Spring 2018

Gerrymandering and the Impossibility of Fair Districting Systems

Danielle Degutz
Bard College

Gerrymandering and the Impossibility of Fair Districting Systems

A Senior Project submitted to
The Division of Science, Mathematics, and Computing
of
Bard College

by
Danielle Degutz

Annandale-on-Hudson, New York
May, 2019

Abstract

Voting district boundaries are often manipulated, or gerrymandered, by politicians in order to give one group of voters an unfair advantage over another during elections. To make sure a system of voting districts is not gerrymandered, the population size, the shape, and the voting efficiency of each party in each district should be taken into consideration. Following recent work of Boris Alexeev and Dustin G. Mixon, we discuss mathematical criteria for each of these three aspects, and we prove how problems arise when attempting to apply all three at once to a districting system–first to a simplified districting system and then to a more realistic districting system.

Contents

Acknowledgments

Thank you to Ethan Bloch, my senior project advisor, for being very helpful and telling me not to worry. Thank you to Lauren Rose, my academic advisor, for making me feel welcome in the math department. Thank you to Mom and Dad for making me go to college and for paying my tuition. Thank you to Grandpa for encouraging me to study mathematics. Thank you to myself for doing all this work.

1
Background Information

1.1 Introduction

Gerrymandering is a way of dividing a geographic area into voting districts, that is, into a districting system, in a way that unfairly favors one political group. Policymakers are usually the people given the task of redrawing district boundaries and they often do so in a way that gives more voting power to their party or takes away voting power from minority groups. To learn more about gerrymandering, see [6].

One strategy used is called packing. When packing is used, many voters from one group are packed into only a small amount of districts so that they cannot affect the outcome of any other districts and will win only the ones they are packed into. One tell-tale sign of packing is when district boundaries are irregularly shaped, so that they include all voters from the same group even when they live in very different areas.

The opposite strategy, cracking, occurs when the voters from one group are divided up, putting a small amount of them in each district. This ensures that they don't have enough people in any district to win it. To determine if cracking is being used, we can calculate the Efficiency Gap, which was created by Nicholas Stephanopoulos and Eric McGhee in [5].

Mathematicians are working to come up with a standard to be used to determine if a districting system is fair or gerrymandered. This is difficult and they are encountering many problems in the process. It is controversial which factors can determine district fairness and which cannot. There have been many court cases on the issue.

Gerrymandering is a difficult mathematical problem because it involves looking at both geometry of districts and the arrangement and amount of voters for each party. In real life cases, one must take into consideration natural barriers that are often irregularly shaped, such as coasts, mountains, and rivers, and human barriers such as infrastructure, city boundaries, and personal property lines. These provide geometrical challenges for creating districts. Additionally, residents of any place do not live in orderly rows, but rather scattered throughout, densely packed in cities and sparsely sprinkled throughout rural farmland, which create challenges for providing fairly represented districts. To complicate the matter further, oftentimes those who belong to the same political party or racial minority live close to each other in one area, rather than spread evenly throughout the land, making it more difficult to provide them with fair representation.

This project will work from [1], and takes ideas from previous work done by Moon Duchin and Mira Bernstein in [2]. In the paper [1], Alexeev and Mixon, provide a theorem and proof that there is a trade-off between regularly shaped, or compactness, of districts and partisan efficiency, or the amount of votes per party which work towards a victory. I rewrite their paper in a different way which will provide more details and be more clear to the reader than the original one. I will provide examples and other proofs which build off of their ideas.

2
Background Information

2.1 Three Criteria for Fair Districts

There are "three well established desiderata" used to flag districts as potentially gerrymandered. I will explain each, both formally and informally, and provide examples for clarity.

2.1.1 One Person, One Vote

This criterion ensures that each district has approximately the same population size. It is very difficult to create districts with exactly the same population, so instead, we create an interval window that the population has to fall between. This is the least controversial, and possibly most important criterion because it makes sure that each voter has the same voting power. For example, the outcome of a district of one voter would be entirely in the hands of that one voter. In a district consisting of one million voters, each one would only have one millionth of a say in the outcome of the election.

Definition 2.1.1. Let T be the population of voters in the state. Let k be the number of districts in the state. Let T_i be the population of voters in district i for all $i \in \{1, \ldots, k\}$. The state satisfies **Criterion (i), One Person One Vote** if there exists $\delta \in [0, 1)$ such that

$$(1-\delta)\left\lfloor \frac{T}{k} \right\rfloor \leq T_i \leq (1+\delta)\left\lceil \frac{T}{k} \right\rceil$$

for all $i \in \{1, \ldots, k\}$. △

Having a smaller δ would mean that the populations of the districts must be closer together. Having a larger δ means there is more leeway.

Example 2.1.2. Let $T = 25$ and $k = 4$. First, suppose $\delta = 0.1$. Then $(1-\delta)\left\lfloor \frac{T}{k} \right\rfloor = (1 - 0.1)\left\lfloor \frac{24}{4} \right\rfloor = 0.9 \cdot 6 = 4.8$ and $(1+\delta)\left\lceil \frac{T}{k} \right\rceil = (1+0.1)\left\lceil \frac{25}{4} \right\rceil = 1.1 \cdot 7 = 7.7$ Therefore, $4.8 \leq T_i \leq 7.7$ for all $i \in \{1, \cdots, k\}$, and hence the population of the four districts must each be between 5 and 7. ♢

2.1.2 Polsby-Popper Compactness

When districts are drawn with the intent of including or excluding a certain demographic, the boundary often looks strange, long and curving to dodge certain areas and include other non-contiguous ones. In this case, the districts are often not compact, meaning they have a large perimeter and a small area. Daniel D. Polsby and Robert D. Popper invented the Polsby-Popper score in [3], which measures the ratio of area to perimeter of a district to determine compactness. A higher ratio indicates that districts are more compact and a smaller ratio indicates that districts are less compact and more likely to be gerrymandered.

Definition 2.1.3. Let k be the number of districts in the state. Let $\{D_1, \ldots, D_k\}$ be the districts in the state. Let R_i be the area of district i and let P_i be the perimeter of district i for all $i \in \{1, \ldots, k\}$. The state satisfies **Criterion (ii), Polsby-Popper Compactness** if there exists $\gamma \in (0, \infty)$ such that

$$\frac{4\pi R_i}{P_i^2} \geq \gamma$$

for all $i \in \{1, \ldots, k\}$. △

It is very difficult to calculate the perimeter of a shape whose boundary is made of irregular curves. For this reason, we include here the Grid Polsby-Popper Ratio which also measures compactness, but for districts with more easy to measure perimeters.

