The Mathematics of Partisan Gerrymandering

Dustin G. Mixon

THE OHIO STATE UNIVERSITY

LA Combinatorics and Complexity Seminar

December 8, 2020
Gerrymander detection by shape

Long history of detecting gerrymandering by shape

- "bizarre shape" quantified by "geographic compactness"
- constraints imposed by bipartisan redistricting commissions

Boston Gazette, 1812

Washington Post, 2014
Gerrymander detection by shape

Does shape provide enough information?

What about context? Does intent matter?

images from Wikipedia
Gerrymander detection by voting results

One modern approach: compare votes vs. seats

- proportionality (not a valid constitutional standard)
- efficiency gap (subject of Gill v. Whitford)
Gerrymander detection by voting results

One modern approach: compare votes vs. seats

- proportionality (not a valid constitutional standard)
- efficiency gap (subject of Gill v. Whitford)

Do voting results provide enough information?

---

**United States presidential election in Massachusetts, 2016**

<table>
<thead>
<tr>
<th>Party</th>
<th>Candidate</th>
<th>Running Mate</th>
<th>Votes</th>
<th>Percentage</th>
<th>Electoral votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>Hillary Clinton</td>
<td>Tim Kaine</td>
<td>1,995,196</td>
<td>60.0%</td>
<td>11</td>
</tr>
<tr>
<td>Republican</td>
<td>Donald Trump</td>
<td>Mike Pence</td>
<td>1,090,893</td>
<td>32.8%</td>
<td>0</td>
</tr>
<tr>
<td>Libertarian</td>
<td>Gary Johnson</td>
<td>William Weld</td>
<td>138,018</td>
<td>4.2%</td>
<td>0</td>
</tr>
<tr>
<td>Green</td>
<td>Jill Stein</td>
<td>Ajamu Baraka</td>
<td>47,661</td>
<td>1.4%</td>
<td>0</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td></td>
<td></td>
<td>3,269,783</td>
<td>100.00%</td>
<td>11</td>
</tr>
</tbody>
</table>

---

**Current representatives**

- 1st congressional district: Richard Neal (D) (since 1989)
- 2nd congressional district: Jim McGovern (D) (since 1997)
- 3rd congressional district: Niki Tsongas (D) (since 2007)
- 4th congressional district: Joseph P. Kennedy III (D) (since 2013)
- 5th congressional district: Katherine Clark (D) (since 2013)
- 6th congressional district: Seth Moulton (D) (since 2015)
- 7th congressional district: Mike Capuano (D) (since 1999)
- 8th congressional district: Stephen F. Lynch (D) (since 2001)
- 9th congressional district: Bill Keating (D) (since 2011)

---

images from Wikipedia
Gerrymander detection by voting results

One modern approach: compare votes vs. seats

- proportionality (not a valid constitutional standard)
- efficiency gap (subject of *Gill v. Whitford*)

Do voting results provide enough information?

<table>
<thead>
<tr>
<th>United States presidential election in Massachusetts, 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Party</strong></td>
</tr>
<tr>
<td>Democratic</td>
</tr>
<tr>
<td>Republican</td>
</tr>
<tr>
<td>Libertarian</td>
</tr>
<tr>
<td>Green</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
</tr>
</tbody>
</table>

Is partisan fairness at odds with geometry?

---

images from Wikipedia
Joint work with

Boris Alexeev
borisalexeev.com

Richard Kueng
Caltech

Soledad Villar
NYU
Part I

compact gerrymandering
How to quantify “bizarre shape”?
Classical notions of geographic compactness

**Isoperimetry.** wasted perimeter

- perimeter
- Polsby–Popper score $\propto \frac{(\text{area})}{(\text{perimeter})^2}$

---

Duchin, sites.duke.edu/gerrymandering/files/2017/11/MD-duke.pdf
Classical notions of geographic compactness

**Isoperimetry.** wasted perimeter

- perimeter
- Polsby–Popper score \( \propto \frac{\text{area}}{(\text{perimeter})^2} \)

**Convexity.** wasted area, *indentedness*

- \( \frac{\text{area}}{\text{(area of convex hull)}} \)
- Reock score \( = \frac{\text{area}}{\text{(area of minimum covering disk)}} \)
Classical notions of geographic compactness

**Isoperimetry.** wasted perimeter

- perimeter
- Polsby–Popper score \( \propto \frac{\text{(area)}}{\text{(perimeter)}^2} \)

**Convexity.** wasted area, *indentation*

- \( \frac{\text{(area)}}{\text{(area of convex hull)}} \)
- Reock score = \( \frac{\text{(area)}}{\text{(area of minimum covering disk)}} \)

