzes the word to see whether  $\boxed{F}$  consumes a  $\textcircled{C}$  or  $\textcircled{H}$  and then treat  $\boxed{F}$  as a  $\boxed{C}$  or  $\boxed{H}$  respectively. This is the only other (still relatively simple) case discussed in greater detail in see<sup>12</sup>. In particular, words corresponding to customers who leave with an unfilled order are not discussed.

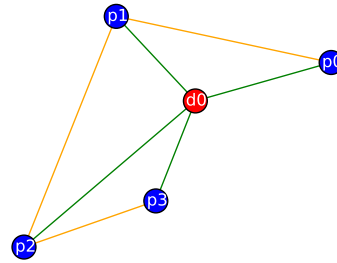
### 2.3.3 Maps and words with more than one loop

In the previous section, we looked at the special case where the planar map  $M$  was always the mating of two trees. This case corresponded nicely with the words  $W$  in  $\Theta$  that reduce to the empty word. Here we see this is one of many relationships between the global geometric properties of a planar map with conformal loop ensembles and the local quantitative observations on the corresponding words. The full results are summarized in Table 2.1.

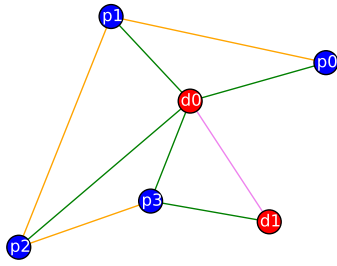
Given this intimate relationship between the symbols that make-up a word  $W$  and the geometry of the corresponding planar map  $M$ , we would expect to find a strong relationship between the length of words and the transboundary modulus, of the corresponding map and a relationship between the ratio of  $\boxed{C} : \boxed{H} : \boxed{F}$ . Some preliminary numerical results regarding this first hypothesis are included below.



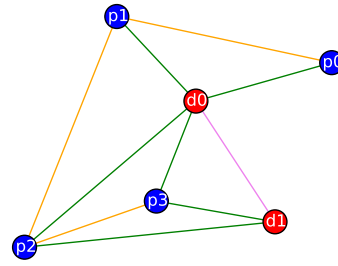
The graph  $M_2$ .



The graph  $M_3$ .



The graph  $M_4$ .

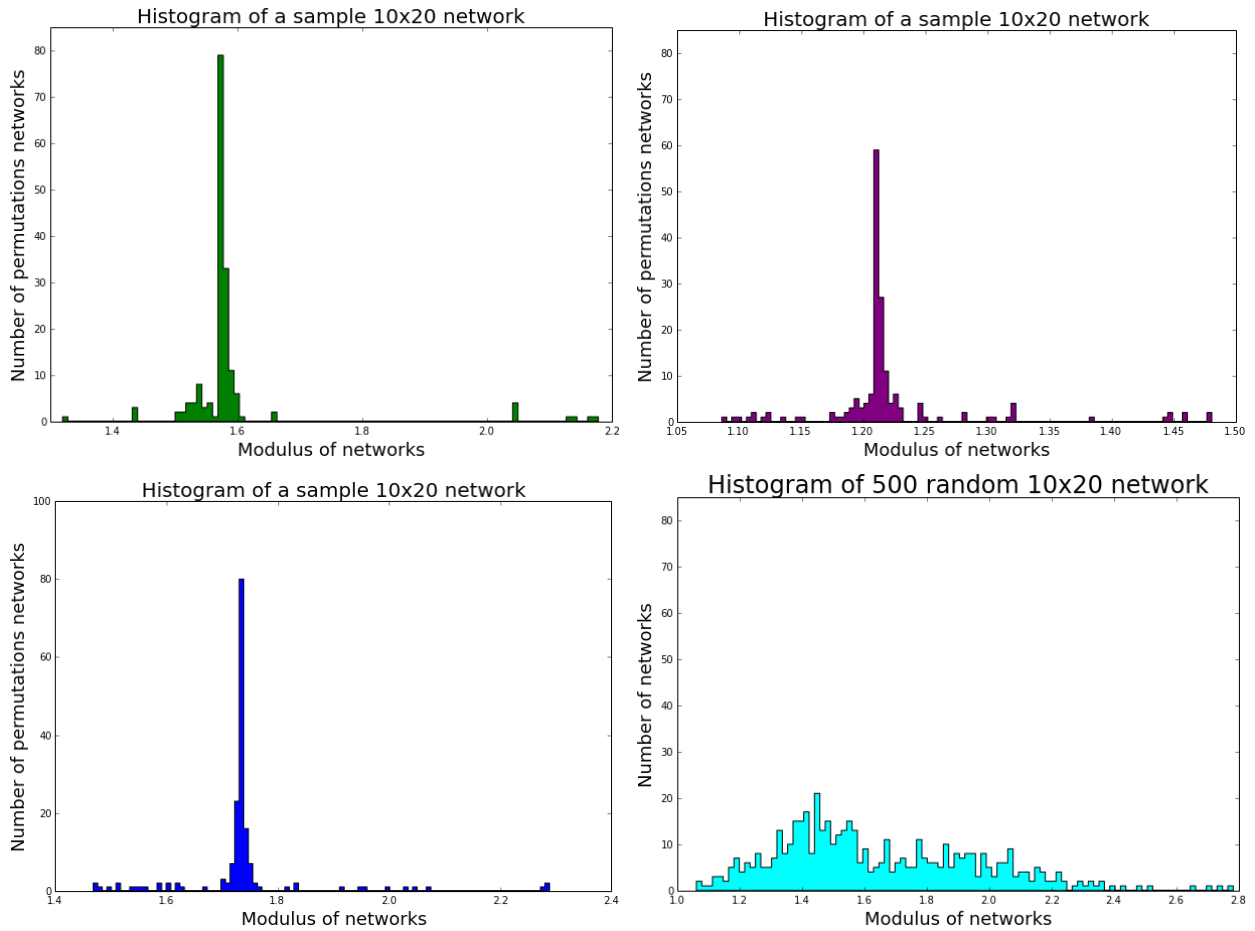


The graph  $M_5$ .

Figure 2.6: Steps two through five of creating the planar map  $M$  from the word in (2.3.3). It's worth noting the creation of the planar map is only unique up to graph isomorphisms. Also, it is particularly hard to draw the map  $M$  with a planar embedding that uses only straight edges. The graph layout shown would either lose the pattern of straight edges, or lose its planar embedding within the next few steps.

ROOTED MAP/EDGE SET: $(M, T)$	CORRESPONDING WORD: $W$
Number of edges in $\tilde{T}$	Number of $\textcircled{\text{H}}$ 's in $W$ .
Number of edges in $\tilde{T}'$	Number of $\textcircled{\text{C}}$ 's in $W$ .
Original number of loops	1 plus number of $\text{F}$ 's in $W$ .
Number of components of $T$	1 plus number of $\text{F}$ 's matched to $\textcircled{\text{C}}$ 's.
Number of components of $T'$	1 plus number of $\text{F}$ 's matched to $\textcircled{\text{H}}$ 's.
Number of edges in $T$	Number of $\textcircled{\text{H}}$ to $\text{H}$ or $\textcircled{\text{C}}$ to $\text{F}$ matches.
Number of edges in $T'$	Number of $\textcircled{\text{C}}$ to $\text{C}$ or $\textcircled{\text{H}}$ to $\text{F}$ matches.

Table 2.1: Summary of quantitative local observations and their corresponding global geometric properties.  $\tilde{T}$  and  $T$  are the sets of primal diagonal edge-sets before and after (respectively) merging the conformal loop ensemble into a single loop. Similarly,  $\tilde{T}'$  and  $T'$  are the dual diagonal edge-sets before and after merging the conformal loop ensemble into a single loop.



# Chapter 3

## Graph centralities and epidemic measures

Dynamics on graphs has long been a central research topic across many applied disciplines. Several graph related quantities have proven successful in studying different applications. In particular, the effective resistance metric appears to be an important tool for studying a variety of dynamics over graphs, including, but not limited to, random walks on graphs, electrical networks networks, Markov chains, and averaging networks<sup>14</sup>. Since the effective resistance is a special case of the  $p$ -modulus, it is also interesting to study different centralities and metrics related to the modulus of families of walks on graphs.

It comes as no surprise that effective resistance is important for all these dynamic processes because they are gradient driven processes. The effective resistance is closely related to the Laplacian matrix of the underlying graph. However, epidemic spreading dynamics is a **branching process** and behaves very differently from gradient driven dynamics.

In this chapter, we seek graph quantities that help describe epidemic dynamics. Centralities are frequently used to determine properties of the underlying topology of a network. In fact, comparing different centralities on the same network can be used to classify the network structure<sup>2</sup>. Herein, we compare common centralities as well as graph metrics to



see how to best understand epidemic dynamics. We use many of the same real-world data sets as in<sup>2</sup> and conclude that surprisingly, regardless of the underlying network structure, numerics indicate that the effective resistance is the most relevant graph quantity to the epidemic spreading. A partial explanation is offered at the end of the chapter.

### 3.1 Dijkstra’s Algorithm

Many of the graph centralities and epidemic models we discuss will unsurprisingly require heavy use of knowing the shortest path between two nodes, whether the shortest path is the shortest number of hops, or the shortest path given some weighted edges. As the “number of hops” case can be recovered from setting all edge weights to one, we here discuss Dijkstra’s algorithm on a weighted graph. Dijkstra’s algorithm is a greedy algorithm for computing the shortest path length between a seed node  $s$  and every other node in  $V$  with the convention that the distance is infinite if there is no path between  $s$  and some node  $a$ .

