CONTINUOUS AND DISCRETE ISOPERIMETRIC INEQUALITIES AND GERRYMANDERING

By

EMILIO ISAIAH DEHOYOS

Bachelor of Science in Mathematics
Oklahoma State University
Stillwater, Oklahoma
2019

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2021

CONTINUOUS AND DISCRETE
ISOPERIMETRIC INEQUALITIES
AND GERRYMANDERING

Thesis Approved:

Dr. Edward Richmond

Thesis Advisor

Dr. Lisa Mantini

Dr. Melissa Mills

ACKNOWLEDGMENTS

First and foremost, I would like to thank my thesis advisor, Dr. Edward Richmond for his invaluable advice, continuous support, and patience during my MS study. His immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. I am also extremely grateful Dr. Lisa Mantini and Dr. Melissa Mills for their continuous support throughout my study. I would like to thank all of the faculty in the department of mathematics at Oklahoma State University. It is their kind generosity, help, and support that have made my study a wonderful time. Finally, I would like to express my gratitude to my family, friends, and roommates throughout the years. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my study.

Acknowledgments reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: EMILIO ISAIAH DEHOYOS

Date of Degree: DECEMBER, 2021

Title of Study: CONTINUOUS AND DISCRETE ISOPERIMETRIC INEQUALITIES AND GERRYMANDERING

Major Field: MATHEMATICS

Abstract: The Polsby-Popper score, in political science literature, analyzes compactness of congressional districts in area is compared to perimeter. By observing some issues with this shape analysis scores, we discretize the Polsby-Popper score via notions of graph theory and discuss the relationship of isoperimetric inequalities and vertex compactness in various lattices of the Euclidean plane. We further introduce a way to compare continuous and discrete scores of congressional districts to detect gerrymandering.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

GERRYMANDERING

1.1 Outline

In this chapter, we introduce and explain the idea of gerrymandering as well as the inner workings of gerrymandering adduced from [11] and [1]. Next, we discuss the different ways to detect gerrymandering using the most commonly cited compactness metric in litigation, i.e., the Polsby-Popper score. Lastly, we analyze and explore issues with the Polsby-Popper score and form a new discrete version of the Polsby-Popper score in order to mitigate these issues.

1.2 Introduction

Historically, Americans have been innovators in the design of electoral systems. In the United States, there is a variety of elections that are conducted by partitioning a region into geographically delimited districts and selecting one winner per district via a plurality election. This is the most common and best-known voting method currently in use in America, which is used to elect the members of the U.S. House Representatives as well as many state and local legislatures. Under plurality voting, an area is divided into a number of geographically defined voting districts, each represented by a single elected official. Voters cast a single vote for their district's representative, with the highest total vote-getter winning election, even if they has received less than half of the vote. It is a winner-take-all method for electing the legislature, meaning that each district is represented solely by the party which earned the most votes in the most recent election.

Now, a suitable partition of the region is called a districting plan, and the act of revising it is called redistricting. States may have their own regulations governing the redistricting process, often including specific requirements for valid plans, and the procedures and outcomes are the subject of considerable debate and legal scrutiny. However, there are two main principles commonly applied to the shape (that is, the geometric form) of districts: jurisdictions should be cut into pieces that are "contiguous" and that are "compact." Those evaluating a plan, whether during its construction and approval or during subsequent challenges, need tools for assessment in the context of these two aims, among others.

The first of these guidelines, contiguity, refers to topological connectedness: a district should not be made up of multiple connected components. This is a widespread and uncontroversial requirement of districting plans. In contrast with that clarity, compactness gestures at the reasonableness of district shapes, and is rarely defined precisely, if at all. Even on such unsteady footing, the notion is critical to any discussion of redistricting (and, in par-

ticular, to any discussion of abusive districting practices broadly known as gerrymandering) because compactness appears as an explicit requirement in many states and is nationally recognized as a traditional districting principle. To date, more than thirty possible definitions of compactness as a shape quality metric have been proposed in the political science literature.

The political relevance of requiring districts to be reasonably shaped—and not unnecessarily elongated or twisting—can be defended in several ways. Geometric eccentricity can signal a districting plan that has been engineered to produce an extreme outcome, often by exploiting demographics and geography in order to maximize representation for one group at the expense of another. A mapmaker can tilt outcomes by packing the out-group into a small number of districts, with wastefully high vote share in those districts, and cracking their leftover population by dispersing it, thus diluting those voters' influence. Either strategy can induce distended district shapes in order to achieve its purpose. A second, related argument for shape guidelines is that any limitation placed on districters is a healthy check on their power. But third, and maybe most fundamental, is that being more compact should mean that districts represent chunks of territory that have a meaningful cohesion and can be traveled efficiently.

1.3 Polsby-Popper Score

We begin by introducing the most commonly cited compactness metric in litigation, the Polsby-Popper score, and we discuss its variants and a few alternatives in the legal literature.

Definition 1.3.1 (Polsby and Popper, 1991) *The Polsby-Popper score of a region R is*

$$PP(R) \coloneqq 4\pi \cdot \frac{A(R)}{P(R)^2},$$

where $A(R)$ denotes the area of region R and $P(R)$ denoted the perimeter of region R.

The reason that experts frequently apply this score to voting districts is the idea referenced above, that tentacled or otherwise eccentric districts are red flags for possible gerrymandering. Shapes with skinny necks or long spurs will have much less area than could have been enclosed by the same boundary length around a plumper shape.

Additionally, The Polsby-Popper score is derived from isoperimetric inequality in $\mathbb{R}^2$, which we prove in Chapter 2. In other words, we have that circles have the most area among all shapes with a given perimeter in $\mathbb{R}^2$. Thus, higher Polsby-Popper scores are therefore deemed preferable to lower ones and hence, by the isoperimetric inequality, the score satisfies $0 \leq PP(R) \leq 1$ for all shapes, with $PP(R) = 1$ realized if and only if R is a circle. The squaring of the perimeter in the denominator of the Polsby-Popper score makes the units of measurement cancel out, so that the score is (theoretically) scale-invariant. In other words, $PP(kR) = PP(R)$, meaning that if one were to enlarge an entire region by a factor of k, its Polsby-Popper score would not register the change. Thus this metric is designed to measure something about the shape, and not the size, of a district.

In spite of the way that specialists as often as possible refer to Polsby-Popper scores, there is no agreement on how these scores ought to be utilized while deciding the legitimacy of a districting plan. To make matters seriously befuddling, legal contexts regularly require the revealing of more than one kind of compactness score. Consider the litigation of the recently new driven case of the legislative redistricting in Pennsylvania. In the court orders of January 22 and 26, 2018, it is required that "[A]ny redistricting plan the parties or intervenors choose to submit to the Court for its consideration shall include ... [a] report detailing the compactness of the districts according to each of the following measures: Reock; Schwartzberg; Polsby-Popper; Population Polygon; and Minimum Convex Polygon [1]." In short, each are defined as the following [5].

1. The Reock Score is the ratio of the area of the district AD to the area of a minimum bounding cirle that encloses the district's geometry. A district's Reock score falls within the range of [0,1] and a score closer to 1 indicates a more compact district [15].

2. The Schwartzberg compactness score is the ratio of the perimeter of the district to the circumference of a circle whose area is equal to the area of the district. A district's Schwartzberg score as calculated below falls with the range of [0,1] and a score closer to 1 indicates a more compact district [16].

3. The Population Polygon is the population of a district divided by the population contained in its convex hull [1].

4. The Minimum Convex Polygon score, also known as the Convex Hull score, is a ratio of the area of the district to the area of the minimum convex polygon that can encloses the district's geometry. A district's Convex Hull score falls within the range of [0,1] and a score closer to 1 indicates a more compact district [1].

1.4 Issues with Scores

As we have discussed, the preferred compactness scores are all shape based. Now, we articulate issues with the scores that are inseparably attached to the utilization of Euclidean geometry on map forms summarized in [5].

1. "Coastline effects (i.e. the coastline paradox): Districts with boundaries produced by natural physical features, such as coastlines, are heavily penalized by compactness scores based on the perimeters of contours because the counter-intuitive observation that the coastline of a landmass does not have a well-defined length [5]." In fact, the answer depends on the length of the ruler you use for the measurements. A shorter ruler measures more of the sinuosity of bays and inlets than a larger one, so the estimated length continues to increase as the ruler length decreases.

2. "Resolution instability: The choice of map resolution can have a dramatic impact on contour-based compactness scores, both independently and in contrast with one another, and there is no finest canonical resolution at which information can be accumulated [5]." The Census Bureau itself flags this issue, warning that "[t]hese boundary

files are specifically designed for small scale thematic mapping ... These files should not be used for ... geographic analysis including area or perimeter calculation."

3. "Coordinate dependence: The choice of map projection and coordinate system can force extreme changes on the form based compactness scores of districts, both exclusively and in contrast with one another [5]." This is because all projections of a sphere on a plane necessarily distort the surface in some way and to some extent which was proved by Carl Friedrich Gauss in 1827 under Theorema Egregium. Map projections can be constructed to preserve some of these properties at the expense of others. Because the curved Earth's surface is not isometric to a plane, preservation of shapes inevitably leads to a variable scale and, consequently, non-proportional presentation of areas. Vice versa, an area-preserving projection can not be conformal, resulting in shapes and bearings distorted in most places of the map. Each projection preserves, compromises, or approximates basic metric properties in different ways.

4. "Empty space effects: Although non-populated regions have no impact whatsoever on electoral outcomes, contour-based compactness scores are highly sensitive to their assignment to districts [5]." As of the 2010 census, the United States consists of 11,078,300 Census Blocks. Of them, 4,871,270 blocks totaling 4.61 million square kilometers were reported to have no population living inside them and, hence, compactness can be wildly skewed by assignments of unpopulated surface area to districts.

1.5 Discretizing Compactness

We will lay out some direction below for how to import isoperimetric ideas from discrete geometry into the study of electoral geography. In doing so, we will flag the decision junctures at which the modeler—who may be working as part of the creation and approval process for new districts, or may be working to assess a districting plan after the fact—is making choices that might impact the qualitative features of the scores. We now set out a few specialized definitions that will be helpful below (see [11] and [5]).

Definition 1.5.1 *Given a graph $G = (V, E)$ and a subset $I \subseteq V$, the internal boundary $\partial_0 I$ of the induced subgraph $G[I]$ is the set of vertices in I that have neighbors outside of I; that is,*

$$\partial_0 I \coloneqq \{x \in I : \{x, a\} \in E, a \in V - I\}.$$

Equivalently, the internal boundary $\partial_0 I$ is the set of vertices in I that are incident to edges in the cut-set of G defined by I. The set of edges for which $x \in I$ and $a \in I^c$ where $I^c \coloneqq \{a : a \in V, v \notin I\}$ is a cut-set.

Definition 1.5.2 *A graph with boundary is a graph G = (V,E) together with a subset of vertices $\partial G \subseteq V$ that are designated as boundary vertices. The total boundary of an induced subgraph G[I] is*

$$\partial I \coloneqq \partial_0 I \cup (I \cap \partial G),$$

which is the interior boundary together with any vertices of I that belong to the boundary of the graph G.