Definition 2.1.4. In a simplified district, one in which district boundaries are formed with rectilinear lines, we can determine compactness of a district by using the **Grid Polsby-Popper Ratio**. Let $\{D_1, \ldots, D_k\}$ be the districts in the state. Let R_i be the area of district i and let P_i be the perimeter of district i for all $i \in \{1, \ldots, k\}$. The **Grid Polsby-Popper Ratio** is $\frac{16R_i}{P_i^2}$. A state satisfies the **Grid Polsby-Popper Compactness** if there exists $\gamma \in (0, \infty)$ such that

$$\frac{16R_i}{P_i^2} \geq \gamma$$

for all $i \in \{1, \ldots, k\}$. △

Example 2.1.5. Here we will calculate the Grid Polsby-Popper ratio of each district in Figure 2.1.1. In Figure 2.1.1, we have X represent one voter for Party $\mathcal{X}$, and O represents one voter for Party $\mathcal{O}$. The bold lines represent the boundaries of the 6 districts. Going clockwise starting in the top left corner, number the districts 1 through 5 and number the district in the center 6.

Districts 1,2, 4 and 5 each have a perimeter of 12 and have a Grid Polsby-Popper Ratio of $\frac{16R_i}{P_i^2} = \frac{16 \cdot 7}{12^2} = 0.78$. Districts 3 and 6 have a perimeter of 14, and are less compact, with a Grid Polsby-Popper Ratio of $\frac{16R_i}{P_i^2} = \frac{16 \cdot 7}{14^2} = 0.57$. ◊

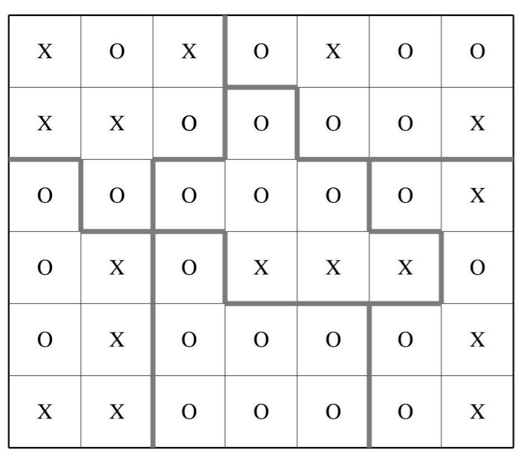

Figure 2.1.1.

2.1.3 Partisan Efficiency

Partisan efficiency measures how efficiently the votes of either party are used. For example, if one party wins a district by a landslide, this is not an efficient use of its votes considering that it only needed slightly more than half the votes to win. After receiving more than half the votes, any remaining voters would have been more useful if they lived in other districts which lost by a small amount of votes. A party is inefficient when the majority of voters belonging to one group are packed into a few districts. Also, if one party loses many districts by only a small amount, then each of the votes that went towards these districts are wasted. They would have been used more efficiently if they were moved to a district that they had a chance of winning and had let their district lose by a large amount.

If a state has a close number of voters for its two different parties, but still has one party win the election by a large amount of districts, then there is reason to suspect that the efficiency gap is large and gerrymandering is occurring. That is why the criterion states that if the difference between total number of votes for a party is less than a certain amount, then we must ensure that the efficiency gap is small. When an efficiency gap of a state is small, there is less of a chance that it is gerrymandered.

Definition 2.1.6. Let T be the population of voters in the state. Let k be the number of districts in the state. For each $i \in \{1, \ldots, k\}$, let T_i be the population of voters in district i, and let A_i and B_i be the number of voters in district i for Party $\mathcal{A}$ and Party $\mathcal{B}$, respectively. Let A and B be the number of voters in the state for Party $\mathcal{A}$ and Party $\mathcal{B}$, respectively.

For each $i \in \{1, \ldots, k\}$, if Party $\mathcal{A}$ wins district i, the wasted votes in that district are defined as

$$W_{A,i} = A_i - \left\lceil \frac{T_i}{2} \right\rceil \qquad \text{and} \qquad W_{B,i} = B_i;$$

similarly if Party $\mathcal{B}$ wins the district. The **Efficiency Gap** is defined as

$$\text{EG} = \frac{1}{T} \sum_{i=1}^{k} (W_{A,i} - W_{B,i}),$$

The state satisfies **Criterion (iii), Partisan Efficiency** if there exist $\alpha, \beta \in (0, \infty)$ such that if $|A - B| < \beta T$, then $|EG| < \frac{1}{2} - \alpha$. $\triangle$

Example 2.1.7. Here we will calculate the Efficiency Gap of the state in Figure 2.1.1. Going clockwise starting in the top left corner, number the districts 1 through 5 and number the district in the center 6. The wasted votes for Party $\mathcal{O}$ and Party $\mathcal{X}$ in each district are as follows: $W_{O,1} = 3$, $W_{X,1} = 0$, $W_{O,2} = 1$, $W_{X,2} = 2$, $W_{O,3} = 0$, $W_{X,3} = 3$, $W_{O,4} = 3$, $W_{X,4} = 0$, $W_{O,5} = 3$, $W_{X,5} = 0$, $W_{O,6} = 0$, $W_{X,6} = 3$. The efficiency gap for the entire state is

$$EG = \frac{1}{42}((3-0) + (1-2) + (0-3) + (3-0) + (3-0) + (0-3)) = \frac{2}{42}.$$

Let $\beta = 0.4$. Then $|O - X| = |26 - 16| = 10 < 16.8 = 0.4 \cdot 42 = \beta \cdot T$. First, let $\alpha = 0.3$. Then $|EG| = \frac{2}{42} = 0.048 < 0.2 = \frac{1}{2} - 0.3 = \frac{1}{2} - \alpha$, and Criterion (iii), Partisan Efficiency is satisfied.

Next, let $\alpha = 0.48$. Then $|EG| = \frac{2}{42} = 0.048 \not< 0.02 = \frac{1}{2} - 0.48 = \frac{1}{2} - \alpha$, and Criterion (iii), Partisan Efficiency is not satisfied. $\lozenge$

3

Gerrymandering in a Simplified County Districting System

3.1 A County Districting System

Realistically, voters live all over the place and it is really difficult to group them together in an orderly way. This makes the task of creating fair voting districts very complicated. But first if we suppose that voters do live in an orderly simplified way, we are able to work with the simplified districts and draw conclusions that will help us when we next look at more realistic districts. To simplify the state, here we assume that each voter lives in a county. In Theorem 3.2.2, we assume each county has an equal population size. In reality, this would be very difficult to achieve, but here it gives us a place to start.