The input to Dijkstra’s algorithm is a triple  $(G, \ell, s)$  where  $G = (V, E)$  is a graph,  $s$  is a seed node, and  $\ell : E \rightarrow [0, \infty]$  is a weighting function, so that  $\ell(a, b)$  is the weight of the edge  $\{a, b\}$ . If  $\{a, b\}$  is not an edge, then  $\ell(a, b) = \infty$ . The goal of Dijkstra’s to output:

- (1) a function  $d : V \rightarrow [0, \infty]$  so that  $d(a)$  is the weighted distance from the seed  $s$  to node  $a$  and
- (2) return a function  $\text{pred}()$ , where  $\text{pred}(s) = \emptyset$  and  $u = \text{pred}(v)$  means that a shortest path from  $s$  to  $v$  reaches  $u$  immediately before reaching  $v$ .

The function  $\text{pred}()$  allows the creation of a shortest path from  $s$  to  $v$  by reversing the chain  $v, \text{pred}(v), \text{pred} \circ \text{pred}(v), \dots, \text{pred}^{(k)}(v), \dots$  where  $\text{pred}^{(k)}(v)$  denotes iterating  $\text{pred}()$ ,  $k$  times starting at  $v$ . The iterations stop when  $\text{pred}()$  returns the empty set.

To implement Dijkstra’s algorithm, we make use of a function  $\delta : V \rightarrow [0, \infty]$  where we claim that  $\delta(v) \geq d(v)$  for all  $v \in V$  and that this inequality holds at every loop. We

also make use of a set  $W \subset V$  where we claim that  $\delta(v) = d(v)$  for all  $v$  in  $W$ . Dijkstra's algorithm can be implemented as in Algorithm 2.

It can be seen that

---

**Algorithm 2** Dijkstra's algorithm

---

**Require:**  $(G, \ell, s)$

```

1:  $W \leftarrow \emptyset$ 
2:  $\delta(s) \leftarrow 0$ 
3:  $\delta(u) \leftarrow \infty$  for all  $s \neq u \in V$ .
4: while There exists  $u \notin W$  satisfying  $\delta(u) < \infty$  do
5:   1a)  $u \leftarrow \operatorname{argmin}\{\delta(u) : u \notin W\}$ 
6:   1b)  $W \leftarrow W \cup \{u\}$ 
7:   1c)
8:   for Each  $v \notin W$  that neighbors  $u$  do
9:     if  $\delta(u) + \ell(u, v) < \delta(v)$  then
10:        $\delta(v) = \delta(u) + \ell(u, v)$ 
11:        $\operatorname{pred}(v) \leftarrow u$ .
12:     end if
13:   end for
14: end while

```

---

One can see that Dijkstra's Algorithm 2 returns the desired shortest path lengths and  $\operatorname{pred}()$  function, as a result of the following proposition.

**Proposition 3.1.1.**

- (1) At every point of the algorithm  $\delta(v) \geq d(v)$  for all  $v \in V$ .
- (2) If  $u = \operatorname{pred}(v)$  and  $\delta(u) = d(u)$  then  $\delta(v) = d(v)$ .
- (3) At the end of each iteration of the while loop,  $\delta(x) = d(x)$  for all  $x \in W$ .

We note that after proving (3), we can reinterpret (2) as: If  $u = \operatorname{pred}(v)$  and  $u \in W$  then  $d(v) = d(u) + \ell(u, v)$ . A proof of Proposition 3.1.1 can be found at <http://www.cs.yale.edu/homes/spielman/365/shortestPaths.pdf>.

In order to determine how helpful a graph quantity is in determining properties of an epidemic process, we must first choose a framework in which to work. Herein, we consider the Susceptible-Infected (SI) model of epidemiology.

## 3.2 The four fold path to the SI model

### 3.2.1 The SI model

The SI Epidemic Model is a model where every interaction between an infected and susceptible node can lead the susceptible node to become infected at a rate  $\beta$  called the infection rate. This means, that if two people, say Alice and Bob interact, and at time  $t$  Alice is infected while Bob is susceptible, then the probability that Alice infects Bob in the time interval  $(t, t + h]$  equals  $\beta h + o(h)$ . Further, in the SI epidemic model, we assume that infections occur independently and once someone becomes infected they remain infected forever. In particular, by independence, the probability that two separate infections occur during a time interval  $(t, t + h]$  is  $o(h)$ .

In realistic models, these interactions are described by links of a contact network  $G = (V, E)$ .

It's well-known that the hypothesis above amount to defining independent Poisson processes with strength  $\beta$ , one for each link in  $E$ , and then equating an infection episode to an arrival in the Poisson process. Since Poisson processes have the lack-of-memory property and can be restarted at any given deterministic time, and even at stopping times, we will be able to let these processes run independently in the background and then consider them when needed.

In order to keep track of the infection, we introduce the state vector  $\omega_t$ , where  $\omega_t(i)$  is the state at time  $t$  of the  $i$ -th node in the network. If  $\omega_t(i) = 0$  we say node  $i$  is susceptible and if  $\omega_t(i) = 1$  we say node  $i$  is infected. In order to make the scenario with Alice and Bob rigorous, we let  $N = |V|$  and  $A = [A(i, j)]_{1 \leq i, j \leq N}$  be the adjacency matrix representing network  $G$ . Then  $A(i, j) = 1$  if node  $i$  can be infected by node  $j$  and zero otherwise. Finally, let  $I_t = \{i : \omega_t(i) = 1\}$  be the set of nodes that are infected by time  $t$ . Since at most one infection occurs during  $(t, t + h]$ , we can partition using the possible infection occurrences and get:

$$\begin{aligned}
\mathbb{P} [ \omega_{t+h}(i) = 1 \mid \omega_t(i) = 0, \omega_t ] &= \sum_{\substack{j \in I_t \\ j \sim i}} (\beta h + o(h)) \\
&= \sum_{j \sim i} \omega_t(j) \beta h + o(h) \\
&= \beta h \sum_{j=1}^N A(i, j) \omega_t(j) + o(h),
\end{aligned}$$

Next we consider the process  $N_t$  which counts the total number of infected nodes in the network at time  $t$ . We find the transition probabilities for  $N_t$  by summing over all susceptible nodes  $S_t = \{i : \omega_t(i) = 0\}$ . In other words, writing  $N_t = \omega_t \cdot \mathbf{1}$  where  $\mathbf{1} = [1 \cdots 1]^T$  is a vector of ones, we have that for  $h > 0$  small,

$$\begin{aligned}
\mathbb{P} [ N_{t+h} - N_t = 1 \mid \omega_t ] &= \mathbb{E} [ (\omega_{t+h} - \omega_t) \cdot \mathbf{1} \mid \omega_t ] \\
&= \sum_{i \in V} \mathbb{E} [ \omega_{t+h}(i) - \omega_t(i) \mid \omega_t ] \\
&= \sum_{i \in V} \mathbb{E} [ \omega_{t+h}(i) - \omega_t(i) \mid \omega_t(i) = 0, \omega_t ] \mathbb{P} [ \omega_t(i) = 0 \mid \omega_t ] \\
&= \sum_{i \in V} \mathbb{P} [ \omega_{t+h}(i) = 1 \mid \omega_t(i) = 0, \omega_t ] \mathbb{E} [ 1 - \omega_t(i) \mid \omega_t ] \\
&= \sum_{i \in V} (1 - \omega_t(i)) \beta h \sum_{j \in V} A(i, j) \omega_t(j) + o(h) \\
&= \beta h \sum_{i, j \in V} A(i, j) \omega_t(j) - \beta h \sum_{i, j \in V} A(i, j) \omega_t(i) \omega_t(j) + o(h) \\
&= \beta h \left( \omega_t^T A \mathbf{1} - \omega_t^T A \omega_t \right) + o(h)
\end{aligned}$$

The second line follows by linearity of expectation. The third line follows because  $\omega_{t+h}(i) - \omega_t(i)$  is either equal to 0 or 1 and can only equal 1 when  $\omega_t(i) = 0$ .

Note that  $A\mathbf{1} = d$  where  $d(j) = \deg(j)$  is the degree of node  $j$ . Write  $D = \text{diag}(d)$  for the diagonal matrix of the node degrees. Then, since  $\omega_t(i)D(i, j)\omega_t(j) = \omega_t(i)D(i, j)$ ,

$$\omega_t^T A\mathbf{1} = \omega_t^T d = \omega_t^T D\mathbf{1} = \omega_t^T D\omega_t.$$

So letting  $L = D - A$  be the combinatorial Laplacian, we get that

$$\omega_t^T A\mathbf{1} - \omega_t^T A\omega_t = \omega_t^T D\omega_t - \omega_t^T A\omega_t = \omega_t^T L\omega_t$$

Then

$$\mathbb{P} [ N_{t+h} - N_t = 1 \mid \omega_t ] = \beta h \omega_t^T L\omega_t + o(h). \quad (3.2.1)$$

We learned to use the Laplacian in this equation from <sup>15</sup>.

Assume that an orientation has been chosen for every edge. Define the  $|E| \times |V|$  **incidence matrix**

$$B(e, x) = \begin{cases} 1 & \text{if } \exists y \sim x \text{ such that } e = (y, x). \\ -1 & \text{if } \exists y \sim x \text{ such that } e = (x, y). \\ 0 & \text{else} \end{cases}$$

We think of  $B$  as a “gradient” operator that turns functions defined on  $V$  into functions defined on  $E$  by assigning the end-points difference to each edge. Then the transpose  $B^T$  is a “divergence” operator that turns functions defined on the edges to ones defined on the vertices.