**Dispersion.** second moment, *sprawl*

- average distance between pairs of points
- moment of inertia

---

Duchin, sites.duke.edu/gerrymandering/files/2017/11/MD-duke.pdf
Classical notions of geographic compactness

**Isoperimetry.** wasted perimeter

- perimeter
- Polsby–Popper score $\propto \frac{\text{area}}{(\text{perimeter})^2}$

**Convexity.** wasted area, *indenteness*

- $\frac{\text{area}}{\text{area of convex hull}}$
- Reock score $= \frac{\text{area}}{\text{area of minimum covering disk}}$

**Dispersion.** second moment, *sprawl*

- average distance between pairs of points
- moment of inertia

**Question 1:** Can we gerrymander with nice shapes?

---

Duchin, sites.duke.edu/gerrymandering/files/2017/11/MD-duke.pdf
Passing to a (cartoon) model

- state = closed disk
- $2n$ voters equispaced along concentric circle
- voter preferences = iid uniform from $\{\pm 1\}$
Passing to a (cartoon) model

- state = closed disk
- $2n$ voters equispaced along concentric circle
- voter preferences = iid uniform from $\{\pm 1\}$

Theorem

Partition a closed disk $C$ of unit radius into two regions $A, B$ whose closures are homeomorphic to $C$. Then

$$\max \left\{|\partial A|, |\partial B|\right\} \geq \pi + 2, \quad \min \left\{\frac{|A|}{|\text{hull}(A)|}, \frac{|B|}{|\text{hull}(B)|}\right\} \leq 1, \quad I_A + I_B \geq \frac{\pi}{2} - \frac{16}{9\pi}.$$ 

Equality is achieved in all three when $A$ and $B$ are complementary half-disks.

Proof ingredients: Jordan curve theorem, basic convexity, calculus

Alexeev, M., J. Appl. Probab., 2018
Passing to a (cartoon) model

- state = closed disk
- $2n$ voters equispaced along concentric circle
- voter preferences = iid uniform from $\{\pm 1\}$

**Theorem**

Let $D_n$ denote the random number of majority-positive districts in an optimal partisan gerrymander by complementary half-disks. Then

$$
\Pr(D_n = 2) = \frac{1}{2} - \Theta\left(\frac{1}{\sqrt{n}}\right), \quad \lim_{n \to \infty} \Pr(D_n = 0) = \frac{1}{1 + e^{\pi}}.
$$

**Proof ingredients:** IVT, Donsker’s invariance principle, Brownian bridges

**Upshot:** 73% of seats (on average) from 50% of votes (on average)

Alexeev, M., J. Appl. Probab., 2018
Math can't be helpful if it only considers gerrymandering in the abstract.

Walch, thenib.com/changing-the-math-on-gerrymandering
Beyond the model

Apply gerrymandered version of Smith’s split-line algorithm:

---

Smith, rangevoting.org/SplitLR.html
Alexeev, M., J. Appl. Probab., 2018
Soberón, Notices Amer. Math. Soc., 2017
Beyond the model

Apply gerrymandered version of Smith’s split-line algorithm:

In general, vote majority $\mapsto$ seat unanimity  (ham sandwich theorem)

**Why not impose partisan fairness?**

---

Smith, rangevoting.org/SplitLR.html
Alexeev, M., J. Appl. Probab., 2018
Soberón, Notices Amer. Math. Soc., 2017
Part II

an impossibility theorem
How to measure partisan fairness?

**Proportionality.** prop votes \( \approx \) prop seats
How to measure partisan fairness?

**Proportionality.** prop votes $\approx$ prop seats

**Efficiency gap.** $\#(\text{red wasted votes}) \approx \#(\text{blue wasted votes})$

- voter locations $= A, B \subseteq [0, 1]^2$
- districts $= D_1 \cup \cdots \cup D_k = [0, 1]^2$
- wasted votes by $A$ in $i \in [k]$: $w_{A,i} := \begin{cases} |A \cap D_i| - \left\lceil \frac{1}{2} |(A \cup B) \cap D_i| \right\rceil & \text{if } |A \cap D_i| > |B \cap D_i| \\ |A \cap D_i| & \text{else} \end{cases}$

- efficiency gap:

$$
\text{EG}(D_1, \ldots, D_k; A, B) := \frac{1}{|A \cup B|} \sum_{i=1}^{k} \left( w_{A,i} - w_{B,i} \right)
$$

---

**Districting system.** \( \text{DIST}: (A, B, k) \mapsto (D_1, \ldots, D_k) =: D \)
Technical definitions

Districting system. \( \text{DIST}: (A, B, k) \mapsto (D_1, \ldots, D_k) =: D \)