Definition 3.1.1. A **County Districting System** is a subdivision of the state into smaller shapes, called **Counties**, of population of at least 2, such that each voting district is a union of these shapes. $\triangle$

Figure 3.1.1 shows an example of a state divided into districts using a County Districting System.

Figure 3.1.1.

3.2 Implication of Criterion (i)

Here, we look at Criterion (i), which ensures that all the districts are around the same size. We show that when voters are organized into counties, it is much easier to ensure that each district is the same size, and therefore, that each voter has the same amount of voting power.

First, we show a lemma that is needed in the proof and then we show the proof.

Lemma 3.2.1. *Let $T, k \in \mathbb{N}$. Suppose $T, k \geq 2$. Then $\lceil \frac{T}{k} \rceil + \lfloor \frac{T}{k} \rfloor \leq T$.*

Proof. For the first case, suppose $T = kr$ for some $r \in \mathbb{N}$. Then $\lceil \frac{T}{k} \rceil + \lfloor \frac{T}{k} \rfloor = \lceil \frac{kr}{k} \rceil + \lfloor \frac{kr}{k} \rfloor = 2r \leq kr = T$. For the second case, suppose that $T = kr + n$ for some $r \in \mathbb{N}$ and some $n \in \{1, \ldots, k-1\}$. Then $\lceil \frac{T}{k} \rceil + \lfloor \frac{T}{k} \rfloor = \lceil \frac{kr+n}{k} \rceil + \lfloor \frac{kr+n}{k} \rfloor = \lceil r + \frac{n}{k} \rceil + \lfloor r + \frac{n}{k} \rfloor = 2r + 1 \leq kr + n = T$. $\square$

Theorem 3.2.2. *Suppose there is population T who lives in a state which is divided into at least 2 districts using the county districting system. Suppose each county has the same population size. Let $\delta \in [0, 1)$. If Criterion (i) holds and $\delta < \frac{1}{T}$, then all districts have the same population.*

Proof. Suppose Criterion (i) holds and suppose $\delta < \frac{1}{T}$. Let k be the number of districts and let $k \geq 2$. Let $\{T_1, \ldots, T_k\}$ be the populations of the districts. Let L_i be the number of counties in district i for all $i \in \{1, \ldots, k\}$. Let S be the population per county and let $S \geq 2$. By Criterion (i), we know

$$(1-\delta)\left\lfloor \frac{T}{k} \right\rfloor \leq T_i \leq (1+\delta)\left\lceil \frac{T}{k} \right\rceil$$

for all $i \in \{1, \ldots, k\}$. Let $i, j \in \{1, \ldots, k\}$. Without loss of generality, suppose $T_i \geq T_j$. Then $T_i \leq (1+\delta)\left\lceil \frac{T}{k} \right\rceil$ and $T_j \geq (1-\delta)\left\lfloor \frac{T}{k} \right\rfloor$. Therefore, $T_i - T_j \leq (1+\delta)\left\lceil \frac{T}{k} \right\rceil - (1-\delta)\left\lfloor \frac{T}{k} \right\rfloor = \left\lceil \frac{T}{k} \right\rceil - \left\lfloor \frac{T}{k} \right\rfloor + \delta(\left\lceil \frac{T}{k} \right\rceil + \left\lfloor \frac{T}{k} \right\rfloor)$. Note that $\left\lceil \frac{T}{k} \right\rceil - \left\lfloor \frac{T}{k} \right\rfloor \in \{0, 1\}$.

By Lemma 3.2.1, we know $\delta(\left\lceil \frac{T}{k} \right\rceil + \left\lfloor \frac{T}{k} \right\rfloor) < \frac{1}{T} \cdot T = 1$. Therefore $T_i - T_j < 2$.

Note $T_i = L_i S$ and $T_j = L_j S$. Then $S(L_i - L_j) = T_i - T_j < 2$. Then $0 \leq L_i - L_j < \frac{2}{S} \leq 1$. Since L_i and L_j are whole numbers, it must be the case that $L_i - L_j = 0$. Then $T_i = T_j$. □

3.3 An Impossibility Theorem for Gerrymandering a County Districting System

The paper [1] by Boris Alexeev and Dustin G. Mixon proposed the idea that the three criteria for a fair districting system that we defined in Chapter 2 cannot all be used at once because there is some arrangement of voters that would never be able to satisfy all three, it is impossible.

Here, we take our simplified districting system, our county districting system, and show that the efficiency gap cannot be used on it because no matter how the districts are drawn, there will always be some arrangement of voters such that the efficiency gap is too large to be considered fair.

First, we provide a lemma that will be needed for the proof.

Lemma 3.3.1. *Let $\alpha, \beta \in (0, \infty)$. There exists an $a, b \in \mathbb{N}$ such that*

1. $a > b$,
2. $a + b$ *is even,*
3. $\frac{a-b}{a+b} < min\{\frac{1}{2}, \alpha, \beta\}$.

Proof. By Archimedes, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \min\{\frac{1}{2}, \alpha, \beta\}$. Let $a = n+2$ and $b = n$. Then

$$\frac{a-b}{a+b} = \frac{(n+2)-n}{(n+2)+n} = \frac{2}{2n+2} = \frac{1}{n+1} < \frac{1}{n} < \min\{\frac{1}{2}, \alpha, \beta\}.$$

Observe that $a + b = 2(n+1)$ which is even. Clearly, $a > b$. □

Theorem 3.3.2. *Let $\alpha, \beta \in (0, \infty)$. For any county districting system, there are populations A and B such that for every choice of districts, Criterion (iii) is violated. That is, there is no county districting system that guarantees a small efficiency gap.*

Proof. Suppose the state consists of n counties. Let a and b be as in Lemma 3.3.1. In each county, there are a voters for Party $\mathcal{A}$ and b voters for Party $\mathcal{B}$.

Suppose there are k districts in the county districting system. Let L_i be the number of counties in district i for all $i \in \{1, \ldots, k\}$. Notice there are n counties in the state and that $\sum_{i=1}^{k} L_i = n$. Let A_i and B_i be the number of voters for Party $\mathcal{A}$ and Party $\mathcal{B}$ in district i, respectively, for all $i \in \{1, \ldots, k\}$. Notice that $aL_i = A_i$ and $bL_i = B_i$. Let T_i be the population of district i, so that $A_i + B_i = T_i$. Let A and B be the number of voters in the whole state who vote for Party $\mathcal{A}$ and Party $\mathcal{B}$, respectively. Notice that $A = an$ and $B = bn$. Let T be the total population of voters, and notice that $A + B = T$.