Moreover, the Laplacian can be written as in the continuous case as the “divergence of the gradient”:

$$L = B^T B.$$

To see this, write  $L(x, y) = \sum_{e, e'} B^T(x, e)B(e', y) = \sum_e B(e, x)B(e, y)$ , and note that if  $x = y$ , then  $B(e, x)B(e, y) = 1$  for every edge incident at  $x$  and is zero otherwise, so we get  $\deg(x)$ ; but if  $x \neq y$  and  $x \sim y$ , then  $B(e, x)B(e, y) = -1$  for the edge between  $x$  and  $y$  and

zero otherwise, so we get  $-1$ .

The matrix  $B$  is sometime referred to as the “square root” of the Laplacian  $L$ . Given a function  $u$  on  $V$ , let  $\nabla u(e) = |u(y) - u(x)|$  if  $e$  connects  $x$  to  $y$ . Then the “gradient norm” of  $u$  is given by

$$u^T L u = u^T B^T B u = (B u)^T (B u) = \|B u\|^2 = \sum_{e \in E} [(\nabla u)(e)]^2.$$

In our case, the set  $V$  is split into two subsets  $S_t$  and  $I_t$  and  $\omega_t$  is the indicator function of the set  $I_t$ . This partition defines a subset of edges called the edge-boundary,  $\partial I_t$ , consisting of all the edges that connect a node in  $I_t$  with a node in  $S_t$ . With this in mind, the quadratic form

$$\omega_t^T L \omega_t = \sum_{e \in E} [(\nabla \omega_t)(e)]^2 = \sum_{e \in \partial I_t} 1 = |\partial I_t|$$

counts the number of edges in  $\partial I_t$ .

So equation (3.2.1) becomes

$$\mathbb{P} [ N_{t+h} - N_t = 1 \mid I_t ] = \beta h |\partial I_t| + o(h). \tag{3.2.2}$$

Note that conditioning on  $\omega_t$  or  $I_t$  is equivalent.

In particular, we see  $N_t$  is a doubly stochastic Poisson process with rate  $\beta$  times the number of active edges at time  $t$ .

Now we determine the probability that a susceptible node  $i$  will be the next node infected after time  $t$ , given  $\omega_t$  and given that an infection occurs. Using the definition of conditional

probability,

$$\begin{aligned}
& \mathbb{P} \left[ \omega_{t+h}(i) - \omega_t(i) = 1 \mid \omega_t, N_{t+h} - N_t = 1 \right] \\
&= \frac{\mathbb{P} \left[ \omega_{t+h}(i) - \omega_t(i) = 1, N_{t+h} - N_t = 1 \mid \omega_t \right]}{\mathbb{P} \left[ N_{t+h} - N_t = 1 \mid \omega_t \right]} \\
&= \frac{\beta h (A\omega_t)(i) + o(h)}{\beta h |\partial I_t| + o(h)} = \frac{(A\omega_t)(i) + o(1)}{|\partial I_t| + o(1)}
\end{aligned}$$

Letting  $h \downarrow 0$  yields,

$$\mathbb{P} \left[ \text{node } i \text{ is the next infected} \mid \omega_t \right] = \frac{\text{number of infected neighbors of node } i \text{ at time } t}{\text{number of active edges at time } t}. \tag{3.2.3}$$

So  $I_t$  evolves by choosing an active edge in  $\partial I_t$  uniformly at random, and then infecting the susceptible endpoint of the chosen edge.

The **arrival times**  $Y_0, Y_1, \dots$  of the SI epidemic  $\omega_t$  are defined by  $Y_0 = 0$  and

$$Y_k = \inf \{ t \geq 0 : N_k = k \} \quad \text{for } k = 1, 2, \dots$$

The **interarrival times**  $T_1, T_2, \dots$  are the times between successive arrivals,

$$T_k = Y_k - Y_{k-1} \quad \text{for } k = 1, 2, \dots$$

Since  $Y_1 < Y_2 < \dots$  the arrival times are not independent of each other. However, as an assumption of the SI model, infection processes are taken to be independent of each other. Hence, the interarrival times  $T_1, T_2, \dots$  are independently distributed random variables. We will see that typically they are decidedly not identically distributed.

**Theorem 3.2.1.** *In the SI epidemic model on the network  $G = (V, E)$  with epidemic pa-*

parameter  $\beta$ , given the set of infected nodes  $I_{Y_k}$ , the next arrival time  $T_{k+1}$  satisfies

$$T_{k+1} \sim \text{Exponential}(\beta|\partial I_{Y_k}|).$$

More precisely,

$$\mathbb{P} [ T_{k+1} \leq t \mid I_{Y_k} = I ] = 1 - e^{-\beta|\partial I|t}.$$

*Proof.* To show that the inter-arrival times are exponentially distributed we will show that they have the memoryless property and then apply Proposition [A.1.2](#).

The proof will rely on equation [\(3.2.2\)](#) and the independence of the competing infection processes.

Fix an arbitrary  $k, Y_k$ , and  $\omega_{Y_k}$ . Then, we observe that for  $a > 0$  the following are equivalent:

1.  $T_{k+1} > a$ ,
2.  $Y_{k+1} - Y_k > a$ ,
3.  $N_{Y_k+a} = k$ ,
4.  $N_{Y_k+a} - N_{Y_k} = 0$ .

In the above, it's easily seen that (1)  $\iff$  (2) by way of the definition of the interarrival time and (2)  $\iff$  (3) via the definition of the arrival times and (3)  $\iff$  (4) by applying the definition of the arrival times once again.

Then, using the equivalence of (1) and (4),

$$\mathbb{P} [ T_{k+1} > a \mid I_{Y_k} = I ] = \mathbb{P} [ N_{Y_k+a} - N_{Y_k} = 0 \mid I_{Y_k} = I ] = 1 - \beta|\partial I|a + o(a),$$

where the final equality follows from [\(3.2.2\)](#) and the fact that the competing infection processes are independent and exponentially distributed.

Moreover, the equivalence of (1) and (4) yields,



$$\begin{aligned}
\mathbb{P} [ T_{k+1} > a + b \mid T_{k+1} > b, I_{Y_k} = I ] &= \mathbb{P} [ N_{Y_k+(a+b)} - N_{Y_k} = 0 \mid N_{Y_k+b} - N_{Y_k} = 0, I_{Y_k} = I ] \\
&= \mathbb{P} [ N_{Y_k+(a+b)} - N_{Y_k+b} = 0 \mid N_{Y_k+b} - N_{Y_k} = 0, I_{Y_k} = I ] \\
&= \mathbb{P} [ N_{Y_k+(a+b)} - N_{Y_k+b} = 0 \mid I_{Y_k} = I ] \\
&= 1 - \beta a |\partial I| + o(a) \\
&= \mathbb{P} [ T_{k+1} > a \mid I_{Y_k} = I ].
\end{aligned}$$

The second line follows from subtracting zero of the form  $N_{Y_k+b} - N_{Y_k} = 0$ . In the third line, we used the independence of the infection process, and in the penultimate line (3.2.2) was once again used.

For  $k = 1, 2, \dots$ , we have seen that  $T_{k+1}$  is exponentially distributed conditional on  $I_{Y_k}$ . Consequently, choose  $\lambda_{k+1}$  so that  $T_{k+1} \sim \text{Exponential}(\lambda_{k+1})$  conditional on  $I_{Y_k}$ . We can then compute the parameter  $\lambda_{k+1}$  given  $I_{Y_k} = I$ . Indeed,

$$\begin{aligned}
\beta |\partial I| &= \lim_{h \downarrow 0} \frac{\beta h |\partial I| + o(h)}{h} \\
&= \lim_{h \downarrow 0} \mathbb{P} [ N_{Y_k+h} - N_{Y_k} = 1 \mid I_{Y_k} = I ] \\
&= \lim_{h \downarrow 0} \mathbb{P} [ T_{k+1} \leq h \mid I_{Y_k} = I ] \\
&= \lim_{h \downarrow 0} \frac{1 - e^{-\lambda_{k+1}h}}{h} = - \lim_{h \downarrow 0} \frac{e^{-\lambda_{k+1}h} - e^0}{h} = \lambda_{k+1}.
\end{aligned}$$

The second line used (3.2.2). The third line used the equivalence between (1) and (4). The fourth line follows since  $T_k$  must be exponentially distributed for some  $\lambda_k$  conditional on  $I_{Y_k}$ . Hence, we have shown that

$$T_{k+1} \sim \text{Exponential}(\beta |\partial I|)$$

given that  $I_{Y_k} = I$ . □

### 3.2.2 Node-event based model

This event based model is the model most commonly used to simulate an SI epidemic. It focuses on the independently competing infection processes of each susceptible node, as oppose to each pairwise competing infection process described in the SI model.

Consider the Markov chain  $\{\psi_k\}_{k=1}^\infty$ , with state space  $\mathcal{A}$  consisting of all subsets  $I$  of  $V$  that are connected and contain the *seed node*  $a$ . To relate to the SI model, we will think of  $I_k$  as being a set of infected nodes. Then we let  $I_k$  be the set of infected nodes,  $\partial I_k$  be the set of active edges,  $\partial I_k(s)$  be the subset of active edges that are connected to susceptible node  $s$ , and  $S_k$  denote the set of susceptible nodes (always respective to the  $k$ th event). Then,  $\psi_{k+1}$  is determined by  $\psi_k$  in the following sense.