One person, one vote. \( \exists \delta \in [0, 1), \forall (A, B, k), D = \text{DIST}(A, B, k) \) satisfies
\[
(1 - \delta) \left\lfloor \frac{|A \cup B|}{k} \right\rfloor \leq |(A \cup B) \cap D_i| \leq (1 + \delta) \left\lceil \frac{|A \cup B|}{k} \right\rceil \quad \forall i \in [k]
\]
Technical definitions

**Districting system.** \( \text{DIST}: (A, B, k) \mapsto (D_1, \ldots, D_k) =: D \)

**One person, one vote.** \( \exists \delta \in [0, 1), \forall (A, B, k), D = \text{DIST}(A, B, k) \) satisfies

\[
(1 - \delta) \left\lfloor \frac{|A \cup B|}{k} \right\rfloor \leq |(A \cup B) \cap D_i| \leq (1 + \delta) \left\lceil \frac{|A \cup B|}{k} \right\rceil \quad \forall i \in [k]
\]

**Polsby–Popper compactness.** \( \exists \gamma > 0, \forall (A, B, k), D = \text{DIST}(A, B, k) \) satisfies

\[
4\pi \cdot \frac{|D_i|}{|\partial D_i|^2} \geq \gamma \quad \forall i \in [k]
\]

Alexeev, M., Amer. Math. Mo., 2018
Technical definitions

**Districting system.** $\text{DIST} : (A, B, k) \mapsto (D_1, \ldots, D_k) =: D$

**One person, one vote.** $\exists \delta \in [0, 1), \forall (A, B, k), D = \text{DIST}(A, B, k)$ satisfies

$$(1 - \delta) \left\lfloor \frac{|A \cup B|}{k} \right\rfloor \leq |(A \cup B) \cap D_i| \leq (1 + \delta) \left\lceil \frac{|A \cup B|}{k} \right\rceil \quad \forall i \in [k]$$

**Polsby–Popper compactness.** $\exists \gamma > 0, \forall (A, B, k), D = \text{DIST}(A, B, k)$ satisfies

$$4\pi \cdot \frac{|D_i|}{|\partial D_i|^2} \geq \gamma \quad \forall i \in [k]$$

**Bounded efficiency gap.** $\exists \alpha, \beta > 0, \forall (A, B, k), D = \text{DIST}(A, B, k)$ satisfies

$$|\text{EG}(D_1, \ldots, D_k; A, B)| < \frac{1}{2} - \alpha \quad \text{whenever} \quad ||A| - |B|| < \beta|A \cup B|$$

Alexeev, M., Amer. Math. Mo., 2018
Impossibility

Theorem

There is no districting system that simultaneously satisfies

- one person, one vote
- Polsby–Popper compactness
- bounded efficiency gap

Proof idea: homogeneous mixture of voters

Alexeev, M., Amer. Math. Mo., 2018
Impossibility

Theorem

There is no districting system that simultaneously satisfies

- one person, one vote
- Polsby–Popper compactness
- bounded efficiency gap

Proof idea: homogeneous mixture of voters

Alexeev, M., Amer. Math. Mo., 2018
Beyond the model

<table>
<thead>
<tr>
<th>Party</th>
<th>Candidate</th>
<th>Running Mate</th>
<th>Votes</th>
<th>Percentage</th>
<th>Electoral votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democratic</td>
<td>Hillary Clinton</td>
<td>Tim Kaine</td>
<td>1,995,196</td>
<td>60.0%</td>
<td>11</td>
</tr>
<tr>
<td>Republican</td>
<td>Donald Trump</td>
<td>Mike Pence</td>
<td>1,090,893</td>
<td>32.8%</td>
<td>0</td>
</tr>
<tr>
<td>Libertarian</td>
<td>Gary Johnson</td>
<td>William Weld</td>
<td>138,018</td>
<td>4.2%</td>
<td>0</td>
</tr>
<tr>
<td>Green</td>
<td>Jill Stein</td>
<td>Ajamu Baraka</td>
<td>47,661</td>
<td>1.4%</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td><strong>Totals</strong></td>
<td></td>
<td><strong>3,269,783</strong></td>
<td>100.00%</td>
<td><strong>11</strong></td>
</tr>
</tbody>
</table>

Current representatives

- 1st congressional district: Richard Neal (D) (since 1989)
- 2nd congressional district: Jim McGovern (D) (since 1997)
- 3rd congressional district: Niki Tsongas (D) (since 2007)
- 4th congressional district: Joseph P. Kennedy III (D) (since 2013)
- 5th congressional district: Katherine Clark (D) (since 2013)
- 6th congressional district: Seth Moulton (D) (since 2015)
- 7th congressional district: Mike Capuano (D) (since 1999)
- 8th congressional district: Stephen F. Lynch (D) (since 2001)
- 9th congressional district: Bill Keating (D) (since 2011)
Beyond the model

\[ EG \approx 0.3 \gg 0.08 \]

\[ \text{MA} \notin \{\text{MD, NC, WI, \ldots}\} \]