Since $a > b$, then $aL_i > bL_i$ for all $i \in \{1, \ldots, k\}$ and therefore, Party $\mathcal{A}$ wins every district.

Using the formula for Efficiency Gap, stated in Criterion (iii), we see for each $i \in \{1, \ldots, k\}$ that $W_{A,i} = aL_i - \left\lceil \frac{(a+b)L_i}{2} \right\rceil$ and $W_{B,i} = bL_i$. Since we let $a + b$ be even, we can remove the ceiling brackets, so $W_{A,i} = aL_i - \frac{(a+b)L_i}{2}$ for all $i \in \{1, \ldots, k\}$.

The Efficiency Gap is

$$EG = \frac{1}{n \cdot (a+b)} \sum_{i=1}^{k} aL_i - \frac{(a+b)L_i}{2} - bL_i = \frac{1}{n \cdot (a+b)} \sum_{i=1}^{k} L_i(a - \frac{a+b}{2} - b)$$

Since we know that $\sum_{i=1}^{k} L_i = n$, this can be simplified to

$$EG = \frac{n}{n \cdot (a+b)}(a - \frac{a+b}{2} - b) = \frac{a-b}{a+b} - \frac{1}{2}.$$

Since from Lemma 3.3.1 we know that $\frac{a-b}{a+b} < \frac{1}{2}$, then $|EG| = |\frac{a-b}{a+b} - \frac{1}{2}| = \frac{1}{2} - \frac{a-b}{a+b}$.

From Lemma 3.3.1, we know that $\alpha > \frac{a-b}{a+b}$. Therefore, it is true that $\frac{1}{2} - \frac{a-b}{a+b} > \frac{1}{2} - \alpha$, and hence $|EG| \nless \frac{1}{2} - \alpha$.

But, from Lemma 3.3.1, we know that $\frac{a-b}{a+b} < \beta$. Hence

$$\frac{an - bn}{an + bn} < \beta,$$

and so

$$an - bn < \beta(an + bn),$$

and substitution tells us that $A - B < \beta T$. Since $A > B$, we see that $|A - B| < \beta T$.

Hence, $|A - B| < \beta T$ holds, but $|EG| < \frac{1}{2} - \alpha$ does not hold. Hence, for any grid districting system, there are populations A and B such that every choice of districts violates Criterion (iii). $\square$

4
Gerrymandering in a Realistic Districting System

4.1 A Realistic Districting System

Now we will look at a districting system in which voters can live anywhere in the state and are not confined to counties or any other way of organizing them. The boundaries of the districts can be curvy, squiggly, straight, or drawn in any other way without restrictions.

Definition 4.1.1. Suppose the state is a square. A **Districting System** is a subdivision of the state into areas called **Districts**.

Figure 4.1.1 shows an example of a state divided into districts using a Districting System.

4.2 An Impossibility Theorem for Gerrymandering

The paper [1] proves the same theorem that we prove here, that using the three criteria for fair districts that we listed in Chapter 2 would not work because no matter how the district boundaries are drawn, there is some arrangement of the population such that one of the criteria would be violated. We chose to rewrite their proof because it left out many details that we found necessary in order to accept the theorem as true and in order for the reader to be able to follow the proof.

The following lemmas address details that were skipped in the original proof.

Figure 4.1.1.

Lemma 4.2.1. *Let $\delta \in [0,1)$. Let $\alpha \in (0, \frac{1}{2})$. Let $\beta, \gamma \in (0, \infty)$. Let $k \in \mathbb{N}$ be such that $k \geq 2$. Let $F = \sqrt{\frac{1-\delta}{2k}}$. There exist $a, b, l, n \in \mathbb{N}$ such that*

1. $\frac{b}{a} < \frac{Fn^2 - 4\pi\gamma^{-1}F\sqrt{2}n - 8\pi^2\gamma^{-1}}{F^2n^2 + 4\pi\gamma^{-1}F\sqrt{2}n + 8\pi^2\gamma^{-1}}$

2. $b < a$

3. $\frac{a-b}{a+b} < \beta$

4. $\frac{a-2b}{a+b} \leq \alpha - \frac{1}{2}$

5. $a + b = l^2$

6. *l is even.*

Proof. By Archimedes, there exists $m_1 \in \mathbb{N}$ such that $\frac{1}{m_1} < \beta$.

Because $\alpha \in (0, \frac{1}{2})$, we know that $\alpha - \frac{1}{2} \in (-\frac{1}{2}, 0)$. Observe $\lim_{c\to\infty} \frac{-c+2}{2c+2} = -\frac{1}{2}$. Therefore, there is some $m_2 \in \mathbb{N}$ such that if $p \in \mathbb{N}$ and $p \geq m_2$, then $\frac{-p+2}{2p+2} < \alpha - \frac{1}{2}$.

Next, let $q = max\{m_1, m_2\}$. Let $m = 2q^2 - 1$. Note that $m \geq m_1$ and $m \geq m_2$.

Let $a = m + 2$ and $b = m$.

Then $b < a$ and $\frac{b}{a} < 1$ and Part (2) is satisfied. Then $\frac{a-b}{a+b} = \frac{1}{m+1} < \frac{1}{m} \leq \frac{1}{m_1} < \beta$, hence Part (3) is satisfied.

Next, because $m \geq m_2$ we know that $\frac{a-2b}{a+b} = \frac{m+2-2m}{m+2+m} = \frac{2-m}{2m+2} = \frac{-m+2}{2m+2} < \alpha - \frac{1}{2}$. Therefore, Part (4) is true.

Let $l = 2q$. Then l is even, so Part (6) is true. Then $a + b = 2m + 2 = (2q)^2 = l^2$. Therefore, Part (5) is true.

Next, we simplify the fraction in Part (1). Let $y = 4\pi\gamma^-1F\sqrt{2}$, and let $z = 8\pi^2\gamma^-1$. Let $r \in \mathbb{N}$.