Given  $s \in S_k$  we let  $\partial I_k(s)$  denote the set of edges connecting  $s$  to some  $v \in I_k$ . Then, for each node  $s \in S_k$  flip an exponential coin  $X_s \sim \text{Exponential}(\beta|\partial I_k(s)|)$ . Let  $X_{s_0}$  be the minimum over the independent random variables  $\{X_s\}_{s \in S_k}$ . We say that  $s_0$  “wins” against the other independent competing infection processes, and hence is the next infected node. In particular,  $I_{k+1} = I_k \cup \{s_0\}$  and then  $\psi_{k+1}$  is defined accordingly. Consequently, the transition probability matrix for the Markov chain  $\psi_k$  coincides with the transition probability matrix in (3.2.3). Indeed,

$$\begin{aligned} \mathbb{P} [ I_{k+1} = I_k \cup \{s_0\} \mid \psi_k ] &= \mathbb{P} [ X_{s_0} = \min\{X_s : s \in S_k\} \mid \psi_k ] \\ &= \frac{\beta|\partial I_k(s_0)|}{\beta \sum_{s \in S_k} |\partial I_k(s)|} = \frac{|\partial I_k(s_0)|}{|\partial I_k|}. \end{aligned}$$

Where the second line follows by Proposition A.1.1 part 2.

Moreover, from Proposition A.1.1 part 1 we see that the random variable  $\min\{X_s : s \in S_k\}$ , is distributed like  $\text{Exponential}\left(\beta \sum_{s \in S_k} |\partial I_k(s)|\right) = \text{Exponential}(\beta|\partial I_k|)$ . Hence, inter-arrival times of the node-event based Markov process and the inter-arrival times of the

the SI model (compare to Theorem 3.2.1) are distributed identically.

### 3.2.3 Edge-event based model

The SI edge-event based model, is a model to simulate an SI epidemic which requires flipping relatively fewer random exponential coins. It's based off of knowing the desired transition probabilities and desired inter-arrival time as opposed to focusing on the independent and competing infection processes of each susceptible node.

Consider the Markov chain  $\{\eta_k\}_{k=1}^\infty$ , with state space  $\mathcal{A}$ , consisting of all subsets  $I$  of  $V$  that are connected and contain the seed node  $a$ . Given  $\eta_k$ , let  $I_k, S_k$ , and  $\partial I_k$  be as before. The Markov chain  $\eta_k$  evolves in the following way.

Given a seed node  $a$ , we set  $\eta_0 = \{a\}$  and iteratively define  $\eta_{k+1}$  by choosing a uniformly random  $e_{k+1} \in \partial I_k$  and infecting its susceptible endpoint  $v_{k+1}$ . In particular,  $I_{k+1} = I_k \cup \{e_{k+1}\}$ . Moreover, we define the random variable  $T_{k+1} \sim \text{Exponential}(\beta|\partial I_k|)$  to be the inter-arrival time. Then, by Theorem 3.2.1, we see the inter-arrival times coincide with the SI model. Moreover, the transition probability matrix coincides as well, since

$$\mathbb{P} [ I_{k+1} = I_k \cup \{v_{k+1}\} \mid \eta_k ] = \mathbb{P} [ e_{k+1} \in \partial I_k(v_{k+1}) \mid \eta_k ] = \frac{|\partial I_k(v_{k+1})|}{|\partial I_k|}$$

is the same as (3.2.3).

### 3.2.4 Variable-lengths model

We describe a continuous time Markov chain with seed  $a$  denoted,  $W_t$ , on the state space  $\mathcal{A}$ , where  $\mathcal{A}$  is as before. In this model, we let time be parametrized by a random distance on the contact network  $G = (V, E)$ . More precisely, for each edge  $e \in E$  we assign a weight given by i.i.d. random variables  $X_e \sim \text{Exponential}(\beta)$ . Then, we define the length of a walk

$\gamma$  to be

$$\ell_X(\gamma) = \sum_{e \in E} \mathcal{N}(\gamma, e) X_e$$

where  $\mathcal{N}(\gamma, e)$  denotes the number of times the walk  $\gamma$  traverses the edge  $e$ . Then, the distance from  $a$  to  $b$ , is  $d(a, b) = \inf_{\gamma: a \rightsquigarrow b} \ell_X(\gamma)$ , and the continuous time process  $W_t$  with seed  $a$  in  $V$  is defined as

$$W_t = \{v \in V : d(a, v) \leq t\}.$$

We let  $\{Y_k\}$  denote the arrival times. That is,  $Y_0 = 0$  and  $Y_k = \inf\{t \geq 0 : N_t = k\}$  where  $N_t = |W_t|$ . Moreover, let  $\{T_k\}$  denote the set of inter-arrival times, so that  $T_{k+1} = Y_{k+1} - Y_k$ . Finally, the sets  $I_t, \partial I_t, S_t$ , and  $\partial I_t(s)$  will be as before.

**Theorem 3.2.2.** *For a contact network  $G = (V, E)$  with i.i.d. edge-weights  $X_e \sim \text{Exponential}(\beta)$  for each  $e \in E$ , the continuous time process  $W_t$  with seed  $a$  in  $V$  is a jump process with probability transition matrix satisfying*

$$\mathbb{P} [ W_{Y_{k+1}} = W_{Y_k} \cup \{s\} \mid W_{Y_k} ] = \frac{|\partial I_{Y_k}(s)|}{|\partial I_{Y_k}|} \quad \text{for each } s \text{ in } S_{Y_k} \quad (3.2.4)$$

and the randomly distributed inter-arrival times

$$T_{k+1} \sim \text{Exponential}(\beta |\partial I|) \quad (3.2.5)$$

given  $I_{Y_k} = I$ .

*Proof.* This proof relies on induction on  $k + 1$ . We first start with the base case  $k = 0$ . Then  $I_{Y_0} = \{a\}$  and  $v \in S_{Y_0}$  if and only if  $v \sim a$ . Then, letting  $e_u$  denote the edge  $\{a, u\}$  for each  $u$  neighboring  $a$ , we see that

$$T_1 = Y_1 - Y_0 = \min\{X_{e_u} + d(a, a) - Y_0 : e_u \in \partial I_0\} = \min\{X_{e_u} : e_u \in \partial I_0\} \sim \text{Exponential}(\beta |\partial I_{Y_0}|)$$

by Proposition A.1.1. Moreover, since each  $X_{e_u}$  is i.i.d., then the location of the minimum

of  $\{X_{e_u}\}$  has a uniform probability distribution over the links that have  $a$  as an endpoint. In particular,

$$\mathbb{P} [ X_{e_s} = \min\{X_e : e \in \partial I_0\} \mid W_{Y_0} ] = \frac{|\partial I_{Y_0}(s)|}{|\partial I_{Y_0}|},$$

coinciding with (3.2.3).

Now, assume that (3.2.4) and (3.2.5) continue to hold for  $n = 2, 3, \dots, k$ . We wish to show that it holds for  $n = k + 1$  as well. For convenience, re-label each link in  $\partial I_{Y_k}$  so that they are called  $e_j = \{y_j, s_j\}$  with  $y_j \in I_{Y_k}$  and  $s_j \in S_{Y_k}$  for  $j = 1, 2, \dots, m_k$ . We note that in the base case  $y_j = a$  for all  $e_j \in \partial I_{Y_0}$ . Now, since  $y_j \in I_{Y_k}$  for all  $j$ , it follows from the definition of  $W_t$  that  $d(a, y_j) \leq Y_k$  for each  $j$ . However, since  $s_j \notin I_{Y_k}$  we also have  $d(a, y_j) + X_{e_j} > Y_k$  for each  $j$ , or alternatively

$$X_{e_j} > Y_k - d(a, y_j) \quad \forall e_j \in \partial I_{Y_k}. \quad (3.2.6)$$

Then, fixing arbitrary  $j \in \{1, 2, \dots, m_k\}$  we observe,

$$\begin{aligned} \mathbb{P} [ X_{e_j} + d(a, y_j) > Y_k + h \mid W_{Y_k} ] &= \mathbb{P} [ X_{e_j} > Y_k - d(a, y_j) + h \mid W_{Y_k} ] \\ &= \mathbb{P} [ X_{e_j} > Y_k - d(a, y_j) + h \mid W_{Y_k}, X_{e_j} > Y_k - d(a, y_j) ] \\ &= \mathbb{P} [ X_{e_j} > h \mid W_{Y_k} ] \\ &= \mathbb{P} [ X_{e_j} > h ] \\ &= e^{-\beta h} \quad \text{for all } j = 1, 2, \dots, |\partial I_{Y_k}|. \end{aligned}$$

The second line follows by (3.2.6) and the third by the strong memoryless property (see Proposition A.1.2). The fourth line follows due to independence of  $X_{e_j}$  and the current state of the Markov chain, and the final equality follows since each edge-weight was chosen i.i.d. and distributed like  $\text{Exponential}(\beta)$ .

In particular, each random variable  $d(a, y_j) + X_{e_j}$  conditioned on the current state of  $W_t$

is exponentially distributed with parameter  $\beta$ , so that the edge achieving the minimum of  $d(a, y_j) + X_{e_j}$  is chosen uniformly at random from the set of active edges. This completes induction on the transition probability matrix. On the other hand, if we let  $Z_j$  be the random variable  $X_{e_j} + d(a, y_j)$  conditioned on  $W_{Y_k}$  and  $e_j \in S_{Y_k}$ , then

$$T_{k+1} = Y_{k+1} - Y_k = \min\{Z_j - Y_k \mid j = \{1, 2, \dots, |\partial I_{Y_k}|\}\} \sim \text{Exponential}(\beta|\partial I_{Y_k}|).$$

This can be seen by applying Proposition [A.1.1](#) to the preceding calculation. □

### 3.2.5 Dijkstra on the Fly

I believe that Dijkstra on the fly is the same model as the exponential weights model, except that instead of tossing all of the random edge lengths at the beginning, we only toss an i.i.d.  $\text{Exponential}(\beta)$  weight for an edge if and when the edge is in the active edge-set the first time. As this model flips fewer exponential random coins, it's faster to compute than the random lengths model.