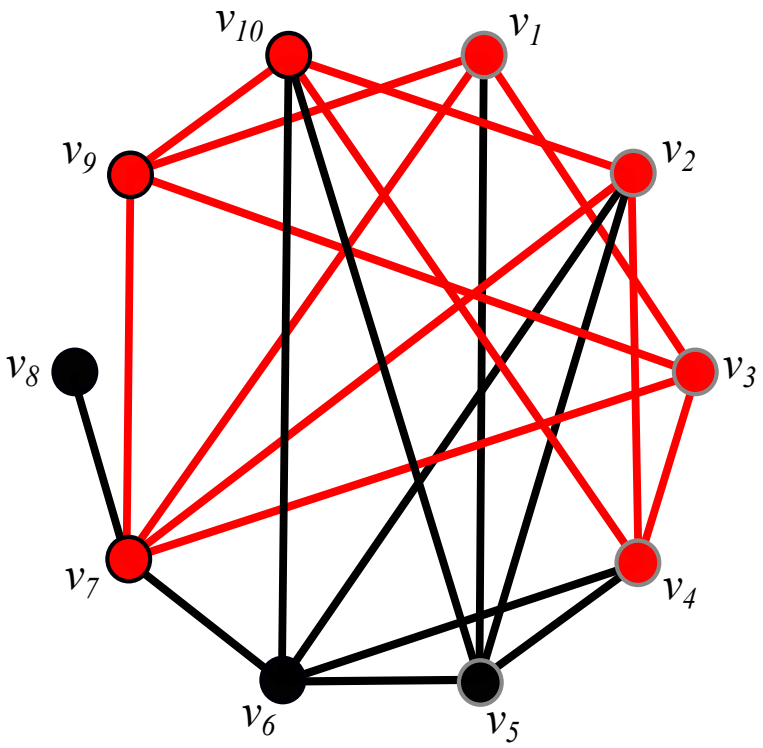

Figure 1: The graph G from the example above. The boundary vertices of the graph, ∂G, are encircled in grey, the vertices in I are encased in red, and the induced subgraph $G[I]$ is colored red.

For example, let $G = (V, E)$ be a graph define its boundary to be $\partial G \coloneqq \{v_1, v_2, v_3, v_4, v_5\}$ (see Figure 1). Consider the subset of vertices $I \coloneqq \{v_1, v_2, v_3, v_4, v_7, v_9, v_{10}\} \subset V$. The internal boundary of $G[I]$ is $\partial I_0 = \{v_1, v_2, v_4, v_7, V_{10}\}$ because each of those vertices is adjacent to at least one element of $I^c = \{v_5, v_6, v_8\}$, meaning that it is incident to an edge in the cut-set. The total boundary of $G[I]$ is

$$\begin{aligned}\partial I &= \{v_1, v_2, v_4, v_7, v_{10}\} \cup (I \cap \{v_1, v_2, v_3, v_4, v_5\} \\ &= \{v_1, v_2, v_3, v_4, v_7, v_{10}\}\end{aligned}$$

A graph may be endowed with a (vertex) weight function, $w : V \longrightarrow \mathbb{R}^+$, assigning nonnegative values to the vertices. Endow G with vertex weights is displayed (see Table 1).

v	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9	v_{10}
$w(v)$	2	7	8	1	5	6	0	2	10	2

Table 1: The graph G in the above example with vertex weights

The principle of discrete compactness scores is to use the combinatorial geometry of these graphs. We start by taking the discrete area to be the order (that is, the number of vertices) of the district subgraph and the discrete perimeter to be the order of its boundary, possibly choosing to weight both of these calculations by population. This immediately suggests two discrete analogs of the Polsby-Popper score of a district. We can let the compactness

be measured by discrete area divided by the square of discrete perimeter; or the same calculation, but weighted by population [5]. That is,

$$\frac{|I|}{(|\partial I|)^2} \tag{1.5.1}$$

and

$$\frac{\sum\limits_{v\in I} w(v)}{\Big(\sum\limits_{v\in \partial I} w(v)\Big)^2}. \tag{1.5.2}$$

Table 1 provides weights of the vertices of G, therefore, one can easily calculate that $\frac{6}{7}$ of the vertices of I are in its total boundary, while the boundary contains only $\frac{2}{3}$ of the vertex weight:

$$\frac{\sum\limits_{v\in I} w(v)}{\Big(\sum\limits_{v\in \partial I} w(v)\Big)^2} = \frac{2+7+8+1+0+2}{2+7+8+1+0+10+2} = \frac{2}{3}.$$

1.6 Motivation for Grid Graphs

It is a basic technique in geometric group theory and algebraic topology to generate combinatorial model spaces out of vertices, faces, and higher-dimensional cells to aid in understanding the geometry of groups and manifolds. The length (L) of a loop in the model space's edges can then be compared to the number (N) of cells in a filling (the faces enclosed by the loop). These quantities L and N are abstract analogs of perimeter and area, respectively, in the context of our prior discussion. Any loop in the square grid, for example, fulfills the inequality $N \leq \frac{1}{16}L^2$, but the triangular grid has $N \leq \frac{1}{6}L^2$. These have been studied since the pioneering work of Dehn in the early twentieth century [11].

The inequalities mentioned can easily be rewritten to relate the order of an induced subgraph to the order of its boundary. The resulting bounds may differ in their coefficients but will have the same exponents. That is, for any subgraph whose boundary ∂I lies on a loop, the inequalities relating N and L can be converted to inequalities relating $|I|$ and $|\partial I|$ that are all of the form

$$|I| \leq c \cdot |\partial I|^k. \tag{1.6.1}$$

Now, observe with equation 1.5.1 and 1.5.2, we can produce scores tending to infinity if we begin with a fixed district I and successively subdivide interior cells while leaving the boundary fixed. This motivates our move to view the discrete area of a district in terms of the vertices of a defining lattice graph, and the discrete perimeter in terms of boundary vertices.

CHAPTER II

ISOPERIMETRIC INEQUALITIES

2.1 Introduction

In this chapter, we first begin by explaining the history as well as idea behind the isoperimetric problem and inequality. Then, we prove the continuous version of the isoperimetric inequality in two-dimensional space. Next, we introduce a discrete version of the isoperimetric inequality in two-dimensional space, which is followed by a constructed proof in the $\mathbb{Z}^2$ lattice graph, that is, the square grid graph. Lastly, we show a consistency proof of a different variation of the discrete isoperimetric problem and inequality in contrasting types of lattice graphs that tile the Euclidean plane.

2.2 A Brief History

Isoperimetric problems are classical and long-established ideas of study in mathematics. In general, these problems ask for sets whose boundary is the smallest for a given volume, i.e., for a given metric space with a notion of volume and boundary, an isoperimetric inequality will give a lower bound on the boundary of a set of fixed volume. Ideally, for any fixed volume, it produces a set of that volume with minimal boundary. The most preeminent isoperimetric inequality states that, in Euclidean space, the unique set of fixed volume with minimal boundary is the Euclidean ball.

A detailed historical description of the classic example of the isoperimetric inequality in ancient time is given in [3]. It dates back to ancient Rome when Publius Vergilius Maro (70 – 19 B.C.) told the story of queen Dido in his epic Aeneid in Book 4: Fate of Queen Dido. Dido, the daughter of the Phoenician king of the 9th century B.C., fled to a haven near Tunis (a city in modern day Tunisia) after the assassination of her husband by her brother. There she asked a local leader for as much land as could be enclosed by bull's hide. After the leader agreed, Dido cut the hide into narrow strips, tied them together and encircled a large tract of land which became the city of Carthage. Dido faced the following mathematical problem, which is also known as the isoperimetric problem: Find among all curves of given length the one which encloses maximal area. Dido found intuitively the right answer.

The Greeks solved this question, by their standards, when Zenodorus (c. 200 – c. 140 BC) proved that a circle has greater area than any polygon with the same perimeter in [4]. Unfortunately, his work was lost. Though we know of his ideas mainly through Pappus (c. 290 – c. 350 AD) in [4]. Pappus introduced the subject of the problem in his Collection (c. 340 AD), Book V, which is considered a literary masterpiece. We quote Pappus's

introduction; "[b]ees, then, know just this fact which is useful to them, that the hexagon is greater than the square and the triangle and will hold more honey for the same expenditure of material in constructing each. But we, claiming a greater share in wisdom than the bees, will investigate a somewhat wider problem, namely that, of all equilateral and equiangular plane figures having an equal perimeter, that which has the greater number of angles is always greater, and the greatest of them all is the circle having its perimeter equal to them [4]."

Next, we come now to the mathematics hero of our story—Jakob Steiner (1796 – 1863) as described in [4]. Steiner invented the technique of symmetrization, which has become a key tool in geometric analysis. The goal is to transform a set such that the volume (i.e. Lebesgue measure) remains unchanged and the perimeter decreases. Steiner began with some geometric constructions that can be easily understood. For example, it can be shown that any closed curve enclosing a region that is not fully convex can be transformed to enclose more area, by "flipping" the concave areas so that they become convex. Furthermore, be shown that any closed curve which is not fully symmetrical can be "tilted" so that it encloses more area. Steiner gave five proofs of the isoperimetric theorem [18]. However, as lovely as they are, they did not represent a rigorous proof of the isoperimetric theorem; that is, all proofs assume the existence of a solution. This is because his strategy was to take a figure that is not a circle and show that its area can be improved. These proofs did not go unpunished (with future banter) as mathematical analyst would smell an existence assumption just around the corner.

This symmetrization together with the calculus of variations enabled Hermann Schwarz (1843 – 1921) to give in 1884 the first rigorous proof of the isoperimetric property of the ball in the class of domains with piecewise analytic boundaries [3, 17]. Ironically, he was also the first who pointed out the lack of an existence theorem in Steiner's proofs. Later, it was solved in the most elegant way—as we will see later—by means of an inequality derived by Hermann Brunn in 1887 (1862 – 1939) and Hermann Minkowski in 1896 (1864 - 1909) for convex sets and then generalized to nonconvex sets by Lazar Aronovich Lyusternik (1899 - 1981) in 1935 [10].

The isoperimetric problem and inequality has influenced the writings of scholars in many different areas because it is so easy to state and understand and yet its solution has been so mathematically challenging for thousands of years with numerous attempts. Though these inequalities are classical problems of study, isoperimetric problems are an active field of research in a number of areas, such as differential geometry, partial differential equations, discrete geometry, probability and graph theory.

2.3 The Isoperimetric Inequality in $\mathbb{R}^n$

Theorem 2.3.1 (Minkowski, 1909) *Given a bounded set $K \subset \mathbb{R}^n$ with surface area $S(K)$ and volume $V(K)$, then*

$$S(K) \geq nV(K)^{\frac{n-1}{n}} V(B)^{\frac{1}{n}},$$

where $B \subset \mathbb{R}^n$ is a unit ball. The equality is strict when K is a ball in $\mathbb{R}^n$.

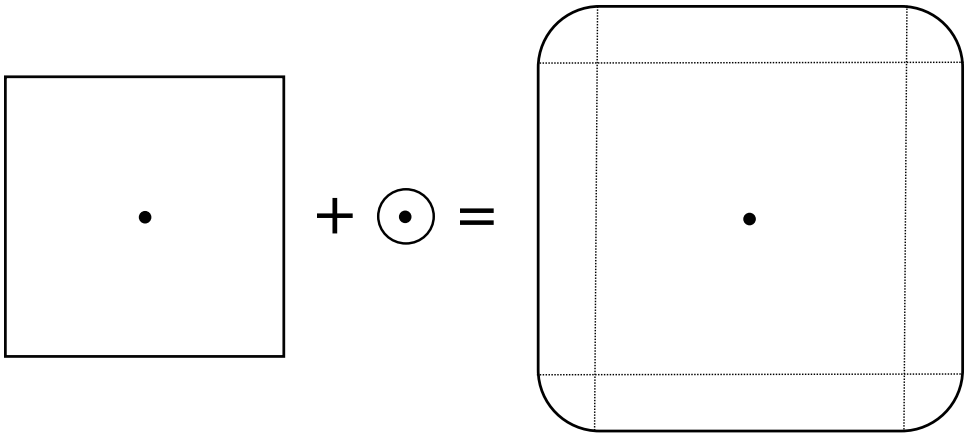

Figure 2: The vector sum of a square and a disk.