Then

$$\frac{Fr^2 - 4\pi\gamma^{-1}F\sqrt{2}r - 8\pi^2\gamma^{-1}}{F^2r^2 + 4\pi\gamma^{-1}F\sqrt{2}n + 8\pi^2\gamma^{-1}} = \frac{F - \frac{y}{r} - \frac{z}{r^2}}{F + \frac{y}{r} + \frac{z}{r^2}}.$$

Observe

$$\lim_{r\to\infty} \frac{F - \frac{y}{r} - \frac{z}{r^2}}{F + \frac{y}{r} + \frac{z}{r^2}} = 1.$$

Recall $\frac{b}{a} < 1$. Therefore, there is some $n \in \mathbb{N}$ such that if $p \in \mathbb{N}$ and $p \geq n$, then $\frac{b}{a} < \frac{F - \frac{y}{p} - \frac{z}{p^2}}{F + \frac{y}{p} + \frac{z}{p^2}}$. In particular, we have $\frac{b}{a} < \frac{F - \frac{y}{n} - \frac{z}{n^2}}{F + \frac{y}{n} + \frac{z}{n^2}}$. Hence

$$\frac{b}{a} < \frac{Fn^2 - 4\pi\gamma^{-1}F\sqrt{2}n - 8\pi^2\gamma^{-1}}{F^2n^2 + 4\pi\gamma^{-1}F\sqrt{2}n + 8\pi^2\gamma^{-1}}.$$

Therefore, Part (1) is true. $\square$

In the following lemma, and again later on, we will be referring to a square lattice in the plane, as seen in figure 4.2.1.

Lemma 4.2.2. *Suppose q points are arranged as vertices of a square lattice of unit u. Let P be the perimeter of a region containing the points. Then $P \geq u\sqrt{q}$.*

Proof. The most compact way for points to be arranged on a square lattice is in a square. The two furthest points from each other on a square are on corners diagonal from each other. The distance between these two points can be found using the Pythagorean Theorem, where the sides of the square are the sides of a right triangle and the distance between the two corner points is the hypotenuse of the right triangle. The length of the side of the square is $(\sqrt{q} - 1)u$.

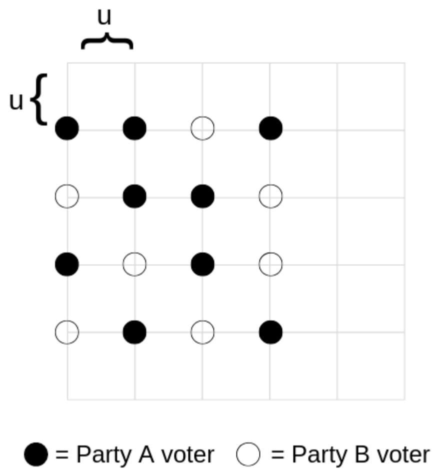

Figure 4.2.1.

Therefore, the distance between the two corner points is $\sqrt{2((\sqrt{q}-1)u)^2} = \sqrt{2}(\sqrt{q}-1)u$. The perimeter is at least twice the distance between the two corner points, so the perimeter is at least $2\sqrt{2}(\sqrt{q}-1)u$. Since $q = a+b$ and because $a > b$ and $b \geq 1$, it must be that $a \geq 2$, and therefore $q \geq 3$. Then $\sqrt{q} \geq \sqrt{3}$, and it follows that $(2\sqrt{2}-1)\sqrt{q} \geq (2\sqrt{2}-1)\sqrt{3}$. Since $2\sqrt{2}-1 \approx 3.166$ and $2\sqrt{2} \approx 2.828$, it is follows that $(2\sqrt{2}-1)\sqrt{q} \geq 2\sqrt{2}$. Adding $\sqrt{q}$ to each side gives us $2\sqrt{2}\sqrt{q} \geq 2\sqrt{2}+\sqrt{q}$, and subtracting $2\sqrt{2}$ from each side gives us $2\sqrt{2}(\sqrt{q}-1) \geq \sqrt{q}$. Therefore, $2\sqrt{2}(\sqrt{q}-1)u \geq \sqrt{q}u$ and so $P \geq \sqrt{q}u$. $\square$

Lemma 4.2.3. *Let $\delta \in [0,1)$. Let $k \in \mathbb{N}$ and suppose $k \geq 2$. Let $F = \sqrt{\frac{1-\delta}{2k}}$. Let P be the perimeter of a district and let R be the area of the district. Suppose $P \geq F$. Let $\epsilon \in (0,\infty)$, and let $E = \frac{\sqrt{2}P}{\epsilon} + 2\pi$. Let $\gamma \in (0,\infty)$. Suppose that Criterion (ii) holds for this district, that is, that $\gamma \leq \frac{4\pi R}{P^2}$.*

1.

$$\frac{F^2 - 4\pi\gamma^{-1}F\sqrt{2}\epsilon - 8\pi^2\gamma^{-1}\epsilon^2}{F^2 + 4\pi\gamma^{-1}F\sqrt{2}\epsilon + 8\pi^2\gamma^{-1}\epsilon^2} \leq \frac{P^2 - 4\pi\gamma^{-1}\epsilon^2 E}{P_1^2 + 4\pi\gamma^{-1}\epsilon^2 E}.$$

2.

$$\frac{P^2 - 4\pi\gamma^{-1}\epsilon^2 E}{P^2 + 4\pi\gamma^{-1}\epsilon^2 E} \leq \frac{R - \epsilon^2 E}{R + \epsilon^2 E}.$$

Proof. **(1)**. Because $P \geq F$, it is true that $F\sqrt{2}P(P-F) \geq 0$ and $2\pi\epsilon(F^2 - P^2) \leq 0$, since F, P, and ϵ are all positive. Then $2\pi\epsilon(F^2 - P^2) \leq F\sqrt{2}P(P-F)$, and by multiplication, that $F^2 2\pi\epsilon - 2\pi\epsilon P^2 \leq F\sqrt{2}P^2 - F^2\sqrt{2}P$. Rearranging the terms gives us $F^2\sqrt{2}P + F^2 2\pi\epsilon \leq F\sqrt{2}P^2 + 2\pi\epsilon P^2$, which is equivalent to $F^2\epsilon(\frac{\sqrt{2}P}{\epsilon} + 2\pi) \leq (F\sqrt{2} + 2\pi\epsilon)P^2$. Substituting in E and multiplying both sides by $4\pi\gamma^{-1}\epsilon$ gives us $F^2 4\pi\gamma^{-1}\epsilon 2E \leq (4\pi\gamma^{-1}F\sqrt{2}\epsilon + 8\pi^2\gamma^{-1}\epsilon^2)P^2$. Next let $A = F^2$, $B = 4\pi\gamma^{-1}F\sqrt{2}\epsilon$, $C = 8\pi^2\gamma^{-1}\epsilon^2$, $x = P^2$, and $y = 4\pi\gamma^{-1}\epsilon^2 E$ and substitute these in, to get $Ay \leq (B+C)x$. By rearranging this, it is equivalent to $Ay - Bx - Cx \leq -Ay + Bx + Cx$, and furthermore, to $Ax + Ay - Bx - By - Cx - Cy \leq Ax - Ay + Bx - By + Cx - Cy$. From factoring, this inequality is equivalent to $(A-B-C)(x+y) \leq (A+B+C)(x-y)$, which is then equivalent to $\frac{A-B-C}{A+B+C} \leq \frac{x-y}{x+y}$. By substituting the original values back in, we know that $\frac{F^2 - 4\pi\gamma^{-1}F\sqrt{2}\epsilon - 8\pi^2\gamma^{-1}\epsilon^2}{F^2 + 4\pi\gamma^{-1}F\sqrt{2}\epsilon + 8\pi^2\gamma^{-1}\epsilon^2} \leq \frac{P^2 - 4\pi\gamma^{-1}\epsilon^2 E}{P_1^2 + 4\pi\gamma^{-1}\epsilon^2 E}$.