The typical Dijkstra's algorithm is described in Section [3.1](#).

For Dijkstra on the fly, you modify the typical Dijkstra's algorithm, see Algorithm [2](#), by inserting a new line after line 8:

$$\mathbf{set} \ell(u, v) = X_{\{u, v\}}$$

where  $X_{\{u, v\}}$  is an i.i.d. random variable and  $X_{\{u, v\}} \sim \text{Exponential}(\beta)$ . Consequently, the modified Dijkstra's algorithm does not require you to input an edge-density function  $\ell$ .

Thus we can see the reason that this equivalence appears to be immediate is that the algorithm is blind to whether we flip an exponential coin when we need it, i.e., when the edge is active (as in the Dijkstra on the fly algorithm), or if we flip a lot of coins and embed them in the function  $\ell(u, v)$  and then only look at them when we need them, i.e., when the

edge  $\{u, v\}$  is an active edge.

### 3.3 Metrics and Centralities

A **graph centrality** is a function  $P : V \rightarrow \mathbb{R}$  that assigns a value to each node. This value represents how important a node is with respect to the graph topology. It is often convenient to normalize a centrality so that it can be thought of as a probability distribution, i.e., the sum of its values over every node adds up to one.

A **graph metric** is a symmetric, non-degenerate function  $d : V \times V \rightarrow [0, \infty]$  that satisfies the triangle inequality. Every graph metric induces a centrality denoted  $P_d$  on the graph. In particular, we define

$$P_d(v_i) := N^{-1} \sum_{\substack{v_j \in V \\ v_j \neq v_i}} d(v_i, v_j)^{-1}.$$

Heuristically, this induced centrality says that if node  $v_i$  is close (with respect to the metric  $d$ ) to many nodes, then it is important to the graph topology.

#### 3.3.1 Metrics

*Epidemic Hitting Time:* We have established a model to describe the epidemic process. There are different properties of the same epidemic model one may wish to study. Herein, we analyze the **epidemic hitting time**. The epidemic hitting time turns out to be a graph metric,  $H$ , where  $H(i, j)$  is the expected time it takes for an infection starting at node  $i$  to infect node  $j$ . To numerically approximate the epidemic hitting time, one can run an  $SI$  simulation starting at node 1 a large number of times and then for each simulation keep track of the time at which each node was infected. Then, repeat this process for nodes  $2, 3, 4, \dots, N$ .

One may wonder why not just use the epidemic hitting time as our graph quantity? It

is the most accurate for understanding how quickly an infection will spread. However, this accuracy comes at a cost. Analytically the epidemic hitting time would be impossible to compute for any network that is large enough to represent a real-world community. Even numerically the epidemic hitting time is very computationally difficult to compute. Even for the modest sized networks analyzed here, the epidemic hitting time sometimes required several weeks to compute.

*Effective Resistance:* The **effective resistance**, which we have seen can be realized as a special case of the  $p$ -modulus on finite graphs, is a graph metric,  $R$ , that naturally arises by consider a graph to be an electrical circuit, where each edge has 1 Ohm of resistance. The pairwise effective-resistance,  $R(i, j)$  is the total effective resistance measured between two nodes in the circuit. The effective resistance, first studied as a graph metric in<sup>16</sup>, can be defined from the graph Laplacian as in Equation (ER). The effective resistance has been shown to be closely related to many properties of random Markov processes<sup>17</sup>, including capturing the covering and commute times of random walks<sup>18</sup> and recurrence/transience of random walks<sup>19</sup>.

$$\begin{aligned}
 L^\dagger &= \left( L + \mathbf{1}\mathbf{1}^T/N \right)^{-1} - \mathbf{1}\mathbf{1}^T/N \\
 R &= \text{diag}(L^\dagger) + \text{diag}(L^\dagger)^T - 2L^\dagger
 \end{aligned}
 \tag{ER}$$

### 3.3.2 Centralities

*Spectral Centrality:* The spectral centrality ( $S$ ) is an important measure of centrality, designed to keep the principle ‘the importance of a node depends on the importance of its neighbors’ in mind. Spectral centrality does well in characterizing simple dynamics like diffusion, and is the basis of Google’s PageRank algorithm<sup>20</sup>. The spectral centrality of a



node  $i$  is defined by

$$S(i) = \lambda^{-1} \sum_{j \sim i} v_j,$$

where  $\lambda$  is the largest eigenvalue of the adjacency matrix and  $v_j$  is the  $j$ th component of the corresponding eigenvector.

*Degree Centrality:* The degree centrality ( $D$ ) is perhaps the simplest measure of network centrality, where a nodes centrality is exactly the number of neighbors a node has. One benefit of degree centrality is its simplicity to compute. However, two immediate shortcomings are that it ignores how important a node's neighbors are, and that it will frequently rank the importance of large numbers of nodes to be the same.

*Betweenness Centrality:* The betweenness centrality ( $B$ ) first introduced in<sup>21</sup> and popularized by Freeman<sup>22</sup> is a method of measuring the importance of a node based off of the percentage of shortest paths that it lies on. It is a measure of the influence a node has over the spread of information through a network. In particular, the betweenness centrality is defined by

$$B(i) = N^{-1} \sum_{j \neq i \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}},$$

where  $\sigma_{jk}$  denotes the number of shortest paths between  $j$  and  $k$  and  $\sigma_{jk}(i)$  is the number of shortest paths from  $j$  to  $k$  that visit node  $i$ . The betweenness centrality is good at determining bottlenecks of networks, however does not take into account non-geodesic paths.

*Communicability Centrality:* The communicability centrality ( $C$ ) is an adaptation of Freeman's betweenness centrality that takes into account all independent walks between two nodes, instead of just the geodesic paths. The pairwise communicability can be computed from the spectrum of the adjacency matrix<sup>23</sup> by

$$C(i) = \sum_{j=1}^N (v_{j,i})^2 e^{\lambda_j},$$

where  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$  are the eigenvalues of the adjacency matrix, and  $v_{k,\ell}$  is the  $\ell$ th

element of the eigenvector corresponding to the  $k$ th eigenvector.

## 3.4 Real-World Networks

In simulations, we used real-world networks. All of which are available for download at [http://ece.k-state.edu/epicenter\\_wiki/index.php/Products](http://ece.k-state.edu/epicenter_wiki/index.php/Products). Included is a brief description of each network and a citation of where they have been popularized.

*Adjnoun*: This is the network of common adjective and noun adjacencies for the novel "David Copperfield" by Charles Dickens, as described by M. Newman. Nodes represent the most commonly occurring adjectives and nouns in the book. Edges connect any pair of words that occur in adjacent position in the text of the book.<sup>24</sup>

*Power*: An undirected unweighted representation of the topology of the western states power grid of the United States compiled by Watts and Strogatz<sup>25</sup>.

*Karate*: The network of friendships between the 34 members of a karate club at a US university, just before the club split, as described by Wayne Zachary<sup>1</sup>.

*Football*: The football network, compiled by Girvan and Newman<sup>26</sup>, is a network of American football games between Division I colleges. A link between two teams represents that a game was played between the two teams during the fall of 2000 season.

*Dolphins*: The network contains an undirected social network of frequent associations between 62 dolphins in a community living off Doubtful Sound, New Zealand, as compiled by Lusseau<sup>27</sup>.

*High Energy Physics Collaboration (HEP)*: The collaboration network of scientists posting preprints on the high-energy physics to the archive at [www.arxiv.org](http://www.arxiv.org) during 1995-1999, as compiled by Newman<sup>28</sup>.

*Lesmis*: The weighted network of coappearances of characters in Victor Hugo's novel "Les Miserables". Nodes represent characters in the novel and edges connect any pair of characters that appear in the same chapter of the book.<sup>29</sup>

*Polblogs*: Political blogosphere Feb. 2005 Data compiled by Adamic and Glance<sup>30</sup>. Links between blogs were automatically extracted from a crawl of the front page of the blog.

*Netscience*: A coauthorship network of scientists working on network theory and experiment, as compiled by Newman<sup>24</sup>. The network was compiled from the bibliographies of two review articles on networks, with a few additional references added by hand. The version used here contains just the largest component of 379 scientists.

*Facebook*: This dataset consists of ‘friends lists’ from Facebook. Facebook data was collected from survey participants using the Facebook app and their information was anonymized. The original data set has 4,039 nodes and 88,234 edges. We looked at a cluster of only approximately 3,000 edges. This data was compiled by Leskovec and McAuley and used in<sup>31</sup>.

*General Relativity and Quantum Cosmology Collaboration (GRQC)*: This graph represents the collaboration network in General Relativity and Quantum Cosmology from e-prints on arXiv during the period January 1993 to April 2003 (124 months). This data was compiled by Leskovec, Kleinberg, and Faloutsos and used in<sup>32</sup>.