Before proving Theorem 2.3.1, some background information is needed to help build to the proof. The vector sum of two bodies, also known as Minkowski addition, is denoted $X+Y = \{x+y : x \in X, y \in Y\}$ of sets X and Y. For example, consider the vector sum of a square K centered at the orgin of side length ℓ and a disk $L = \epsilon B$ of radius ϵ, also centered at the origin (see Figure 2). The vector sum $K + L$ is a rounded square composed of a copy of K, four rectangles of area $\ell\epsilon$, and four quarter-disks of radius ϵ.

The volume $V(K + L)$ of $K + L$ is

$$\begin{aligned} V(K+L) &= V(K) + 4\ell\epsilon + V(L) \\ &\geq V(K) + 2\sqrt{\pi}\ell\epsilon + V(L) \\ &= V(K) + 2\sqrt{V(K)V(L)} + V(L) \\ &= (V(K)^{\frac{1}{2}} + V(L)^{\frac{1}{2}})^2 \end{aligned}$$

which implies that,

$$V(K+L)^{\frac{1}{2}} \geq V(K)^{\frac{1}{2}} + V(L)^{\frac{1}{2}}.$$

In fact, this is the Brunn-Minkowski inequality for $n = 2$.

Lemma 2.3.1 (Minkowski, 1909) *For convex bodies M and N in $\mathbb{R}^n$, then the inequality*

$$V(M+N)^{\frac{1}{n}} \geq V(M)^{\frac{1}{n}} + V(N)^{\frac{1}{n}}$$

is satisfied.

To understand the connection between these two inequalities in Theorem 2.3.1 and Definition 2.3.1, observe that for a square in $\mathbb{R}^2$, namely K, and a unit ball in $\mathbb{R}^2$, namely B, as shown in Figure 1, then

$$V(K+L) = V(K+\epsilon B) = V(K) + 4\ell\epsilon + V(\epsilon B) = V(K) + 4\ell\epsilon + V(B)\epsilon^2.$$

Therefore, we have the following limit:

$$\lim_{\epsilon \to 0^+} \frac{V(K + \epsilon B) - V(K)}{\epsilon} = 4\ell,$$

which is the perimeter of K. This critical observation unlocks the door to the utmost important ingredient of Minkowski's mixed volumes and Brunn-Minkowski theory [12].

Definition 2.3.1 (Steiner, 1842) *The Minkowski–Steiner formula of the surface area $S(M)$ of a convex body M in $\mathbb{R}^n$ is*

$$S(M) = \lim_{\epsilon \to 0^+} \frac{V(M + \epsilon B) - V(M)}{\epsilon}.$$

This definition allows us define the most important part of the proof of Theorem 2.3.1 and ultimately completes our proof.

Proof of Theorem 2.3.1. Now, using l'Hôpital rule, we see that

$$\begin{aligned} S(K) &= \lim_{\epsilon \to 0^+} \frac{V(K + \epsilon B) - V(K)}{\epsilon} \\ &\geq \lim_{\epsilon \to 0^+} \frac{(V(K)^{\frac{1}{n}} + \epsilon V(B)^{\frac{1}{n}})^n - V(K)}{\epsilon} \\ &= nV(K)^{\frac{(n-1)}{n}} V(B)^{\frac{1}{n}}. \end{aligned}$$

■

Surely this alone is good reason for appreciating the Brunn-Minkowski inequality and its efficiency. Additionally, a general case for Theorem 2.3.1 (for $n = 2$) obviously states that among all planar regions with a given perimeter, the circle encloses the greatest area. More specifically, in $\mathbb{R}^2$, the isoperimetric inequality states that for the length L of a closed curve and the area A of the planar region that it encloses, that $L^2 \geq 4\pi A$ and that equality holds if and only if the curve is a circle. Additionally, one should notice that proving Theorem 2.3.1 allows us to see the deep connection in Definition 1.3.1 over the Polsby-Popper score with respect to score being in between $0 \leq PP(R) \leq 1$ for all shapes, with $PP(R) = 1$ realized if and only if R is a circle and the scale-invariant property. These were the motivating factors when coming up with such a mathematical definition in political literature. Additionally, thus also allows us to deduce that equation 1.6.1, is just as in the classical Isoperimetric Theorem, in which $k = 2$ and $c = \frac{1}{4\pi}$.

2.4 The Isoperimetric Inequality in $\mathbb{Z}^2$

Now, we can look at the same problem in a discrete setting. Given a graph $G = (V, E)$ and a finite subset W of vertices of G we define the area of W as the number $|W|$ of vertices in W. For the perimeter of W, there are different notions (see Figure 3). Common notions use the number of neighbor vertices or the number of boundary vertices for the perimeter. It is also reasonable to consider the number of outgoing edges to measure the perimeter. Depending on the graphs we consider and the notion of perimeter, each might define a

different isoperimetric problem. The isoperimetric problem is called (neighbor or boundary) vertex isoperimetric problem if one of the first two notion is used for the perimeter, and edge isoperimetric problem if the last notion is used.

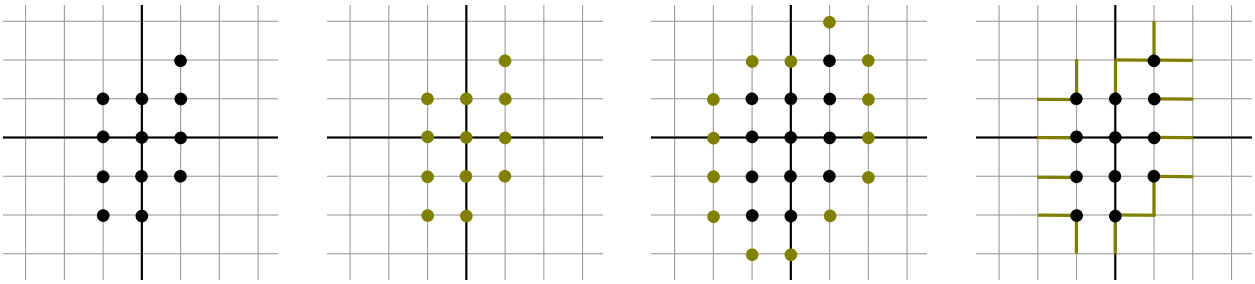

Figure 3: In the left figure, a subset W of vertices in $\mathbb{Z}^2$ is shown, then (from left to right) boundary vertices (as defined in Definition 2.4.1), neighbor vertices and outgoing edges (as defined in Definition 2.5.2 and 2.5.3) are shown as the different measures of perimeter of W.

For example, one of the more broadly known discrete vertex isoperimetric inequalities is Harper's Theorem, which was proven in 1966 by K. H. Harper [9]. Harper's Theorem. This theorem involves the discrete cube $\{0,1\}^n$ and the ℓ_1 metric; that is, the taxicab metric defined by the norm

$$||(x_1, x_2, \ldots, x_n)||_1 = \sum_{i=1}^{n} |x_i|.$$

A graph in this topological space is defined as the vertex set $V = \{0,1\}^n$, where the edge set E consists of points whose distance in the ℓ_1 metric is 1. That is,

$$E = \{(u, v) : u \in \{0,1\}^n, v \in \{0,1\}^n, ||u - v||_1 = 1\}.$$

Here, the vertex boundary of a set $\Gamma \subset \{0,1\}^n$ consists of all points whose distance from Γ in the graph metric is no more than 1. That is,

$$\partial\Gamma = \{v \in V : d(v, \Gamma) \leq 1\}.$$

As usual, the distance between a point x and a set Γ is defined as

$$d(v, \Gamma) = \inf_{\gamma \in \Gamma} d(x, \gamma) = \inf_{\gamma \in \Gamma} \{\text{length of shortest path from } x \text{ to } \gamma\}$$

In other words, the vertex boundary includes both Γ and all of Γ's neighbors. Harper's Theorem demonstrates that a set of minimum vertex boundaries can be attained on an initial segment by defining an ordering on the discrete cube [9]. The most common approach is "compression." The goal is to define a well-ordering on the collection of vertices in general. Then, in terms of this ordering, we can think of an initial segment as being maximum compressed. You create a more compressed set of the same size from any other set of the same size whose vertex border is no larger than the one you started with. With this technique, we can show that the nonunique isoperimetric sets are given by initial segments in the ordering. This technique works when, as is the case here, the isoperimetric sets are

nested [20]. Thus, the sets achieving the minimal boundary are nested. This proof has become the basis in all of the discrete isoperimetric proofs, which we will see.

Now, consider the graph $G = (V, E)$ with vertex set in $\mathbb{Z}^2$. Two vertices $x, y \in \mathbb{Z}^2$ are joined by an edge when the distance in the square metric is 1. In other words, when

$$\rho(x, y) = \max_{i=\{1,2\}} |x_i - y_i| = 1,$$

where $x_i = (x_1, x_2)$ and $y_i = (y_1, y_2)$.

Definition 2.4.1 *We define the boundary of a set $A \subset \mathbb{Z}^2$ by*

$$\partial A = \{x \in \mathbb{Z}^2 : \rho(x, a) = 1, \text{ for some } a \in A\}. \tag{2.4.1}$$

Observe that the notation of the boundary set is defined such that $A \subset \partial A$ and similar to the one used in Harper's Theorem. Now, we wish to define our own version of a well-ordering on the set of vertices in $\mathbb{Z}^2$, which is similar to that used to prove Harper's Theorem.

Definition 2.4.2 *Suppose X is a finite ordered set. An initial segment of size k is the first k elements in X, according to the given well-ordering $<$ with unique smallest element:*

For $k = 1$, the well-ordering $<$ is defined as:

$$0 < 1 < -1 < 2 < -2 < 3 \cdots$$

For $k = 2$, the well-ordering $<$ is defined as follows: for $x = (x_1, x_2) \in \mathbb{Z}^2$ define

$$M(x) \coloneqq \max_{<}\{x_1, x_2\}$$

where the maximum is according to the previously defined well-ordering on $\mathbb{Z}^1$. Then, for $x, y \in \mathbb{Z}^2$ with $x \neq y$, if $M(x) < M(y)$, then $x < y$. If $M(x) = M(y)$, define

$$i_x \coloneqq \min\{i : x_i = M(x)\}.$$

That is, for any $x \in \mathbb{Z}^2$, let i denote the smallest index for which $x_i = M(x)$ and let $i_x \coloneqq i$. Then, if $i_y < i_x$, it follows that $x < y$. Lastly, if $M(x) = M(y)$ and $i_x = i_y$, then define x' as x without $i_x - th$ coordinate; hence, $x' \in \mathbb{Z}^1$. Then, we finally say that $x < y$ when $x' < y'$.

1. $(0, 0)$	2. $(0, 1)$	3. $(1, 0)$
4. $(1, 1)$	5. $(0, -1)$	6. $(1, -1)$
7. $(-1, 0)$	8. $(-1, 1)$	9. $(-1, -1)$
10. $(0, 2)$	11. $(1, 2)$	12. $(-1, 2)$.