Political geography explanation?

images from Wikipedia
Part III

fair redistricting is hard
Background

Sometimes, the set of compliant maps is **complicated**

**Constraints for Wisconsin State Assembly:**

- all districts have equal population
- at most 58 counties can be split in different districts
- at most 62 municipalities can be split
- the average Reock score is at least 0.39
- the average Polsby–Popper score is at least 0.28
- at least 6 districts are at least 40% black (among citizens of voting age)
- districts 8 and 9 do not change (previously ordered by a federal court)

How hard is it to find a fair map among compliant maps?
Computational complexity

- NP-complete
  - Hamilton cycle
  - Steiner tree
  - Graph 3-coloring
  - Satisfiability
  - Maximum clique
  - Graph connectivity
  - Primality testing
  - Matrix determinant
  - Linear programming

- NP-hard
  - Matrix permanent
  - Halting problem
  - Factoring
  - Graph isomorphism

- NP
- P

image from quantumgazette.blogspot.com
Fair maps among compliant maps

Compliant maps

▶ all districts have approximately the same population
▶ mild notion of geographic compactness

Fair maps

▶ both parties receive at least some level of representation

Fair maps among compliant maps

Compliant maps
▶ all districts have approximately the same population
▶ mild notion of geographic compactness

Fair maps
▶ both parties receive at least some level of representation

**Theorem**
Deciding whether there exists a fair redistricting among compliant maps is NP-hard.
Proof idea: Reduction from planar 3-SAT

3-SAT: Decide whether there exists a boolean assignment satisfying a formula of the form

\[(\neg x_1 \lor x_2 \lor \neg x_4) \land (\neg x_2 \lor \neg x_4 \lor \neg x_3) \land \ldots\]
Proof idea: Reduction from planar 3-SAT

3-SAT: Decide whether there exists a boolean assignment satisfying a formula of the form

\[(\neg x_1 \lor x_2 \lor \neg x_4) \land (\neg x_2 \lor \neg x_4 \lor \neg x_3) \land \ldots\]

Planar 3-SAT: Bipartite graph with \(V = \{\text{variables, clauses}\}\)

---

Proof idea: Reduction from planar 3-SAT

3-SAT: Decide whether there exists a boolean assignment satisfying a formula of the form

\[ (\neg x_1 \lor x_2 \lor \neg x_4) \land (\neg x_2 \lor \neg x_4 \lor \neg x_3) \land \ldots \]

Planar 3-SAT: Bipartite graph with \( V = \{ \text{variables, clauses} \} \)

Planar 3-SAT is NP-complete

---

Proof idea: Reduction from planar 3-SAT

instance of planar 3-SAT $\rightarrow$ instance of fair redistricting

- Town Pop D R
  - Big: $\frac{L}{2}$ $\frac{2\gamma}{3}L$ $\frac{2\gamma}{3}L$ 0
  - Small: $\frac{L}{2} + \frac{\gamma L}{6}$ $\frac{L}{4}$ $\frac{L}{4} + \frac{\gamma L}{6}$ 0
  - Adjacent: $\frac{L}{2} + \frac{\gamma L}{4}$ $\frac{L}{4}$ $\frac{L}{4} + \frac{\gamma L}{4}$ 0
  - Edge: $\frac{L}{2} - \frac{\gamma L}{4}$ $\frac{L}{4}$ $\frac{L}{2} + \frac{\gamma L}{4}$ 0

Population per district $\in [L, L+\gamma]$.

D wins at most $2k$ districts even with almost half of the vote and Total pop $\gg 2k$
Formula is satisfiable iff D wins $2k$ districts.

Important considerations

Worst-case complexity

- says very little about real-world maps
- identifies limitations of general-purpose redistricting protocols
Political geography can bring tension between shape and fairness

- Sometimes, you can gerrymander with nice shapes
- Sometimes, you need strange shapes to obtain fairness
- Sometimes, a fair map exists, but it’s hard to find
Questions?

**Partisan gerrymandering with geographically compact districts**
B. Alexeev, D. G. Mixon

**An impossibility theorem for gerrymandering**
B. Alexeev, D. G. Mixon

**Fair redistricting is hard**
R. Kueng, D. G. Mixon, S. Villar

**Utility Ghost: Gamified redistricting with partisan symmetry**
D. G. Mixon, S. Villar
arXiv:1812.07377

Also, Google **short fat matrices** for my research blog.