*Summary of Network Structure Properties:*

Network	Number of Nodes	Number of Edges	Average Node Degree	Degree Second Moment	Degree Assortativity	Average Clustering Coefficient	Diameter	Density	Spectral Radius	Algebraic Connectivity	Normalized Effective Resistance
Adjnoun	112	425	7.589	104.53	-0.1293	0.1569	5	0.0683	13.15	.695	8.927
Dolphins	62	159	5.129	34.903	-0.0436	0.309	8	0.0841	7.194	0.173	11.725
Football	115	615	10.695	115.217	0.153	0.407	4	0.0938	10.818	1.460	2.447
Hep	8361	15751	3.768	32.731	0.294	0.329	19	4.5E-4	23	-8.84E-14	Inf
Karate	34	78	4.588	35.647	-0.476	0.256	5	0.139	6.726	0.468	6.029
Lesmis	77	254	6.597	79.532	-0.1652	0.499	5	0.087	12.006	0.205	11.876
Netscience	1589	2742	3.451	23.947	0.467	0.693	17	2.1E-3	19.024	-1.94E-14	Inf
Polblogs	1490	19090	25.624	2499.357	-0.229	0.226	9	8.6E-3	46.514	NAN	1.98E+32
Power	4941	6594	2.669	10.333	3.46E-3	0.103	46	5.4E-4	7.483	7.59E-4	9670.86

## 3.5 Numerical Methods and Results

For each network described above, the epidemic hitting time was approximated by averaging the results of 100 SI simulations, then symmetrizing the resulting matrix by averaging it with its transpose. This creates the epidemic hitting time graph metric, from which the corresponding network centrality was computed as described at the beginning of Section 3.3. These results were compared to the effective resistance metric as well as the spectral, degree, betweenness, and communicability centralities. Comparisons were done in two ways.

*Total variation distance of probability measures:* One comparison was to compute the total variation between two centralities after normalizing the centralities so that they can be interpreted as a probability measure on the nodes of the graph. The total variation,  $\delta$  between two distributions,  $P, Q$  on the probability space  $(\Omega, \mathcal{F})$  is defined by

$$\delta(P, Q) := \sup_{A \in \mathcal{F}} \left| \sum_{e \in A} P(e) - Q(e) \right|.$$

In the case of a finite graph, this is equivalent<sup>33</sup> to

$$\frac{1}{2} \sum_{e \in E} |P(e) - Q(e)|.$$

In particular, the total variation distance of probability measures is a value between zero and two, where zero means the distributions overlap completely, and two means they are mutually singular.

*Spearman's Rank Correlation Coefficient:* The Spearman correlation coefficient is the Pearson correlation coefficient for ranked variables. Hence, it takes values between negative one and one and should be interpreted like the Pearson correlation coefficient<sup>34</sup>. In particular, here it can be interpreted as how well a monotonic function can be defined between two distributions, where the sign indicates whether the function would be monotonically increasing or decreasing. Below are the results of these comparisons.

<b>Network</b>	<b>ER</b>	<b>BC</b>	<b>DC</b>	<b>CC</b>	<b>SC</b>
Adjoun	0.9918	0.8514	0.9814	0.9732	0.9727
Dolphins	0.9846	0.7483	0.8579	0.8533	0.8521
Facebook	0.9603	0.4960	0.6715	0.8559	0.8875
Football	0.8725	0.6142	0.6753	0.3683	0.6471
GRQC	0.9651	0.5181	0.7536	0.7969	0.7960
HEP	0.9554	0.5387	0.6956	0.8834	0.8976
Karate	0.9125	0.7814	0.8540	0.9419	0.9525
Lesmis	0.9900	0.7076	0.9490	0.9643	0.9288
Netscience	0.9744	0.3800	0.4543	0.6389	0.6583
Polblogs	0.996	0.8824	0.9933	0.968	0.9681
Power	0.8277	0.3133	0.3247	0.4207	0.5919

Figure 3.1: A table of values of the Spearman Rank coefficient between each centrality and the epidemic hitting time centrality. We note that in every network except the Karate club network, the Effective Resistance outperforms every other centrality. We also remark that in<sup>2</sup>, there was a similar observation of the Karate club network not behaving like the other networks, likely due its being a small network with a few dominating nodes.

The raw data is pretty convincing that the effective resistance is a strong indicator of the influence that a node has over the spread of an epidemic throughout a social network. However, the tables give little insight as to why the effective resistance outperforms other graph quantities. To try to build this intuition as to why, we look at the heatmaps of the Lesmis network Figure 3.3 that represents the importance of each node according to the respective graph quantity. We see that effective resistance and the epidemic hitting time are the only two graph quantities that assign relative importance to peripheral nodes of the network. Since in section 1.1.4 we showed that the effective resistance is equivalent to the pairwise 2-modulus, this coincides well with our intuition that the modulus assigns high values to nodes on short walks, but also takes into account the many longer walks.

### 3.5.1 Properties of the Epidemic Hitting Time

While we have witnessed that the effective resistance is a good indicator of the importance of a node in an epidemic process, we still know very little about the epidemic hitting time metric. Here are a few properties of the epidemic hitting time metric. Throughout this

Network	ER	BC	DC	CC	SC
Adjnoun	0.0304	0.4834	0.2123	0.4377	0.2172
Dolphins	0.0288	0.4155	0.1631	0.3424	0.347
Facebook	0.0308	0.8277	0.4633	0.7698	0.5904
Football	0.0080	0.1701	0.0233	0.1113	0.0681
GRQC	0.0485	0.6710	0.3498	0.9669	0.8935
Hep	0.0492	0.6296	0.2698	0.9633	0.9635
Karate	0.0259	0.5841	0.2181	0.3271	0.1617
Lesmis	0.0296	0.6902	0.2363	0.5282	0.3490
Netscience	0.0252	0.7706	0.2452	0.6294	0.7331
Polblogs	0.0214	0.6436	0.3734	0.6293	0.3875
Power	0.0433	0.7098	0.2220	0.2972	0.9846

Figure 3.2: A table of values of the total variation between each centrality and the epidemic hitting time centrality. We note that in every network, the Effective Resistance outperforms every other centrality. This is in large part due to the weight of the peripheral nodes. The effective resistance and epidemic hitting time assign a large enough weight to these peripheral nodes, while the other centralities underestimate their importance.

section we let  $\rho : E \times E \rightarrow [0, \infty]$  and  $d : E \times E \rightarrow [0, \infty]$  be the shortest walk length and exponential weights length of a graph  $G = (V, E)$ .

**Proposition 3.5.1.** *The epidemic hitting time is a metric.*

*Proof.* Using the equivalence of the SI model and the exponential lengths model, we see that the epidemic hitting time can be thought of as an average of graph metrics and is hence a metric itself.  $\square$

We say that a graph  $G_1 = (V_1, E_1)$  is a refinement of  $G = (V, E)$  if  $V \subset V_1$  and  $E \subset E_1$ .

**Remark 3.5.1.** We note that if  $G_1$  is a refinement of  $G$ , then for all  $a, b \in V$ , the epidemic hitting time  $\tau_{a,b}(G) \geq \tau_{a,b}(G_1)$ . This is because adding more edges can only shorten a graph, and adding new vertices cannot length the shortest walk between two nodes.

**Proposition 3.5.2.** *If  $G$  is a tree, then the epidemic hitting time  $H$  and the shortest path metric  $d$  are scalar multiples of each other. Furthermore, for an arbitrary graph  $\tilde{G}$ , the epidemic hitting time is bounded above by the shortest path metric.*