Table 2: The first 12 terms for $\mathbb{Z}^2$ in the given well-ordering $<$ are shown.

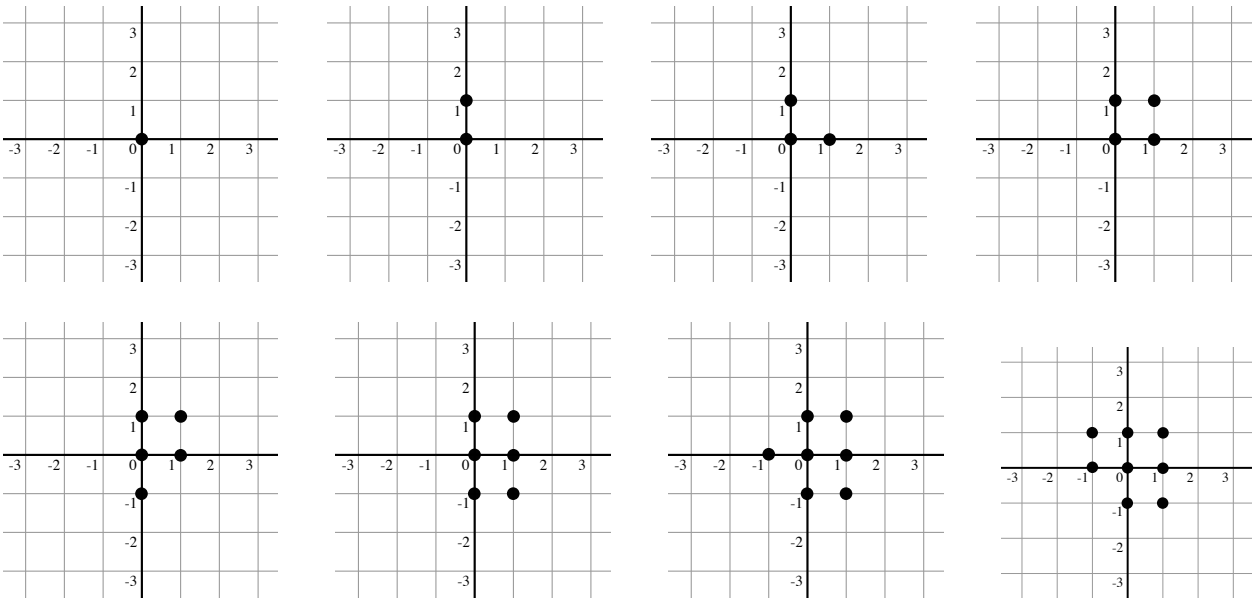

Figure 4: Sets of sizes 1 through 8 of minimal vertex boundary in $\mathbb{Z}^2$.

Theorem 2.4.1 (Veomett and Radcliffe, 2012) *Let I be an initial segment in $\mathbb{Z}^2$, and A a finite, nonempty subset of $\mathbb{Z}^2$. If $|I| = |A|$, then $|\partial I| \leq |\partial A|$.*

Prior to proving Theorem 2.4.1, we must run-through a few points. First, we define a set $A \subset \mathbb{Z}^2$ and prove that the boundary of the initial segment of size $|A|$ is no larger than the boundary of I using a similar technique of compression that originated in Harper's Theorem. Additionally, compression relies heavily on the fact that sets of minimum boundary are nested. However, compression by itself is not enough to prove Theorem 2.4.1, instead we show that compressing a set in $\mathbb{Z}^2$ and a meticulous and particular type of "organization" eventually results in a set of minimal boundary. Note that the sets of minimal boundary are not unique but instead there exists sets of minimal boundaries that are nested [20].

Clearly, sets of size k^2 with minimum vertex boundary in $\mathbb{Z}^2$ are squares of side length k. In addition, if a set has size which is not a perfect square, then the set achieving optimality will be a rectangular box, or a rectangular box with a strip on one side of the box. Optimal sets of sizes 1 through 8 in $\mathbb{Z}^2$ are shown in Figure 4 and note the well-ordering $<$ for $\mathbb{Z}^2$.

Lemma 2.4.1 (Veomett and Radcliffe, 2012) *Let I be an initial segment in $\mathbb{Z}^1$, and A a finite, nonempty subset of $\mathbb{Z}^1$. If $|I| = |A|$, then $|\partial I| \leq |\partial A|$.*

Proof. Let $A \subset \mathbb{Z}^1$ be finite and nonempty. One can verify that

$$|\partial A| = |A| + 2$$

if and only if A is a connected segment of $\mathbb{Z}^1$.

Additionally, A is not a connected segment of $\mathbb{Z}^1$ if and only if

$$|\partial A| > |A| + 2.$$

Observe that every initial segment in $\mathbb{Z}^1$ according to the given well-ordering $<$ of $\mathbb{Z}^1$ is a connected segment by definition. ■

Now, we need one last definition, similarly used in Harper's Theorem [20], and we are ready to prove Theorem 2.4.1

Definition 2.4.3 *For $A \subset \mathbb{Z}^2, i \in \{1,2\}$, and $j \in \mathbb{Z}$, define $A_{1^j} \coloneqq \{x \in \mathbb{Z} : (j, x_2) \in A\}$ and $A_{2^j} \coloneqq \{x \in \mathbb{Z} : (x_1, j) \in A\}$. That is, A_{i^j} is the set of vectors in $\mathbb{Z}$, such that when j is inserted into the respected coordinate then the resulting vector is in A.*

Proof of Theorem 2.4.1. Suppose $A \subset \mathbb{Z}^2$ is finite and nonempty. Hence, there are only finitely many $j \in \mathbb{Z}$ for which A_{1^j} and A_{2^j} are nonempty. In addition, we can observe that

$$\sum_{j \in \mathbb{Z}} |A_{i^j}| = |A|.$$

Define C_i as the i-th compression of A by taking $\mathbb{Z}^1$ dimensions and fixing the i-th coordinate (where $i \in \{1,2\}$) such that $|(C_i)_{i^j}| = |A_{i^j}|$ for each j and $(C_i)_{i^j}$ is an initial segment in $\mathbb{Z}^1$. Now, since

$$\sum_{j \in \mathbb{Z}} |A_{i^j}| = |A|$$

and

$$\sum_{j \in \mathbb{Z}} |(C_i)_{i^j}| = |C_i|,$$

then $|C_i| = |A|$. Observe that the following hold true by definition of boundary,

$$\partial(A)_{i^j} = \partial(A_{i^{j-1}}) \cup \partial(A_{i^j}) \cup \partial(A_{i^{j+1}})$$

and

$$\partial(C_i)_{i^j} = \partial((C_i)_{i^{j-1}}) \cup \partial((C_i)_{i^j}) \cup \partial((C_i)_{i^{j+1}}).$$

Then, from construction of C_i, we have $|(C_i)_{i^j}| = |(A)_{i^j}|$. Additionally, since $(C_i)_{i^j}$ is an initial segment in $\mathbb{Z}^1$ and from Lemma 2.4.1, we have that $|\partial((C_i)_{i^j})| \leq |\partial(A_{i^j})|$ for any $j \in \mathbb{Z}$.

If the union of $\partial(A)_{i^j}$ was disjoint, we would be finished. However, it may not be so we note that $\partial((C_i)_{i^j}), \partial((C_i)_{i^{j-1}}), \partial((C_i)_{i^{j+1}})$ are all initial segments by definition and hence ordered. So, one of the inequalities $|\partial((C_i)_{i^j})| \leq |\partial((A)_{i^j})|$, $|\partial((C_i)_{i^{j-1}})| \leq |\partial((A)_{i^{j-1}})|$, or $|\partial((C_i)_{i^{j+1}})| \leq |\partial((A)_{i^{j+1}})|$ is enough to guarantee and prove that $|\partial((C_i)_{i^j})| \leq |\partial(A)_{i^j}|$.

Now, $j \in \mathbb{Z}$ was arbitrary and then,

$$\sum_{j \in \mathbb{Z}} |(\partial A)_{i^j}| = |\partial A|$$

and

$$\sum_{j \in \mathbb{Z}} |(\partial C_i)_{i^j}| = |\partial C_i|.$$

Therefore, $|\partial C_i| \leq |\partial A|$. It's important to note that since $(C_i)_{i^j}$ is an initial segment in $\mathbb{Z}^1$, then C_i is an initial segment in $\mathbb{Z}^2$, which is I. This shows the synonymous statement that $|\partial I| \leq |\partial A|$.

■

2.4.1 Classes of Non-Isomorphic Compact Graphs

Hence, we have shown that for $i \in \{1, 2\}$, the i-th compression of A is a set of the same size with no larger boundary. However, this does not imply that A is an initial segment. For example, the sets seen in Figure 5 are both in $\mathbb{Z}^2$ of size 10 of minimal boundary, which are clearly compressed in every coordinate, but the right figure follows the well-ordering as the left does not. However, it is important to note that the two graphs shown in Figure 5 are graph isomorphic to one another because the left set is just a counter-clockwise rotation of the right set by 90 degrees.

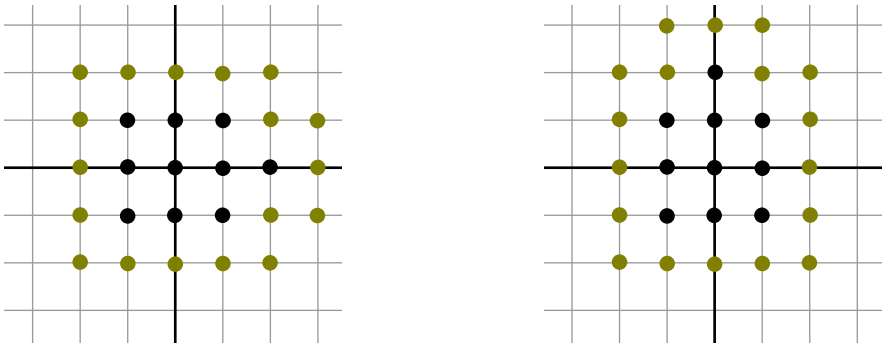

Figure 5: The vertex set on the left is an initial segment, while the set on the right is not. However, both are comprised of minimal boundary and compressed.

Now, this begs the question, can we find sets that are optimal in boundary that are not graph isomorphic? The answer is quite easy to see. In Figure 6, there are two vertex sets, each with 8 interior vertices (with minimal boundary) and compressed in every direction. Observe that one of the sets follows the initial well-ordering but these graphs are not graph isomorphic. This is because within the interior vertices, the right set contains a vertex that would share an edge with every other interior vertex. No such vertex exists in the left set of Figure 6.