(2). Since Criterion (2) holds, we know $\gamma \leq \frac{4\pi R}{P^2}$, which, by multiplying each side by $\epsilon^2 E$ and by rearranging, is equivalent to $P^2\epsilon^2 E \leq 4\pi\gamma^{-1}\epsilon^2 ER$. Now, let $x = P^2$, let $v = R$, let $y = 4\pi\gamma^{-1}\epsilon^2 E$, and let $z = \epsilon^2 E$, and substitute in these values. Hence, $xz \leq yv$. By multiplying each side by 2, adding xv and $-yz$ to each side, and rearranging terms, we see that $xv + xz - yv - yz \leq xv + yv - zx - zy$. We can factor, to see that $(x-y)(v+z) \leq (v-z)(x+y)$, and furthermore, that $\frac{x-y}{x+y} \leq \frac{v-z}{v+z}$. Now we can substitute back in the original values to see that $\frac{P^2 - 4\pi\gamma^{-1}\epsilon^2 E}{P^2 + 4\pi\gamma^{-1}\epsilon^2 E} \leq \frac{R - \epsilon^2 E}{R + \epsilon^2 E}$. □

Lemma 4.2.4. *Let $T, n \in \mathbb{N}$. Suppose $k \geq 2$ and $T \geq 2k$. Then $\lfloor \frac{T}{k} \rfloor \geq \frac{T}{2k}$.*

Proof. First, suppose that $T = kr$ for some $r \in \mathbb{N}$. Because $r > 0$, it is true that $r > \frac{r}{2}$. We know that $r = \lfloor \frac{kr}{k} \rfloor = \lfloor \frac{T}{k} \rfloor$. We also know that $\frac{r}{2} = \frac{kr}{2k} = \frac{T}{2k}$. Hence, $\lfloor \frac{T}{k} \rfloor \geq \frac{T}{2k}$. Next, suppose that $T = kr + n$ for some $r \in \mathbb{N}$ and some $n \in \{1, \dots, k-1\}$. Since $\frac{n}{k} < 1$ and $\frac{r}{2} > 1$, we know that $\frac{n}{k} < \frac{r}{2}$ and therefore $\frac{r}{2} + \frac{n}{k} < r$. Because $\frac{n}{k} > 0$, we know $\frac{n}{2k} < \frac{n}{k}$, and it follows that $\frac{r}{2} + \frac{n}{2k} < r$. Since $\lfloor \frac{T}{k} \rfloor = \lfloor \frac{kr+n}{k} \rfloor = \lfloor r + \frac{n}{k} \rfloor = r$ and $\frac{T}{2k} = \frac{kr+n}{2k} = \frac{kr}{2k} + \frac{n}{2k} = \frac{r}{2} + \frac{n}{2k} < r$, it is the case that $\lfloor \frac{T}{k} \rfloor \geq \frac{T}{2k}$. □

Theorem 4.2.5. *Let $\delta \in [0, 1)$ and let $\alpha, \beta, \gamma \in (0, \infty)$. For every possible arrangement of at least 2 districts in a square, there exists an arrangement of Party $\mathcal{A}$ voters and Party $\mathcal{B}$ voters that violates one of Criteria (i), (ii), and (iii).*

Proof. Without loss of generality, suppose our square is 1×1. We assume Criteria (i) and (ii) are true and show that Criterion (iii) does not hold. Let n, a, b, and l be as in Lemma 4.2.1. Let $\epsilon = \frac{1}{n}$. Begin by dividing a 1×1 square into a grid of n^2 smaller squares with edge length ϵ, as shown in Figure 4.2.2.

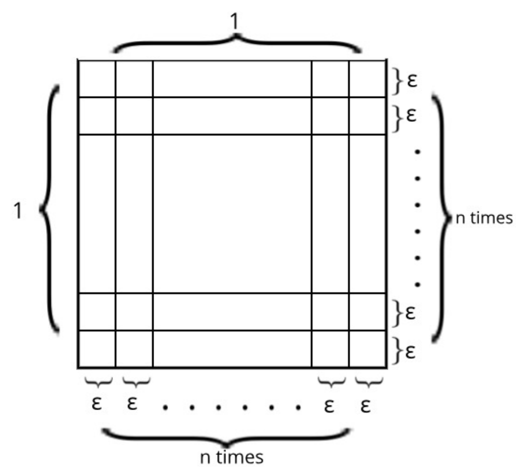

Figure 4.2.2.

Further divide each $\epsilon \times \epsilon$ square into a grid with l^2 smaller squares of edge length $\frac{1}{nl}$, as shown in Figure 4.2.3.

Suppose Party $\mathcal{A}$ has a voters in each $\epsilon \times \epsilon$ square and suppose Party $\mathcal{B}$ has b voters in each $\epsilon \times \epsilon$ square. Since Lemma 4.2.1 states that $a + b = l^2$, it would be possible to distribute the voters so that each $\frac{1}{ln} \times \frac{1}{ln}$ square contains one voter in its center, either from Party $\mathcal{A}$ or Party $\mathcal{B}$. Let us suppose that the voters are distributed in that way. Let T be the total population of voters in the state. It follows that $n^2(a + b) = n^2 l^2 = T$.

Let $k \in \mathbb{N}$ and let $k \geq 2$. Let $i \in \{1, \ldots, k\}$. Next, partition the 1×1 square into k districts, denoted $D_1, \ldots, D_k$, in a way so that Criteria (i) and (ii) are satisfied.