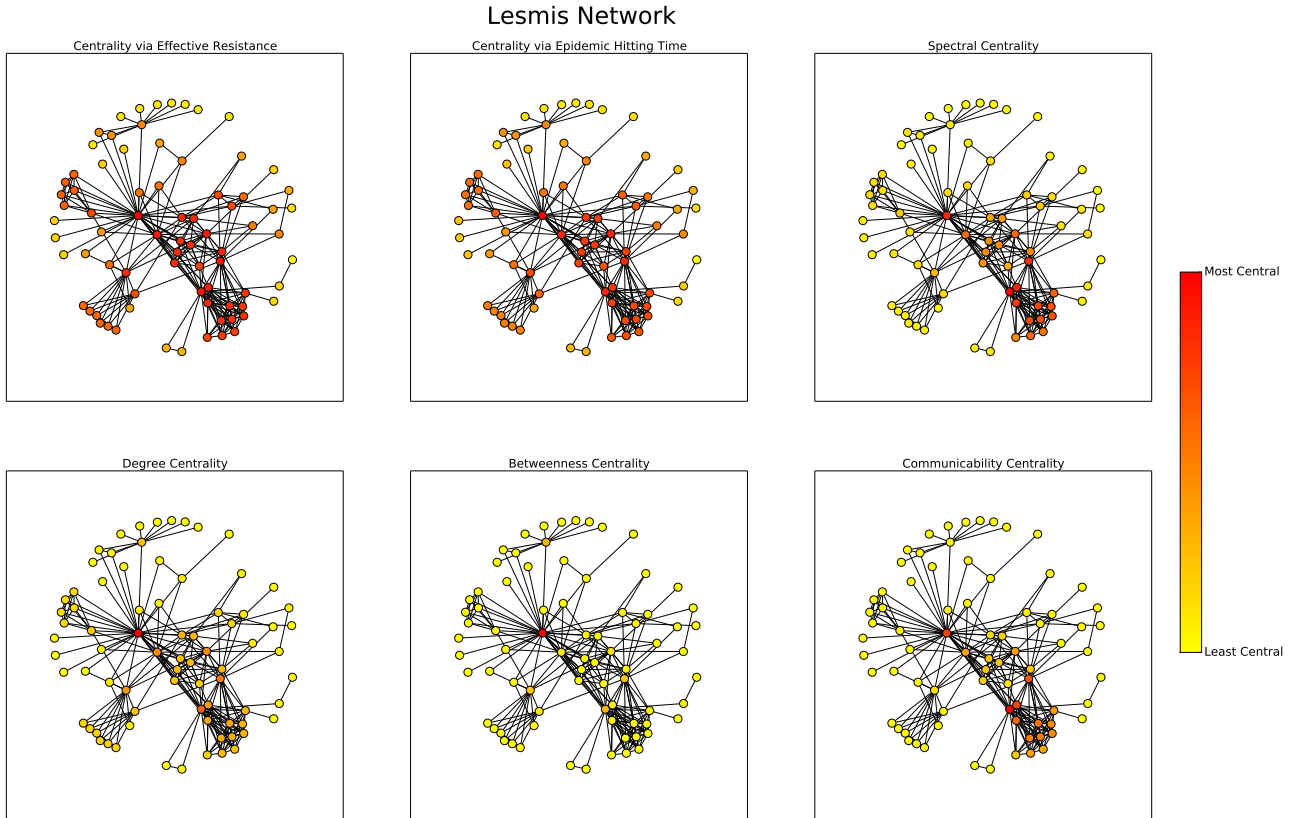


Figure 3.3: Each figure shows the importance of the nodes in the Lesmis network, based off of a different centrality. We observe that not only do the effective resistance and epidemic hitting time centralities visually appear to match up quite well, they are also the only two centralities that assign relative importance to the peripheral nodes. See<sup>3</sup>.

*Proof.* This also follows most clearly from the exponential lengths model of the SI epidemic. On a tree, since there is a unique simple path from  $a$  to  $b$ , then the expectation of the distance  $d(a, b)$  can be computed using linearity of the expectation. That is, if  $\gamma_{ab}$  denotes the unique simple path starting at  $a$  and ending at  $b$ ,

$$\mathbb{E}(d(a, b)) = \beta^{-1} \sum_{e \in E} \mathcal{N}(e, \gamma_{ab}) = \beta^{-1} \rho(a, b),$$

where  $\rho : V \times V \rightarrow \mathbb{N}$  is the unweighted graph distance. □

**Example 3.5.1.** If  $G = (V, E)$  is a star-shaped graph, that is one node connected to  $N - 1$  other nodes each with degree 1 and if  $c$  is the central node, then

$$H(c, v) = \beta^{-1} \quad \text{for all } v \in V \setminus \{c\}. \quad (3.5.1)$$

This follows since the graph distance from  $c$  to  $v$  on the star-shaped graph (a special instance of a tree) is 1 for all  $v \in V \setminus \{c\}$ .

**Proposition 3.5.3.** *For the complete graph  $K_N$  on  $N$  nodes, the epidemic hitting time satisfies*

$$H_N(i, j) \equiv \frac{1 - \delta_{ij}}{N - 1} \sum_{n=1}^{N-1} \sum_{k=1}^n \frac{1}{k(N - k)}. \quad (3.5.2)$$

Therefore, for an arbitrary graph  $G$ , the epidemic hitting time is bounded below by the expression in (3.5.2).

*Proof.* Let  $a$  be the initially infected node. Then for any  $b \neq a$ , we want to compute  $H(a, b) = \mathbb{E}(\tau_{ab})$ . Let  $X_k \sim \text{Bernoulli}(\frac{1}{N-k})$  be i.i.d. random variables. Then  $\{X_k\}_{k=1}^{N-1}$  are Bernoulli random variables with the same probability of a ‘win’ as the probability that  $b$  is infected  $k$ th given  $b$  was not previously infected. Let  $Y_b$  be the random variable that represents which infection  $b$  was. Then,

$$\begin{aligned} \mathbb{P}(Y_b = k) &= \mathbb{P}(X_1 = 0, X_2 = 0, \dots, X_{k-1} = 0, X_k = 1) \\ &= \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0) \cdots \mathbb{P}(X_{k-1} = 0)\mathbb{P}(X_k = 1) \\ &= \left(1 - \frac{1}{N-1}\right) \left(1 - \frac{1}{N-2}\right) \cdots \left(1 - \frac{1}{N-(k-1)}\right) \frac{1}{N-k} \\ &= \left(\frac{N-2}{N-1}\right) \left(\frac{N-3}{N-2}\right) \cdots \left(\frac{N-k}{N-(k-1)}\right) \frac{1}{N-k} = \frac{1}{N-1}. \end{aligned}$$

This result should not be surprising, based off the symmetry of the complete graph.

Moreover, after  $k - 1$  nodes have been infected, this means there are  $k$  total nodes infected (the seed node) and  $N - k$  susceptible nodes. Since the graph is complete, this



means there are  $k(N - k)$  edges connecting susceptible nodes to infected nodes, so that  $t_k$  the  $k$ th inter-arrival time is distributed like  $\text{Exponential}(k(N - k)\beta)$ , and consequently, the time until the  $k$ th infection is  $T_k := \sum_{j=1}^k \frac{1}{\beta k(N-k)}$ . Conditioning on  $Y_b$  attains,

$$\mathbb{E}(\tau_{ab}) = \sum_{k=1}^{N-1} \mathbb{E}(\tau_{ab} | Y_b = k) \mathbb{P}(Y_b = k) = \frac{1}{N-1} \sum_{k=1}^{N-1} T_k.$$

Since  $b$  was arbitrary in  $K_N$  other than not  $b \neq a$ , and since  $H$  is a metric so that  $H(i, j) = 0$  whenever  $i = j$ , we have shown (3.5.2).  $\square$

Combining our refinement observation (Remark 3.5.1) with Propositions 3.5.2 and 3.5.3, we have proven the corollary,

**Corollary 3.5.4.** *Given a contact graph  $G = (V, E)$  with  $a, b$  being nodes in  $V$ , then*

$$\beta^{-1} \rho(a, b) \leq \tau_{a,b} \leq H_N(a, b).$$

The variable lengths model not only lends itself to better demonstrating desirable properties of the epidemic hitting time, but it also offers a partial explanation as to why numerics indicate the effective resistance is so closely correlated to the epidemic hitting time. In the work of Lyons, Pemantle, and Peres<sup>35</sup>, it was shown that the expected shortest path on a graph with i.i.d. exponentially distributed random weights is bounded below by the effective resistance on the graph with unit resistance. In particular, this shows that the epidemic hitting time is bounded below by the effective resistance. This result confirms the recent finding of other authors (for example Sikic et al.<sup>36</sup>) that the impact of peripheral nodes is typically underestimated in epidemic models.

# Bibliography

- [1] Wayne W Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, pages 452–473, 1977.
- [2] José Ricardo Furlan Ronqui and Gonzalo Travieso. Analyzing complex networks through correlations in centrality measurements. *arXiv preprint arXiv:1405.7724*, 2014.
- [3] Mile Sikic, Hrvoje Stefancic, Nino Antulov-Fantulin, and Alen Lancic. Epidemic centrality and the underestimated epidemic impact on network peripheral nodes. Technical report, 2011.
- [4] Nathan Albin, Megan Brunner, Roberto Perez, Pietro Poggi-Corradini, and Natalie Wiens. Modulus on graphs as a generalization of standard graph theoretic quantities. *arXiv preprint arXiv:1504.02418*, 2015.
- [5] D. Drasin, A. A. Gol'dberg, and P. Poggi-Corradini. Quasiconformal mappings in value-distribution theory. In *Handbook of complex analysis: geometric function theory. Vol. 2*, pages 755–808. Elsevier Sci. B. V., Amsterdam, 2005. doi: 10.1016/S1874-5709(05)80022-8. URL [http://dx.doi.org/10.1016/S1874-5709\(05\)80022-8](http://dx.doi.org/10.1016/S1874-5709(05)80022-8).
- [6] Josh Ericson, Pietro Poggi-Corradini, and Hainan Zhang. Effective resistance on graphs and the epidemic quasimetric. *Involve, a Journal of Mathematics*, 7(1):97–124, 2013.
- [7] W Ellens. *Effective resistance and other graph measures for network robustness*. PhD thesis, Masters thesis, Mathematical Institute, University of Leiden, 2011.
- [8] R Tyrrell Rockafellar. *Convex analysis*. princeton landmarks in mathematics, 1997.

- [9] Martin Aigner. *Discrete mathematics*. American Mathematical Society, Providence, RI, 2007. ISBN 978-0-8218-4151-8. Translated from the 2004 German original by David Kramer.
- [10] Yury Makarychev. A short proof of kuratowski’s graph planarity criterion. *Journal of Graph Theory*, 25(2):129–131, 1997.
- [11] Bertrand Duplantier, Jason Miller, and Scott Sheffield. Liouville quantum gravity as a mating of trees. *arXiv preprint arXiv:1409.7055*, 2014.
- [12] Scott Sheffield. Quantum gravity and inventory accumulation. *arXiv preprint arXiv:1108.2241*, 2011.
- [13] Gilles Schaeffer. Bijective census and random generation of eulerian planar maps with prescribed vertex degrees. *Electron. J. Combin.*, 4(1):20, 1997.
- [14] Arpita Ghosh, Stephen Boyd, and Amin Saberi. Minimizing effective resistance of a graph. *SIAM review*, 50(1):37–66, 2008.
- [15] Piet Van Mieghem. Exact markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold. *arXiv preprint arXiv:1402.1731*, 2014.
- [16] Douglas J Klein and M Randić. Resistance distance. *Journal of Mathematical Chemistry*, 12(1):81–95, 1993.
- [17] David Aldous and Jim Fill. Reversible markov chains and random walks on graphs, 2002.
- [18] Ashok K Chandra, Prabhakar Raghavan, Walter L Ruzzo, Roman Smolensky, and Prason Tiwari. The electrical resistance of a graph captures its commute and cover times. *Computational Complexity*, 6(4):312–340, 1996.
- [19] Peter G Doyle and J Laurie Snell. Random walks and electric networks. *AMC*, 10:12, 1984.