Conjecture 2.4.1 *Let M be any positive integer. Then, there exists a graph G in the integer lattice with n interior vertices with minimal boundary and K classes of non-isomorphic compact graphs of G such that $K \geq M$.*

For example, for 10 interior vertices (as seen in Figure 5), observe that for the right set following the initial segment of $\mathbb{Z}^2$, one could slide the top middle vertex over one unit to the left or right and produce a new set that is not graph isomorphic to either sets in Figure 5. Additionally, one could also form a rectangle that is compressed in every direction (similar to the left set of Figure 6) with the 10 vertices. That is, one could take the 10 interior vertices and form a compressed rectangle of length 5 vertices and width 2 vertices, which would still have only 18 boundary vertices. Notice, that the three sets aforementioned containing 10

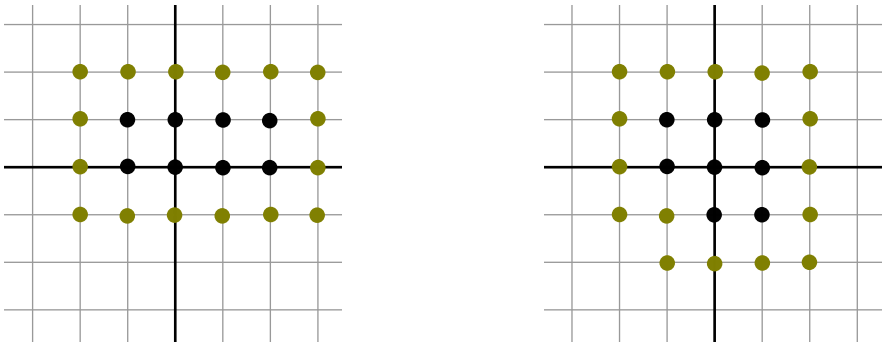

Figure 6: The vertex set on the left is an initial segment, while the set on the right is not. However, both are comprised of minimal boundary and compressed.

interior vertices are graph isomorphic. So, how can we create these compact rectangular sets?

First, let R be a compressed rectangular vertex graph with $n = a \cdot b$ interior vertices where a and b represent the side lengths of the rectangle and $b \leq a$. Now, one must note that in order to form these compressed rectangles of minimal boundary, the sides a and b of the rectangle must "close enough" otherwise the rectangle will not have minimal boundary. We have seen in the example above that if $a = 5$ and $b = 2$, then we get a compressed rectangle with minimal boundary by inspection. However, notice that if $a = 10$ and $b = 1$, then we do *not* get a compressed rectangle with minimal boundary; in fact, it would have 26 boundary vertices. Now, how "close enough" do these numerical values of a and b have to be?

Conjecture 2.4.2 *The upper bound of the ratio of the longer side, a, to the shorter side, b, of R must be in between 2.5 and 3.*

Clearly, the rectangle graph with $a = 5$ vertices and width $b = 2$ vertices (with a total of $n = 10$ vertices) has a ratio of 2.5 and satisfies the additional necessary requirement of being compressed. But, notice that the only set of minimal boundary that is compressed with $n = 3$ interior vertices is shown in Figure 4 (or a variation of this set via rotations and reflections). Clearly, the R with $a = 3$ vertices and $b = 1$ vertex fails to have minimal boundary by inspection. This question of what the exact value of the upper bound of the ratio of a to b of R is left for the reader to discover.

2.5 The Isoperimetric Inequality in the Hexagonal Grid

Recall that a graph is defined as $G = (V, E)$, where V is a (not necessarily finite) set, the vertex set, and E is a binary relation on V, the edge relation. We define the area of A as the cardinality $|A|$ of A and introduce three notions for perimeter as forementioned in Section 2.4. These definitions, which come from [7], are an integral part for proving Theorem 2.5.1.

Definition 2.5.1 *The cardinality of the set of boundary vertices is:*

$$\partial A \coloneqq \{x \in A : \{x, a\} \in E, a \in V - A\}.$$

The reader should note the differences between Definition 2.4.1 and Definition 2.5.1 as we use this above definition of boundary for all of Section 2.5.

Definition 2.5.2 *The number of neighbor vertices, where the set of neighbor vertices of A is the set:*

$$N(A) \coloneqq \{x \in V - A : \{x, a\} \in E, a \in A\}.$$

Definition 2.5.3 *The number of outgoing edges is:*

$$E(A) \coloneqq \{\{x, a\} \in E : x \in A, a \in V - A\}.$$

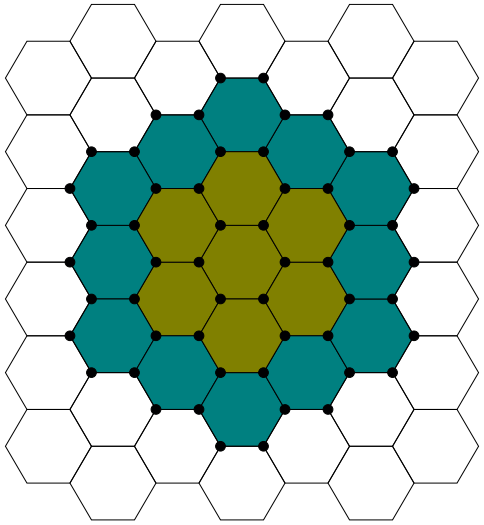

Figure 7: G_2 colored in olive and the surrounding layer of hexagons in teal.

The infinite hexagonal grid is the grid graph formed by a tiling of regular hexagons, in which three regular hexagons meet at each vertex (see Figure 7). We define the finite hexagonal grid of radius r, denoted by G_r, inductively as a subgraph of the infinite hexagonal grid, e.g., for $r = 1$, the hexagonal grid G_1, is a cycle of length 6, depicted by a regular hexagon. The graph G_r consists all vertices and edges of G_{r-1} and the layer of regular hexagons surrounding G_{r-1}. We also denote the infinite hexagonal grid by G_∞.

Theorem 2.5.1 (Grußien, 2012) *Let G_∞ be the infinite hexagonal grid. For any finite subset $A \subset V(G_\infty)$ we have:*

1. $|N(A)| \leq c \cdot \sqrt{|A|}$ *for* $c = \sqrt{6}$, *and c is optimal.*
2. $|E(A)| \leq c \cdot \sqrt{|A|}$ *for* $c = \sqrt{6}$, *and c is optimal.*
3. $|\partial A| \leq c \cdot \sqrt{|A| + \frac{c^2}{4}} - \frac{c^2}{2}$ *for* $c = \sqrt{6}$, *and c is optimal.*

We must first change our set A into a more suitable one with the same cardinality and at most as many neighbor vertices before proving Theorem 2.5.1. After that, we set a lower constraint for the number of neighbor vertices and an upper bound for the number of vertices in A. Finally, we combine these bounds to get Theorem 2.5.1's inequality [6].

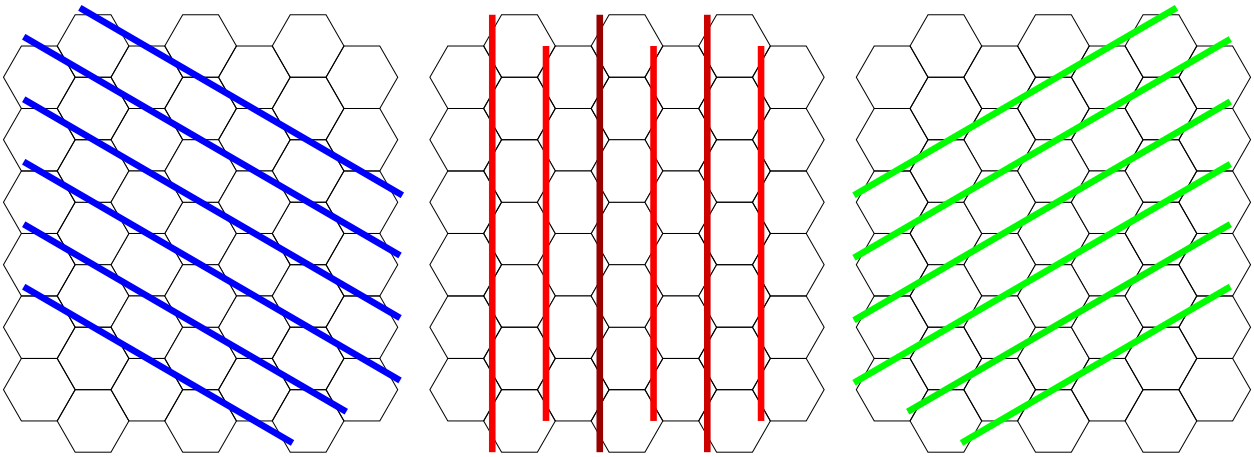

Figure 8: The three distinct directions and their corresponding rows. The first direction is colored blue, the second is colored red, and the third is colored green.

Let's begin with the first section, which entails transforming A. Let us fix three non-zero vectors $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$ in the tiling of regular hexagons, so that there is a vector parallel to each edge of the tiling. We refer to the vectors $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$ as direction vectors, or short directions, and instead of $\vec{v_1}$, $\vec{v_2}$ and $\vec{v_3}$, we refer to direction 1, 2, and 3. By considering the connected components, we can partition the vertices of the grid by moving all edges perpendicular to a direction vector. A connected component like this is referred to as a row. As a result, the infinite hexagonal grid is divided into discontinuous rows of vertices by a direction vector. Furthermore, it is clear that the intersection of two rows of different directions always consists of two adjacent vertices.

Definition 2.5.4 *Let W be a subset of $V(G_\infty)$ that is finite. If a vertex of G_∞ belongs to the set W, it is black with regard to W, and if it does not belong to W, it is white with respect to W.*

Definition 2.5.5 *If no vertex of a row belongs to W, it is white with regard to W; if at least one vertex of a row belongs to W, it is gray with respect to W. If a row of direction i is white and separates two gray rows of direction i, we label it bad.*

We'll just talk about black and white vertices and white and gray rows to keep things simple, and make sure it's apparent which set we're talking about from the context. Consider $A \subset V(G_\infty)$, a finite subset. We can assume that there are no bad rows in any direction with regard to A:

Lemma 2.5.1 *We can find a set A' of the same cardinality with less or equal numbers of neighbor vertices and no bad rows in all directions for the set A* [7].

Proof. Let's have a look at set A. Assume that there is a row that in bad R_i in direction $i \in \{1, 2, 3\}$. The white row then divides the set A into two sets A_1 and A_2, as seen in Figure 9. Because of the hexagonal grid graph's topology, the sets of neighbor vertices $N(A_1)$ and $N(A_2)$ are disjoint. As a result, the vertices of A_2 can be moved closer to the vertices of A_1 without increasing the number of neighbor vertices. This is accomplished by correcting

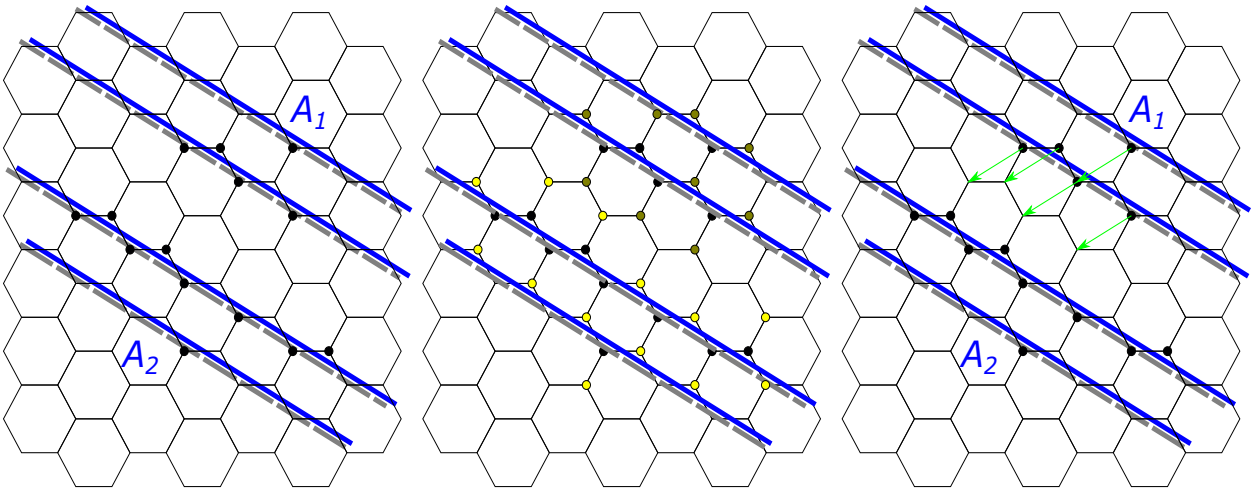

Figure 9: In the left illustration, a white row in direction 1 divides A into A_1 and A_2. In the middle figure, their corresponding neighbor vertices are shown. Finally, movement along direction 3 rows is demonstrated to delete the white row of direction 1.

some $j \in \{1,2,3\} - \{i\}$ and bringing all black vertices $v \in A_2$ along their row of direction j to vertices closer to R_i. When we say we've moved a vertex, we imply we've replaced it with its shifted counterpart. It's worth noting that we don't get any bad rows for direction j this way. In short, we also say we eliminate a bad row of direction i (by moving A_2) agreeable to direction j.