Let $i \in \{1, \ldots, k\}$. The boundary of D_i is denoted ∂D_i, and the length of the boundary, which is the perimeter of D_i, is denoted P_i. We find the tube of D_i, denoted U_i, by taking ∂D_i and

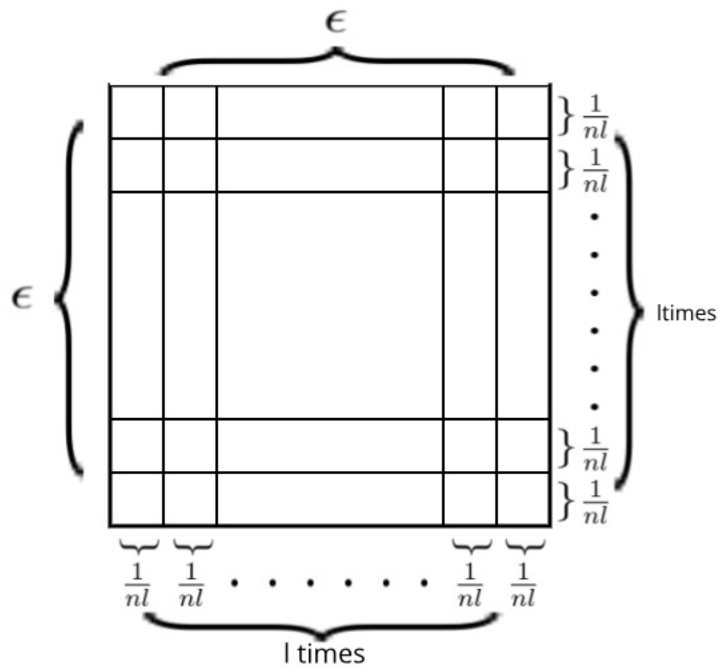

Figure 4.2.3.

thickening it by $\epsilon\sqrt{2}$ on either side. We see that ∂D_i is contained in a union of $\epsilon \times \epsilon$ squares in a rectilinear shape. As shown in Figure 4.2.4, we see that U_i contains this union of $\epsilon \times \epsilon$ squares.

It can be shown that U_i has area of at most $\sqrt{2}\delta_i\epsilon + 2\pi\epsilon^2$; we omit the details. The area of one $\epsilon \times \epsilon$ square is ϵ^2, so the tube has at most $\frac{\sqrt{2}\delta D_i\epsilon+2\pi\epsilon^2}{\epsilon^2} = \frac{\sqrt{2}\delta D_i}{\epsilon} + 2\pi$ $\epsilon \times \epsilon$ squares. Let E_i be the greatest possible number of $\epsilon \times \epsilon$ squares in U_i. Then $E_i = \frac{\sqrt{2}\delta D_i}{\epsilon} + 2\pi$.

Now we will determine the fewest and greatest amount of votes for Party $\mathcal{A}$ and Party $\mathcal{B}$ in D_i. First, suppose all of the voters for Party $\mathcal{B}$ that belong to the union of $\epsilon \times \epsilon$ squares that ∂D_i intersects with lie inside of district i. Then D_i contains the greatest amount of votes for Party $\mathcal{B}$, which is the amount of $\epsilon \times \epsilon$ squares contained in D_i plus the amount of $\epsilon \times \epsilon$ squares contained in U_i times the amount of votes for Party $\mathcal{B}$ per $\epsilon \times \epsilon$ square. Let R_i be the area of D_i. Then the amount of $\epsilon \times \epsilon$ squares in D_i is $\frac{R_i}{\epsilon^2}$ and the greatest amount of votes for Party $\mathcal{B}$ is $b(\frac{R_i}{\epsilon^2} + E_i)$.

Next, suppose the voters for Party $\mathcal{A}$ that live in the $\epsilon \times \epsilon$ squares that intersect with ∂D_i all lie outside of D_i. Then D_i contains the fewest possible votes for Party $\mathcal{A}$, which is the amount

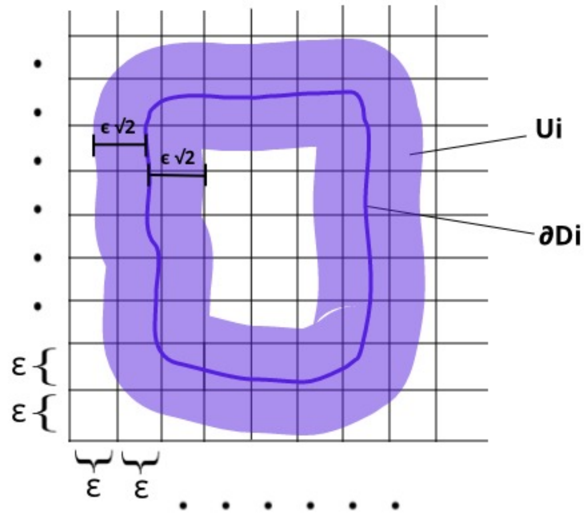

Figure 4.2.4.

of $\epsilon \times \epsilon$ squares contained in D_i minus the amount of $\epsilon \times \epsilon$ squares contained in U_i times the amount of votes for Party $\mathcal{A}$ per $\epsilon \times \epsilon$ square, which is $a(\frac{R_i}{\epsilon^2} - E_i)$.

Recall from Definition 2.1.1 that T_i is the total population for D_i and $(1-\delta)\left\lfloor \frac{T}{k} \right\rfloor \leq T_i \leq (1+\delta)\left\lceil \frac{T}{k} \right\rceil$. Since $T, k \in \mathbb{N}$, and $k \geq 2$ and $T \geq 2k$, from Criterion (i) and Lemma 4.2.4 it follows that

$$T_i \geq (1-\delta)\left\lfloor \frac{T}{k} \right\rfloor \geq (1-\delta)\frac{T}{2k} = (1-\delta)\frac{n^2 l^2}{2k}.$$

Since Lemma 4.2.2 tells us that the smallest possible perimeter around q points arranged on a square lattice of unit u is $u\sqrt{q}$, we know that we can take the number of points arranged on a square lattice, multiply it by the unit length of the lattice squared and then take the square root of that to find the smallest possible perimeter around around the points. Since D_i contains $(1-\delta)\frac{n^2l^2}{2k}$ points arranged on a square lattice of unit $\frac{1}{nl}$, it follows that $P_i \geq \sqrt{\frac{1-\delta}{2k}}$.