- [20] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1):107–117, 1998.
- [21] Alex Bavelas. A mathematical model for group structures. *Human organization*, 7(3):16–30, 1948.
- [22] Linton C Freeman. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41, 1977.
- [23] Ernesto Estrada and Naomichi Hatano. Communicability in complex networks. *Physical Review E*, 77(3):036111, 2008.
- [24] Mark EJ Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, 74(3):036104, 2006.
- [25] Duncan J Watts and Steven H Strogatz. Collective dynamics of small-world networks. *nature*, 393(6684):440–442, 1998.
- [26] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.
- [27] David Lusseau, Karsten Schneider, Oliver J Boisseau, Patti Haase, Elisabeth Slooten, and Steve M Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54(4):396–405, 2003.
- [28] Mark EJ Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001.
- [29] Donald Ervin Knuth, Donald Ervin Knuth, and Donald Ervin Knuth. *The Stanford GraphBase: a platform for combinatorial computing*, volume 37. Addison-Wesley Reading, 1993.

- [30] Lada A Adamic and Natalie Glance. The political blogosphere and the 2004 us election: divided they blog. In *Proceedings of the 3rd international workshop on Link discovery*, pages 36–43. ACM, 2005.
- [31] Jure Leskovec and Julian J Mcauley. Learning to discover social circles in ego networks. In *Advances in neural information processing systems*, pages 539–547, 2012.
- [32] Jure Leskovec, Jon Kleinberg, and Christos Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2, 2007.
- [33] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer. *Markov chains and mixing times*. 2009.
- [34] Ann Lehman. *JMP for basic univariate and multivariate statistics: a step-by-step guide*. SAS Institute, 2005.
- [35] Russell Lyons, Robin Pemantle, and Yuval Peres. Resistance bounds for first-passage percolation and maximum flow. *Journal of Combinatorial Theory, Series A*, 86(1): 158–168, 1999.
- [36] Mile Šikić, Alen Lančić, Nino Antulov-Fantulin, and Hrvoje Štefančić. Epidemic centrality is there an underestimated epidemic impact of network peripheral nodes? *The European Physical Journal B*, 86(10):1–13, 2013.

# Appendix A

## Probability

### A.1 The Exponential Distribution

A continuous random variable  $X$  is said to have the exponential distribution with parameter  $\lambda$  if its probability density function,  $f_X(t)$  is defined as in (A.1.1).

$$f_X(t) = \lim_{h \rightarrow 0} \frac{\mathbb{P}(X \leq t+h) - \mathbb{P}(X \leq t)}{h} = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0. \end{cases} \quad (\text{A.1.1})$$

**Remark A.1.1.** It is straightforward to compute the integral to calculate the cumulative distribution function,

$$F_X(t) = \mathbb{P}(X \leq t) = \begin{cases} 1 - e^{-\lambda t} & t \geq 0 \\ 0 & t \leq 0. \end{cases} \quad (\text{A.1.2})$$

The expectation of  $X$ , denoted  $\mathbb{E}(X)$ , is

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t) dt = \lambda^{-1}. \quad (\text{A.1.3})$$

The exponential distribution has many nice properties. Most notable are the memoryless property and strong memoryless property. However, we first summarize a few other.

**Proposition A.1.1.** (Properties of exponentially distributed random variables). *If  $X_1, X_2, \dots, X_n$  are independent and distributed as an Exponential( $\lambda_j$ ) for  $j = 1, \dots, n$ . Then:*

*i)  $Y = \min\{X_1, X_2, \dots, X_n\}$  is distributed as an Exponential( $\lambda_1 + \lambda_2 + \dots + \lambda_n$ ).*

*ii) Moreover, for  $k = 1, \dots, n$ :*

$$\mathbb{P}(X_k = \min\{X_1, \dots, X_n\}) = \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n}$$

*Proof.* Throughout this problem  $X_1, \dots, X_n$  are as in the statement of the proposition.

**i)** Let  $Y = \min\{X_1, \dots, X_n\}$ . Since each exponentially distributed random variable  $X_i$  has the cumulative distribution function

$$F_{X_i}(x) = \mathbb{P}(X_i \leq x) = 1 - e^{-\lambda_i x},$$

we compute the cumulative distribution function of  $Y$  as follows:

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = 1 - \mathbb{P}(Y \geq t) = 1 - \mathbb{P}(\min\{X_1, X_2, \dots, X_n\} \geq t) \\ &= 1 - \mathbb{P}(X_1 \geq t, X_2 \geq t, \dots, X_n \geq t) = 1 - \prod_{i=1}^n \mathbb{P}(X_i \geq t) = 1 - \prod_{i=1}^n e^{-\lambda_i t} \\ &= 1 - e^{-t \sum_{i=1}^n \lambda_i}. \end{aligned}$$

Independence was used in the second line. Next, we observe

$$f_Y(t) = \frac{d}{dt} F_Y(t) = (\lambda_1 + \dots + \lambda_n) e^{-t(\lambda_1 + \dots + \lambda_n)},$$

as desired.

**ii)** To compute the probability that  $X_k$  equals the minimum, we condition on the value of

$X_k$ .

$$\begin{aligned}
\mathbb{P}(X_k = \min\{X_1, \dots, X_n\}) &= \int_0^\infty \mathbb{P}(X_j > X_k \text{ for } j \neq k | X_k = t) \lambda_k e^{-\lambda_k t} dt \\
&= \int_0^\infty \mathbb{P}(X_j > t \text{ for } j \neq k) \lambda_k e^{-\lambda_k t} dt \\
&= \lambda_k \int_0^\infty e^{-\lambda_k t} \prod_{\substack{j=1 \\ j \neq k}}^n \mathbb{P}(X_j > t) dt \\
&= \lambda_k \int_0^\infty e^{-\lambda_k t} \prod_{\substack{j=1 \\ j \neq k}}^n e^{-\lambda_j t} dt \\
&= \lambda_k \int_0^\infty e^{-t(\lambda_1 + \dots + \lambda_n)} dt = \frac{\lambda_k}{\lambda_1 + \dots + \lambda_n}.
\end{aligned}$$

In the third line we used independence. □

**Proposition A.1.2.** (The memoryless property of exponentially distributed random variables).

A random variable  $X$  has the distribution  $\text{Exponential}(\lambda)$  for some  $\lambda > 0$  if and only if  $X$  has the **memoryless property**, that is for  $a, b > 0$ :

$$\mathbb{P}(X > a + b | X > b) = \mathbb{P}(X > a).$$

Moreover,  $X$  also has the **strong memoryless property**. In other words, for any non-negative random variable  $Y$  that is independent from  $X$

$$\mathbb{P}(X > a + Y | X > Y) = \mathbb{P}(X > a).$$

*Proof.* Since  $F_X(t) = 1 - e^{-\lambda t}$  whenever  $t > 0$ , we observe

$$\mathbb{P}(X > a+b | X > b) = \frac{\mathbb{P}(X > a+b, X > b)}{\mathbb{P}(X > b)} = \frac{\mathbb{P}(X > a+b)}{\mathbb{P}(X > b)} = \frac{e^{-\lambda(a+b)}}{e^{-\lambda b}} = e^{-\lambda a} = \mathbb{P}(X > a).$$

For the other direction we recall the fact that is a fun exercise in analysis. Namely, if



$f : [0, \infty) \rightarrow (0, 1]$  satisfies,  $f(x + y) = f(x)f(y)$ , then  $f(x) = a^x$  for some  $a < 1$ . Moreover, you can find  $\lambda > 0$  so that  $f(x) = e^{-\lambda x}$ .

So, we suppose that  $X$  is a memoryless random variable, and consider the function  $G(a) = \mathbb{P}(X > a)$ . Due to the memoryless property for  $a > 0$ , we have that,

$$G(a) = \mathbb{P}(X > a) = \mathbb{P}(X > a+b|X > b) = \frac{\mathbb{P}[X > a+b, X > b]}{\mathbb{P}[X > b]} = \frac{\mathbb{P}[X > a+b]}{\mathbb{P}[X > b]} = \frac{G(a+b)}{G(b)}.$$

We conclude from the above that,  $G(a + b) = G(a)G(b)$ . Moreover, since  $G(x)$  is a probability, then,  $0 \leq G(x) \leq 1$  for all  $x \geq 0$ . Consequently,  $G(x) = e^{-\lambda x}$  for some  $\lambda$ . Since  $G(x) = 1 - F_X(x)$ , we see that  $X$  is distributed exponential with parameter  $\lambda$  for some positive  $\lambda$ .

To see the strong memoryless property, let  $Y$  be a non-negative random variable independent of  $X$ . Conditioning on the outcome of  $Y$ , we observe,

$$\begin{aligned} \mathbb{P}(X > Y + a|X > Y) &= \int_0^\infty \mathbb{P}(X > Y + a | X > Y, Y = t)f_Y(t) dt \\ &= \int_0^\infty \mathbb{P}(X > t + a | X > t)f_Y(t) dt \\ &= \int_0^\infty \mathbb{P}(X > a)f_Y(t) dt \\ &= \mathbb{P}(X > a) \end{aligned}$$

where we used independence in the second line. □