Definition 2.5.6 *Let W be an arbitrary non-empty finite subset of vertices of G_∞. Let R_i^1 and R_i^2 be the two outermost white rows of direction i that have a black vertex in its neighborhood. We define the parallelogram of a subset W to be the only connected component of*

$$G_\infty - \{R_1^1 \cup R_2^1 \cup R_1^2 \cup R_2^2\}$$

that is finite, and denote it by P_W as seen in Figure 10.

We can inductively build the set A' from A. To begin with, we take out every bad row (consistently) of direction 1 agreeable to direction 2, and from that point forward, all bad rows of direction 2 agreeable to direction 1. We get a set A^*, for which, on the off chance that there was somewhere around one bad rown of course 1 or 2, the parallelogram P_{A^*} is a proper subset of the parallelogram of A. Presently, there may be bad row in regards to heading 3. We can take out each such bad row R_3 by moving the vertices of one part of $P_{A^*} - R_3$ pleasant to the course of the rows that contain a longest side of the parallelogram. As a result the parallelogram of the set we get is a subset of P_{A^*}.We can rehash this methodology until we get a set A' with no bad row of bearing 1 or 2. It is not difficult to see that A' has a similar cardinality and at most however many neighbor vertices as A, and as we eliminated all bad rows of course 3 in the last advance, the set A' has no bad row. ■

In the following we assume that A is such that there are no bad rows for all directions:

First, we'll establish a lower bound for $|N(A)|$. In direction i, let l_i be the number of gray rows. Let $l_3 \geq \max\{l_1, l_2\}$, without losing generality. Within each gray row, A has

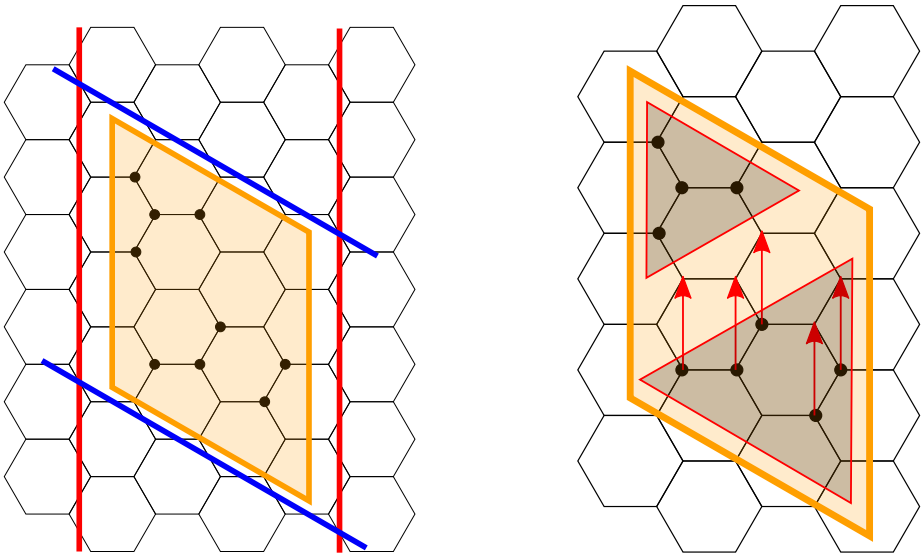

Figure 10: The left graphic depicts a parallelogram of a subset W. The right figure shows the elimination of a bad row in direction 3 by transferring the vertices of one set of agreeable to direction 2.

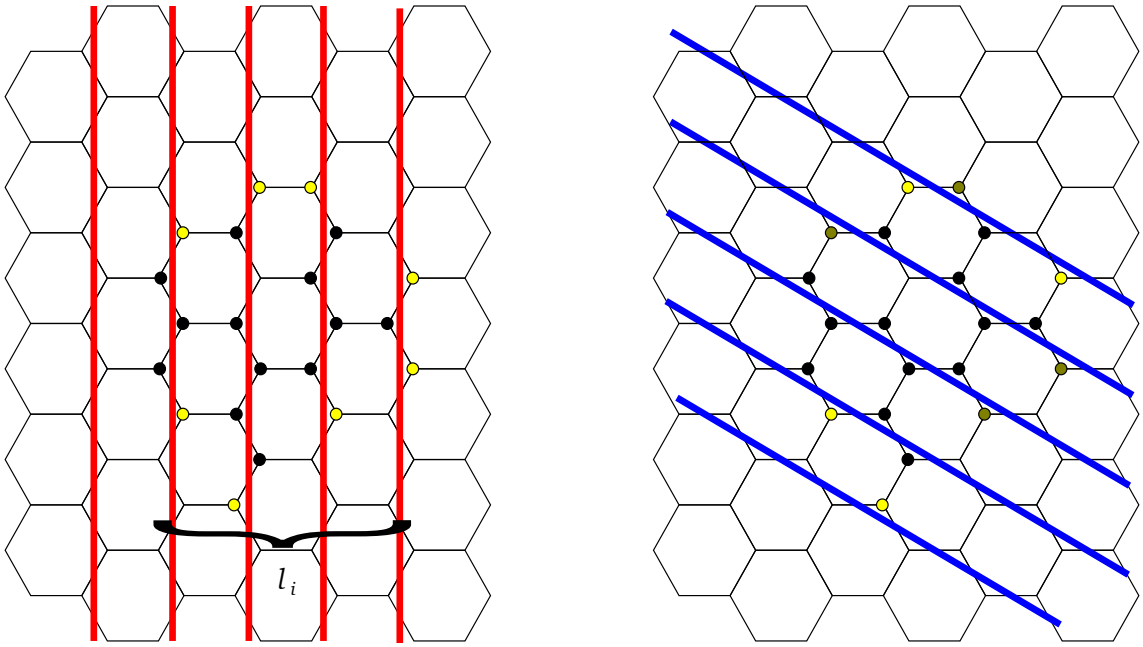

Figure 11: In the left figure, the outermost neighbors for direction 2 are shown in yellow. Then, in the right figure, the outermost neighbors for direction 2 and rows of direction 1 are shown; the olive colored neighbors are also outermost neighbors for direction 1.

two outermost neighbors in any direction. As a result, direction i determines the $2l_i$ farthest neighbors. Figure 11 depicts all of the direction 2's outermost neighbors. Of course, the gray rows' outermost neighbors are neighbor vertices in each direction, and so belong to $N(A)$. As a result, the number of outermost neighbors in each direction is a lower bound on $|N(A)|$. We evaluate all three directions in order to raise this lower limitation. If we take a look at

the outermost neighbors for direction i, we observe that some of these neighbors might also occur as outermost neighbors for other directions [6].

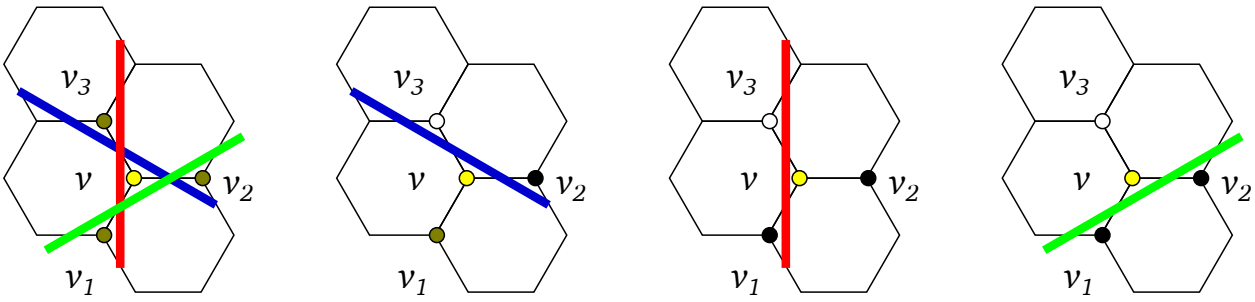

Figure 12: Counting the number of directions in which vertex v is the farthest neighbor.

Lemma 2.5.2 *For at most two directions, every neighbor is an outermost neighbor* [7].

Proof. For three directions, assume vertex v is the outermost neighbor and refer to Figure 12. This will be a proof via contradiction so let v_1, v_2, and v_3 be v's neighbors, with the vertices v_2 and v_3 in direction 1, v_1 and v_3 in direction 2, and v_1 and v_2 in direction 3 occurring in the same row as v. Due to the fact that v is an outermost neighbor in direction 1, either v_2 is a black vertex and v_3 is white, or vice versa. We can assume v_2 is a black vertex and v_3 is a white vertex because G_∞ is symmetric. Because v_3 is white and v is an outermost neighbor for direction 2, v_1 must be a black vertex. But then v is no outermost neighbor for direction 3, which is a contradiction. ■

Applying Lemma 2.5.2, we obtain the following lower bound for the number of vertices in $N(A)$:

$$\begin{aligned} |N(A)| &\geq \frac{1}{2}(2l_1 + 2l_2 + 2l_3) \\ &= l_1 + l_2 + l_3. \end{aligned} \tag{2.5.1}$$

In terms of l_1, l_2, and l_3, we want to find an upper bound for the number of vertices in A. For example, a possible upper bound on the number of vertices in A is $2l_1l_2$. The number of vertices in a parallelogram with l_1 rows in direction 1 and l_2 rows in direction 2 equals the number of vertices. (For directions 1 and 2, remember that there are no white rows between gray rows.) However, l_3 is not taken into account in this bound. Knowing the number of gray rows in the third direction, we can rule out part of the parallelogram's third-direction rows. Since the outermost rows contain the least vertices, we exclude them for an upper bound, and as the number of rows of direction 3 in the parallelogram are $l_1 + l_2$, we exclude $l_1 + l_2 - l_3$ outermost rows.