Now, recall from Lemma 4.2.1 that $\frac{a}{b} \leq \frac{F^2 - 4\pi\gamma^{-1}F\sqrt{2}\epsilon - 8\pi^2\gamma^{-1}\epsilon^2}{F^2 + 4\pi\gamma^{-1}F\sqrt{2}\epsilon + 8\pi^2\gamma^{-1}\epsilon^2}$. Since Criteria (i) and (ii) are satisfied, since $P_i \geq \sqrt{\frac{1-\delta}{2k}}$, since $k \geq 2$ and since $\epsilon \in (0, \infty)$, by Lemma 4.2.3 we know that $\frac{a}{b} \leq \frac{F^2 - 4\pi\gamma^{-1}F\sqrt{2}\epsilon - 8\pi^2\gamma^{-1}\epsilon^2}{F^2 + 4\pi\gamma^{-1}F\sqrt{2}\epsilon + 8\pi^2\gamma^{-1}\epsilon^2} \leq \frac{P_i^2 - 4\pi\gamma^{-1}\epsilon^2 E}{P_1^2 + 4\pi\gamma^{-1}\epsilon^2 E} \leq \frac{R_i - \epsilon^2 E}{R_i + \epsilon^2 E}$. Therefore $\frac{a}{b} \leq \frac{R_i - \epsilon^2 E}{R_i + \epsilon^2 E}$. By multiplying the top and bottom of the fraction on the right side of the inequality by $\frac{1}{\epsilon^2}$ and rearranging, we see that $a(\frac{R_i}{\epsilon^2} - E) \geq b(\frac{R_i}{\epsilon^2} + E)$. Let A_i be the amount of votes for Party $\mathcal{A}$ in D_i and let

B_i be the amount of votes for Party $\mathcal{B}$ in D_i. Since $a(\frac{R_i}{\epsilon^2} - E_i)$ is the fewest possible votes for Party $\mathcal{A}$ in D_i and $b(\frac{R_i}{\epsilon^2} + E_i)$ is the greatest amount of votes for Party $\mathcal{B}$ in D_i, it follows that $A_i \geq a(\frac{R_i}{\epsilon^2} - E_i) \geq b(\frac{R_i}{\epsilon^2} + E_i) \geq B_i$, and so for each District i, Party $\mathcal{A}$ wins the vote.

Since Party $\mathcal{B}$ loses every district, according to Definition 2.1.6, all of its votes are wasted. It has bn^2 votes, since there are n^2 $\epsilon \times \epsilon$ squares in the state and b votes per $\epsilon \times \epsilon$ square. Party $\mathcal{A}$ wastes $an^2 - \lceil \frac{T}{2} \rceil$ votes. Since $b < a$, then $\frac{n^2 2b}{2} < \frac{n^2(a+b)}{2} < \left\lceil \frac{n^2(a+b)}{2} \right\rceil = \lceil \frac{T}{2} \rceil$, that is, $bn^2 < \lceil \frac{T}{2} \rceil$. Then $an^2 - \lceil \frac{T}{2} \rceil < an^2 - bn^2$ and Party $\mathcal{A}$ wastes fewer than $an^2 - bn^2$ votes.

Then, by Definition 2.1.6 and Lemma 4.2.1,

$$EG < \frac{an^2 - 2bn^2}{n^2(a+b)} = \frac{a-2b}{a+b} \leq \alpha - \frac{1}{2} < 0.$$

Hence,

$$|EG| > |\frac{a-2b}{a+b}| \geq \frac{1}{2} - \alpha.$$

By Lemma 4.2.1, we know $a > b$ and we know $\frac{a-b}{a+b} < \beta$. Hence, $|A - B| = |an^2 - bn^2| = n^2(a-b) < \beta n^2(a+b) = \beta T$. For Criterion (iii) to be satisfied, it must be the case that if $|A - B| < \beta T$, then $|EG| < \frac{1}{2} - \alpha$. But we see that this criterion does not hold. $\square$

5

Ideas for Further Study

5.1 Other Definitions of Compactness

The impossibility theorem for gerrymandering showed that using the three well-known criteria for fair districting together does not work. Where can we look from here to keep searching for solutions? The criteria given in Chapter 2 are not the only ways to measure compactness, efficiency and individual voting power. For example, there are other well known methods for calculating compactness of a district. One of these methods uses the Roeck Compactness Ratio, from [4], where the area of the district is compared to the smallest possible circle it can fit inside. If we wanted to use the Roeck Compactness ratio as a criterion in place of Polsby-Popper, Criterion (ii) might look something like this:

Definition 5.1.1. Let k be the number of districts in the state. Let $\{D_1, \ldots, D_k\}$ be the districts in the state. Let R_i be the area of district i and let C_i be the area of the smallest possible circle that can contain district i for all $i \in \{1, \ldots, k\}$. The state satisfies **Criterion (ii), Roeck Compactness** if there exists $\gamma \in (0, \infty)$ such that

$$\frac{R_i}{C_i} > \gamma$$

for all $i \in \{1, \ldots, k\}$. $\triangle$

In this case, the number γ would be between 0 and 1, where if $\gamma = 1$, the district would be as compact as possible and if γ were very small, the district would not be compact. This also is true if we used Polsby-Popper Compactness. Therefore, the impossibility theorem for gerrymandering may still hold if we replaced Polsby-Popper Compactness with Roeck Compactness.

There are even more methods for determining compactness besides these two. And there may be other methods of determining voter efficiency that are yet to be discovered.

Perhaps there will never be a standard for fair districting systems that can apply to all districts and each one's fairness will have to be determined individually.

Perhaps, allowing a party who wins 51% of the votes to take the entire district is just not fair and voting districts should be done away with.

Bibliography

[1] Boris Alexeev and Dustin G. Mixon, *An Impossibility Theorem for Gerrymandering*, Amer. Math. Monthly **125** (2018), 878–884.

[2] Moon Duchin and Mira Bernstein, *A Formula Goes to Court: Partisan Gerrymandering and the Efficiency Gap*, Noticies of the American Math Society **64** (2017), 1020–1024.

[3] Daniel D. Polsby and Robert D. Popper, *The Third Criterion: Compactness as a Procedural Safeguard Against Partisan Gerrymandering*, Yale Law and Policy Review **9** (1991), 301–353.

[4] Ernest C. Reock Jr., *A Note: Measuring Compactness as a Requirement of Legislative Apportionment*, Midwest Journal of Political Science **5** (1961), 70–74.

[5] Nicholas Stephanopoulos and Eric McGhee, *Partisan Gerrymandering and the Efficiency Gap*, Chicago Unbound (2014), 1–3.

[6] *FairVote*, `https://www.fairvote.org/gerrymandering#gerrymandering_key_facts`.