Because $l_3 \geq \max\{l_1, l_2\}$, we know that $l_1 + l_2 - l_3 \leq \min\{l_1, l_2\}$. Thus, if k rows are removed from one side, $k \leq \min\{l_1, l_2\}$ and $\sum_{i=1}^{k} 2i-1 = k^2$ vertices are removed (see Figure 14). We must now distinguish between two scenarios: If $l_1 + l_2 - l_3$ is even, we can delete $2(\frac{1}{2}(l_1 + l_2 - l_3))^2$ vertices from the parallelogram by excluding $2(\frac{1}{2}(l_1 + l_2 - l_3))^2$ rows from

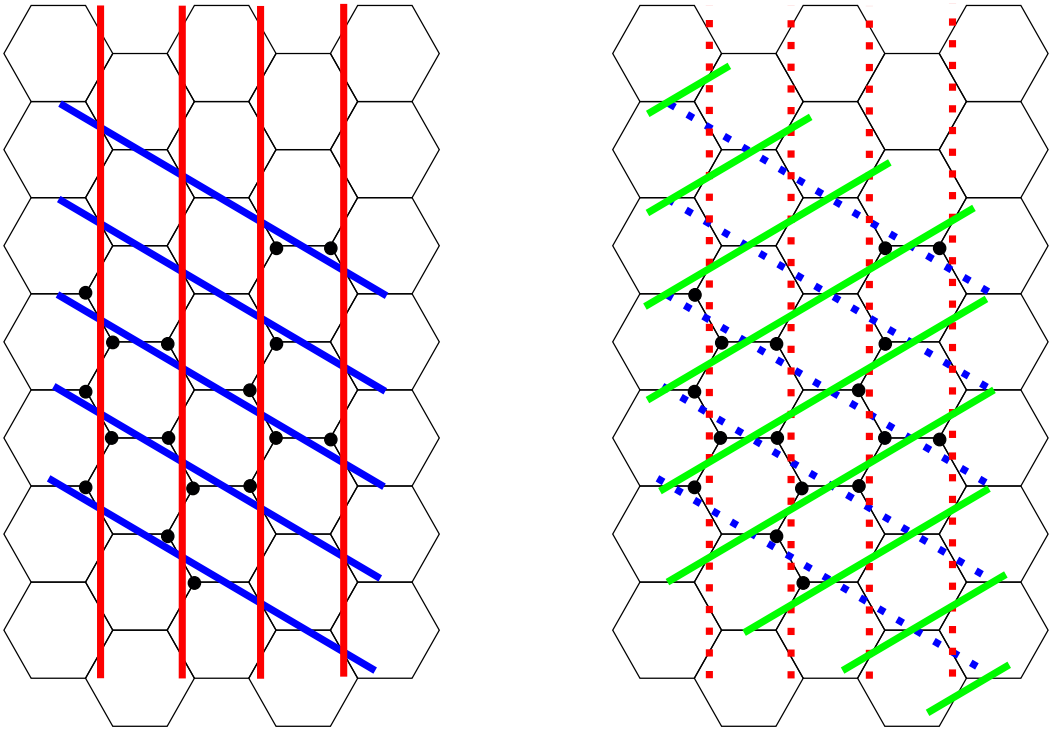

Figure 13: The left picture displays an upper bound on the number of vertices in A using a parallelogram, while the right figure shows l_3 and the possibility of excluding some of the parallelogram's third direction rows to establish an upper bound.

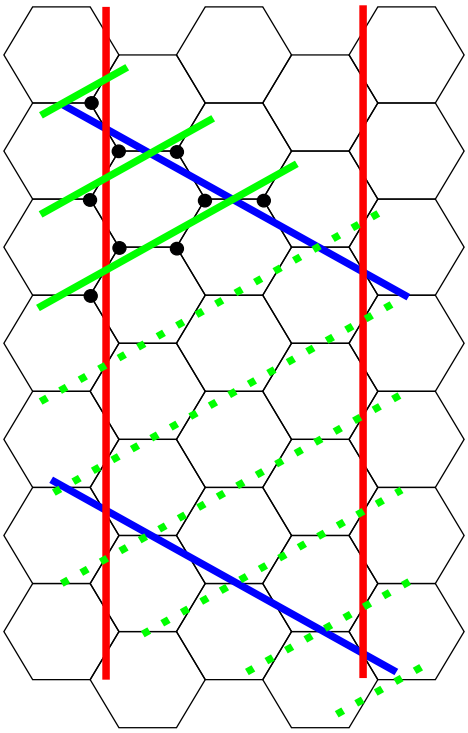

Figure 14: Counting the number of vertices in a parallelogram and comparing to the number of rows removed.

each side. If $l_1 + l_2 - l_3$ is odd, the best result is obtained by eliminating $\frac{1}{2}(l_1 + l_2 - l_3 - 1)$ rows on one side and $\frac{1}{2}(l_1 + l_2 - l_3 + 1)$ rows on the other. Hence, we can bound the number of vertices that we can remove by

$$(\frac{1}{2}(l_1 + l_2 - l_3 - 1))^2 + (\frac{1}{2}(l_1 + l_2 - l_3 + 1))^2 = \frac{1}{2}(l_1 + l_2 - l_3)^2 + \frac{1}{2}.$$

Therefore, we can remove at least $\lceil \frac{1}{2}(l_1 + l_2 - l_3)^2 \rceil$ vertices from the parallelogram, and can give a better upper bound on the number of vertices in A:

$$\begin{aligned} |A| &\leq 2l_1l_2 - \frac{1}{2}(l_1 + l_2 - l_3)^2 \\ &= -\frac{1}{2}(l_1^2 + l_2^2 + l_3^2) + (l_1l_2 + l_1l_3 + l_2l_3). \end{aligned} \tag{2.5.2}$$

Now, we can prove Theorem 2.5.1.

Proof of Theorem 2.5.1 [7]. We use the lower bound on $|N(A)|$ and the upper bound on $|A|$ to show that the first inequality of Theorem 2.5.1 holds:

$$\begin{aligned} |N(A)|^2 &\geq (l_1 + l_2 + l_3)^2 \\ &\geq 6 \cdot (-\frac{1}{2}(l_1^2 + l_2^2 + l_3^2) + (l_1l_2 + l_1l_3 + l_2l_3) \\ &\geq 6 \cdot |A| \end{aligned}$$

Equations (2.5.1) and (2.5.2) show the first and third inequality, respectively. A simple calculation can be used to verify the second inequality. We use an example to demonstrate how tight the inequality is. Consider the subset $A \coloneqq V(G_r)$ of vertices of the infinite grid G_∞ for an arbitrary $r \in \mathbb{N}$. The inequality's tightness arises immediately from the fact that if G_∞ is an infinite hexagonal grid and the subgraph G_r is a finite hexagonal grid of radius r, with a simple induction proof showing $|N(V(G_r))| = 6r, |E(V(G_r))| = 6r$, and $|V(G_r)| = 6r^2$.

We can get the same lower bound for the number of outgoing edges as we can for the number of neighbor vertices for the second inequality if we use outermost outgoing edges instead of outermost neighbor vertices. Since $|N(V(G_r))| = |E(V(G_r))|$, the set $A \coloneqq V(G_r)$ is also an example for the inequality of Theorem 2.5.1 being tight.

Now, for the last inequality of Theorem 2.5.1, let us consider a finite subset $A \subset G_\infty$. Let X be the set $A - \partial A$. Then, the neighbor vertices $N(X)$ are a subset of ∂A. Thereofore, we have $|N(X)| \leq |\partial A|$. Using this inequality and the fact that we just proved $|N(A)| \geq c \cdot \sqrt{|A|}$ for $c = \sqrt{6}$, then we have,

$$|\partial A| \geq |N(X)| \geq c \cdot \sqrt{|X|} = c \cdot \sqrt{|A| - |\partial A|}. \tag{2.5.3}$$

Therefore, an easy calculation shows that the inequality can be translated into the following:

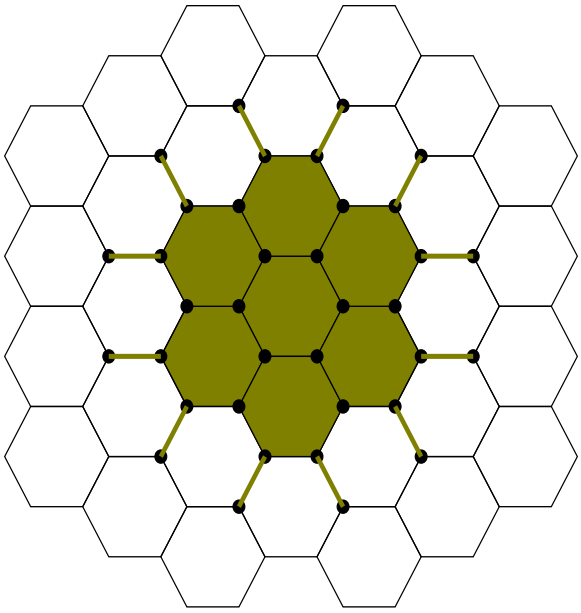

Figure 15: The set $V(G_r) \cup N(V(G_r))$ for $r = 2$, for which the isoperimetric inequality is tight.

$$|\partial A| \geq c \cdot \sqrt{|A| - |\partial A|}$$

$$|\partial A|^2 \geq c^2 \cdot (|A| - |\partial A|)$$

$$\left(|\partial A| + \frac{c^2}{2}\right)^2 \geq c^2 \cdot |A| + \frac{c^4}{4}$$

$$|\partial A| + \frac{c^2}{2} \geq c \cdot \sqrt{|A| + \frac{c^2}{4}}$$

$$|\partial A| \geq c \cdot \sqrt{|A| + \frac{c^2}{4}} - \frac{c^2}{2}.$$

We'll use another example to demonstrate how narrow the gap is: consider the subset $A \coloneqq V(G_r) \cup N(V(G_r))$ of vertices of the infinite grid G_∞ (as seen in Figure 15) for an arbitrary $r \in \mathbb{N}$. Because $|N(X)| = |B(A)|$ in this example, and the inequality in the first part of the proof is tight for the subset $X = V(G_r)$, the inequality in (2.5.3) is also tight, completing the proof of Theorem 2.5.1.

■

2.6 The Isoperimetric Inequality for Triangles

Now, we wish to prove the continuous isoperimetric inequality for triangles.

Theorem 2.6.1 (Svrtan and Veljan, 2012) *Among all triangles of given perimeter, the equilateral one has the largest area. In terms, of in terms of perimeter p and area T, the continuous isoperimetric inequality for triangles states that*

$$12\sqrt{3} \cdot T \leq p^2.$$

Proof. The proof is based on Heron's formula, which states that

$$T^2 = s(s-a)(s-b)(s-c),$$

where a, b, c are the sides of a triangle, and $s = \frac{(a+b+c)}{2} = \frac{p}{2}$, the semiperimeter. Since s is constant, the question is to maximize $\frac{T^2}{s} = (s-a)(s-b)(s-c)$ for $a+b+c = 2s$. We will use the arithmetic and geometric means inequality, which states for any list of n nonnegative real numbers $x_1, x_2, \ldots, x_n$, then we have

$$\sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} \leq \frac{x_1 + x_2 + \ldots + x_n}{2}$$

and that equality holds if and only if $x_1 = x_2 = \ldots = x_n$.
Now, observe that,

$$(s-a) + (s-b) + (s-c) = 3s - (a+b+c) = 3s - 2s = s,$$

and we can deduce that,

$$\frac{T^2}{s} = (s-a)(s-b)(s-c) \leq \left(\frac{s}{3}\right)^3.$$

Finally, we have that $T^2 \leq \frac{s^4}{27}$ with equality only when $(s-a) = (s-b) = (s-c)$, i.e., $a = b = c$. Therefore,

$$T^2 \leq \frac{s^4}{27}$$

$$T \leq \frac{s^2}{3\sqrt{3}}$$

$$3\sqrt{3} \cdot T \leq s^2$$

$$12\sqrt{3} \cdot T \leq p^2,$$

since $s = \frac{p^2}{2^2}$. To sum up, the area of a triangle with perimeter $2s$ never exceeds, $\frac{s^2}{3^{\frac{3}{2}}}$; as is easy to verify, it exactly equals $\frac{s^2}{3^{\frac{3}{2}}}$ for an equilateral triangle with side $\frac{2s}{3}$, since the area of an equilateral triangle is

$$A_T = \frac{\sqrt{3}}{4} \cdot {a_p}^2$$

where A_T and a_p represent the area and side length of an equilateral triangle respectively. ■

Clearly, a similar statement admits an equivalent formulation: Among all triangles with given area, the equilateral one has the least perimeter [19].

This proof can help us deduce the following theorem, which is proved in [4]:

Theorem 2.6.2 (Blåsjö, 2014) *A regular n-gon has greater area than all other n-gons with the same perimeter.*

Likewise, it has been proven that hexagons in the triangular lattice have maximal volume among all sets of a given boundary in any triangulation with minimal degree 6. We state them, respectively, but prove only the continuous version. The reader should compare the continuous and discrete triangular isoperimetric inequality and note the similarities and differences between Section 2.6 and Section 2.5.

Theorem 2.6.3 (Angel, Benjamini, and Horesh, 2018) *The discrete isopetimetric inequality states that for any disc triangulation with V vertices and n boundary vertices as defined in section 1.5, and with all internal degrees at least 6 has*

$$V \leq \lfloor \frac{(n+3)^2}{12} \rfloor.$$

Equality is achieved in the Euclidean triangular lattice.

This theorem is fundamental in the fact that it is a discrete isoperimetric inequality without the lattices as seen in section 2.4 and 2.5.

2.7 The Isoperimetric Inequality for Congressional Districts

Now, we want to see if there exists a discrete isoperimetric inequality for the congressional districts. We first take a look at Oklahoma's congressional districts of the $113th$ (see Figure 16). We compare and contrasts the discrete scores with the continuous scores using the Polsby-Popper method. Note that when using the discrete version, each vertex represents a voting precinct and the edges represent voting precincts sharing a perimeter boundary with another voting precinct, that is, "touching it." The boundary of the congressional district in the discrete score is all of the voting precincts in a districts sharing an edge with another congressional district. However, this boundary notion can change from state to state depending on the calculations needing to observe.

District	Discrete Score	Continuous Score
$1st$	0.06640625000	0.2069
$2nd$	0.04137436224	0.2703
$3rd$	0.03389591130	0.2498
$4th$	0.05273586343	0.2481
$5th$	0.06928697404	0.2678

Table 3: The various scores from Oklahoma's 5 congressional districts of the $113th$ congress.

It's important to note the ranking from worst to best vary when comparing the discrete score and the continuous scores. However, the discrete scores seem to be relatively close in numerical values when looking from a linear scale since the scores can essentially become as large or as small as one chooses, as aforementioned in Chapter 1. However, each score can tell us vital information about each district depending on the state and whether or not it is truly gerrymandered.

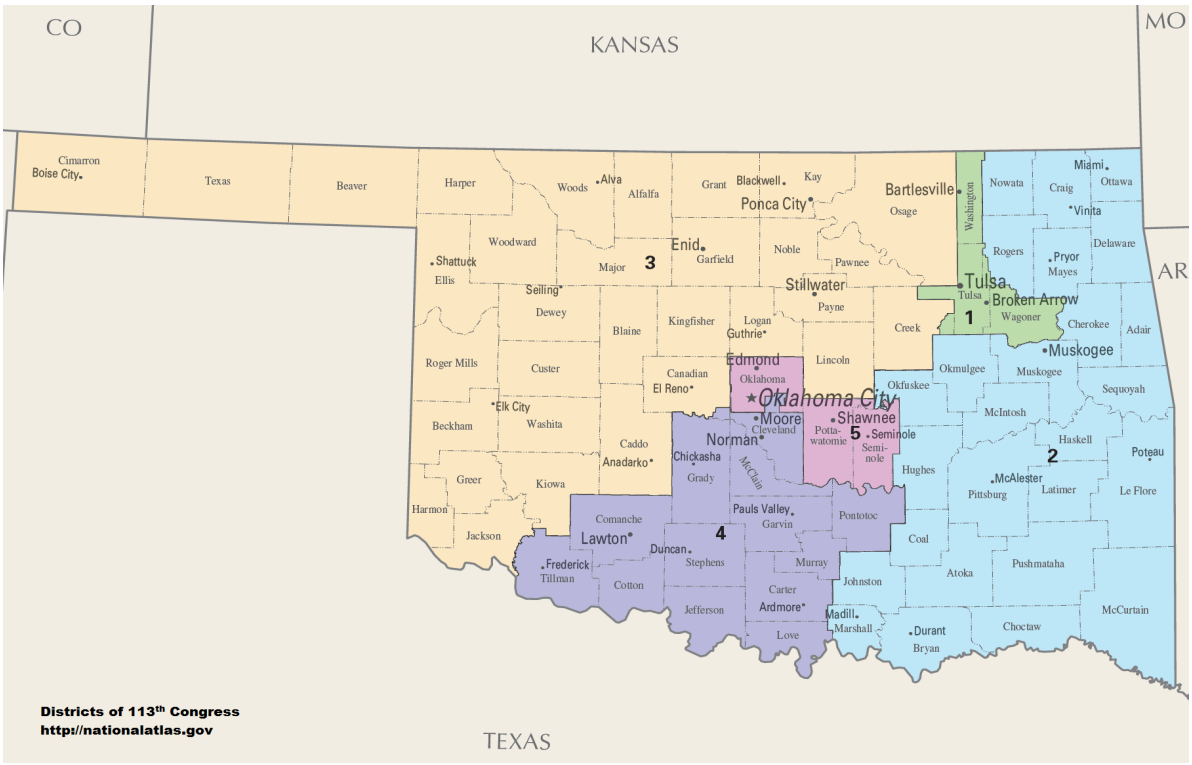

Figure 16: Oklahoma's congressional districts of the $113th$ congress since 2013 [13].

2.8 Further Questions

We now end the chapter with some posing questions.

1. How would the isoperimetric inequality change if weights were added on various vertices in different types of lattices? Hypothetically, this question could relate to gerrymandering in the sense that each weight could represent the voting precincts or population of various cities. Additionally, how would the isoperimetric inequality change if we added different stipulation when defining boundary or edges? For example, one could connect edges diagonally in the square lattice and therefore changing the most optimal shape.

2. What does the isoperimetric inequality look like in semiregular tilings, k-uniform tilings, or non-edge-to-edge tilings of the Eulcidean plane with convex regular polygons? Observe that the lattices presented in this paper have been regular tilings of convex regular polygons: six equilateral triangles, four squares or three regular hexagons at a vertex, yielding the three regular tessellations.

3. What would three-dimensional gerrymandering look like and how could we relate that to the isoperimetric inequality? Clearly, the isoperimetric inequality in section 2.4 could be extended (with the right well-ordering) to three-dimensional space, but what exactly does that mean for gerrymandering? Could extraterrestrial life be gerrymandering galaxies?

REFERENCES

[1] Marie Albiges, *Are pa.'s state house and senate maps gerrymandered? depends on how you measure them*, WHYY (2021).

[2] Omer Angel, Itai Benjamini, and Nizan Horesh, *An isoperimetric inequality for planar triangulations*, Discrete Comput. Geom. **59** (2018), no. 4, 802–809. MR 3802305

[3] Catherine Bandle, *Dido's problem and its impact on modern mathematics*, Notices Amer. Math. Soc. **64** (2017), no. 9, 980–984. MR 3699773

[4] Viktor Bl åsjö, *The isoperimetric problem*, Amer. Math. Monthly **112** (2005), no. 6, 526–566. MR 2142606

[5] Moon Duchin and Bridget Eileen Tenner, *Discrete geometry for electoral geography*, 2018.

[6] Sylvain Gravier, *Tilings and isoperimetrical shapes. II. Hexagonal lattice*, Acta Univ. Palack. Olomuc. Fac. Rerum Natur. Math. **40** (2001), 79–92. MR 1904687

[7] Berit Grußien, *Isoperimetric inequalities on hexagonal grids*, 2012.

[8] T. C. Hales, *The honeycomb conjecture*, Discrete Comput. Geom. **25** (2001), no. 1, 1–22. MR 1797293

[9] H. H. Harper, *Optimal numberings and isoperimetric problems on graphs."*, J. Combinatorial Theory, **1** (1966), 385–393.

[10] Pedro José Herrero Piñeyro, *History of the use of elementary geometry in solving the classical isoperimetric problem*, Gac. R. Soc. Mat. Esp. **15** (2012), no. 2, 335–354. MR 2977072

[11] Takehiro Ito, Naoyuki Kamiyama, Yusuke Kobayashi, and Yoshio Okamoto, *Algorithms for gerrymandering over graphs*, Theoret. Comput. Sci. **868** (2021), 30–45. MR 4246209

[12] H. Minkowski, *Ausgewählte Arbeiten zur Zahlentheorie und zur Geometrie*, Teubner-Archiv zur Mathematik, vol. 12, BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1989, Mit D. Hilberts Gedächtnisrede auf H. Minkowski, Göttingen 1909. Edited and with a foreword and appendices by E. Krätzel and B. Weissbach. MR 1064523

[13] The National Atlas of the United States of America, *Congressional districts - 113th congress*, February 22, 2014.

[14] Daniel Polsby and Robert Popper, *The third criterion: Compactness as a procedural safeguard against partisan gerrymandering*, Yale L. and Pol'y Rev. **9** (1991).

[15] E. C. Reock, *A note: measuring compactness as a requirement of legislative apportionment.*, Midwest Journal of Political Science **5** (1961), 70–74.

[16] J. E. Schwartzberg, *Reapportionment, gerrymanders, and the notion of "compactness."*, Minnesota Law Reviwe **59** (1966), 443–452.

[17] H. A. Schwarz, *Gesammelte mathematische Abhandlungen. Band I, II*, Chelsea Publishing Co., Bronx, N.Y., 1972, Nachdruck in einem Band der Auflage von 1890. MR 0392470

[18] J. Steiner, *Sur le maximum et le minimum des figures dans le plan, sur la sphère et dans l'espace en général. Premier mémoire*, J. Reine Angew. Math. **24** (1842), 93–152. MR 1578318

[19] Dragutin Svrtan and Darko Veljan, *Non-Euclidean versions of some classical triangle inequalities*, Forum Geom. **12** (2012), 197–209. MR 2955631

[20] Ellen Veomett and A. J. Radcliffe, *Vertex isoperimetric inequalities for a family of graphs on $\mathbb{Z}^k$*, Electron. J. Combin. **19** (2012), no. 2, Paper 45, 18. MR 2946103

VITA

Emilio Isaiah DeHoyos

Candidate for the Degree of

Master of Science

Thesis: CONTINUOUS AND DISCRETE ISOPERIMETRIC INEQUALITIES AND GERRYMANDERING

Major Field: Mathematics

Biographical:

Education:

Completed the requirements for the Master of Science in Mathematics at Oklahoma State University, Stillwater, Oklahoma in December, 2021.

Completed the requirements for the Bachelor of Science in Mathematics at Oklahoma State University, Stillwater, Oklahoma in 2019.

Experience:

Employed by Oklahoma State Univeristy in the position of Learning Assisstant for College Algebra and Preparation for Calculus in Stillwater, Oklahoma from August 2018 to May 2019.

Employed by Oklahoma State Univeristy in the position of Tutor at the Mathematics Learning Success Center in Stillwater, Oklahoma from January 2019 to May 2019.

Employed by Oklahoma State Univeristy in the position of Teaching Assistant in Stillwater, Oklahoma from August 2019 to present (Fall 2021).

Professional Memberships:

Member of Pi Mu Epsilon as of May, 2019.