Gerrymandering and Convexity

Jonathan K. Hodge, Emily Marshall, and Geoff Patterson

doi:10.4169/074683410X510317

Jonathan Hodge (hodgejo@gvsu.edu) is an Associate Professor of Mathematics at Grand Valley State University. He is a graduate of Calvin College (B.S., 1998) and Western Michigan University (Ph.D., 2002). His favorite time of the year is summer, when he gets to enjoy the sun and do research with undergraduates through GVSU's summer mathematics REU.

Emily Marshall (emily.a.marshall@vanderbilt.edu) earned her B.A. degree from Dartmouth College and is currently a mathematics graduate student at Vanderbilt University. Her research interests are still very broad and she is excited to develop them further as she gets deeper into the Vanderbilt program. In her free time she enjoys running and exploring the city of Nashville.

Geoff Patterson (geoff@math.hawaii.edu) received his Bachelors of Science in Mathematics with a minor in Statistics at Grand Valley State University, Michigan. He is now pursuing his Ph.D., also in Mathematics, at the University of Hawaii at Manoa. His research interests include applied mathematics and computability.

The American Heritage Dictionary [**1**] defines *gerrymandering* as the act of "dividing a geographic area into voting districts so as to give unfair advantage to one party in elections." Figure 1 illustrates this phenomenon.

Imagine the filled circles represent voters from one party, and the unfilled circles represent voters from an opposing, minority party. Notice that in the district subdivision on the top-left, the party that is dominant in the population at large defeats the minority party by a 5-4 vote in each of the four districts, thus winning all four contested seats. Redrawing the district boundaries, however, changes the outcome significantly. In the most surprising case, the dominant party loses three of the four districts. If this were to occur, the minority party would control a majority of the seats in the legislature. Furthermore, the minority party could maintain control by using their legislative majority to ensure that subsequent redistricting plans would not return seats to the dominant party. The inevitable result of this process is a collection of irregularly-shaped districts whose boundaries are determined not by natural geographic or administrative divisions, but by political expediency. Barack Obama put it this way in his popular book, *The Audacity of Hope* [9, p. 103]: "These days, almost every congressional district is drawn by the ruling party with computer-driven precision to ensure that a clear majority of Democrats or Republicans reside within its borders. Indeed, it's not a stretch to say that most voters no longer choose their representatives; instead, representatives choose their voters."

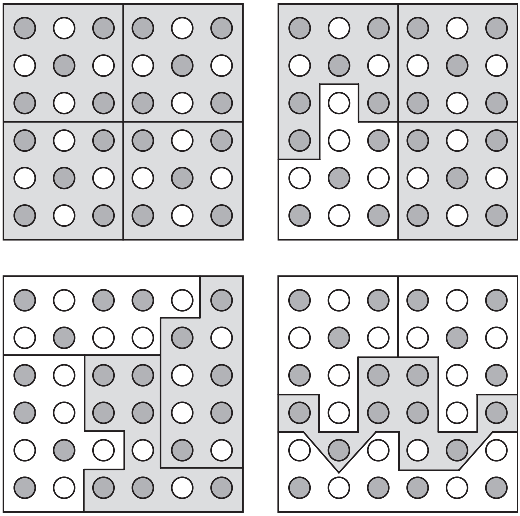

Figure 1. Examples of gerrymandering

Gerrymandering is a hot political topic, but it is also one that involves some interesting mathematics. In particular, the study of gerrymandering has led to the development of a variety of quantitative measures of *shape compactness*, all of which are designed to indicate the degree to which a geometric shape or geographic region meets certain standards of regularity.

Maceachren [**8**] provides a nice summary of several such measures, classifying them into four categories: (i) those that are based on ratios of perimeter to area; (ii) those that compare the characteristics of circles related to the given shape (for instance, inscribed or circumscribed circles); (iii) those that compare the given shape to other standard shapes, such as squares or other polygons; and (iv) those that in some way measure how far the area of the given shape is dispersed from its center. Taylor [**12**] suggests that most measures of shape compactness reflect four essentially distinct characteristics of shape: *elongation*, *indentation*, *separation*, and *puncturedness*. He argues that indentation is the most relevant to gerrymandering and proposes an *indentation index* based on angles formed by the boundary of the shape or district in question.

Chambers and Miller [**6**] maintain that "the sign of a heavily-gerrymandered district is bizarre shape" and that "the most striking feature of bizarrely shaped districts is that they are extremely non-convex." They introduce a convexity-based "measure of bizarreness" and calculate this measure for the congressional districts in Connecticut, Maryland, and New Hampshire. A comparison to the familiar Schwartzberg measure (which essentially computes the ratio of a district's perimeter and area) suggests significant differences between convexity-based and traditional measures of shape compactness.

Here we study a simplified version of Chambers and Miller's measure, which we call the *convexity coefficient*. We present both theoretical results and empirical data based on calculations of the convexity coefficient for all 435 U.S. congressional districts. We then propose modifications to the convexity coefficient to account for irregular state boundaries and non-uniform population distributions. Finally, we present evidence that suggests the potential for non-intuitive outcomes when straight-line divisions are used to divide a state into districts.

The convexity coefficient

Chambers and Miller measure the compactness of a district by calculating the "probability the district contains the shortest path connecting a randomly selected pair of its points." [**6**] Here *shortest path* is interpreted to mean the shortest path *within the state*, so the shortest path may or may not be a straight line. There may also be more than one shortest path between a pair of points, if the state is not simply connected. To account for this, Chambers and Miller calculate the probability that *at least one* of these shortest paths is contained entirely within the district.

We will begin with a simpler measure that uses line segments in lieu of shortest paths.

Definition 1. Let D be a subset of $\mathbb{R}^2$. The *convexity coefficient* of D, denoted $\chi(D)$, is the probability that the line segment connecting two random points in D is contained entirely within D.

Recall that a region within $\mathbb{R}^2$ is said to be *convex* if the line segment connecting any two of its points is contained entirely within the region. Thus, our measure calculates the extent to which a region achieves or fails to achieve the property of convexity.

From the standpoint of gerrymandering, we believe, like Chambers and Miller, that convexity is important. For example, attempts to group voters by common characteristics (such as race or political preference), or to exclude certain groups of voters from a congressional district, often lead to districts with significant indentations and/or geographic separations, such as Illinois' 4th congressional district (see Figure 2). This district combines two geographically separated areas whose populations are mainly Hispanic (74.5%). The western border of the district consists of a portion of Interstate 294 but little of the surrounding area. This clever use of the interstate ensures that the district is pathwise connected, or *contiguous*, a legal requirement in most states.

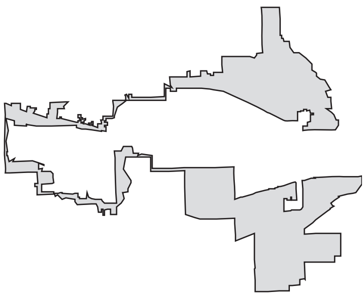

Figure 2. Illinois' 4th congressional district

It is not difficult to estimate an upper bound for the convexity coefficient of a district like this. Note that whenever two points are chosen so that one point is in the northern portion of the district and the other is in the southern portion, the line segment connecting these two points necessarily exits the district. Since roughly half of all such pairs of points fall into this category, it seems that the convexity coefficient can be no more than 0.5. This upper bound is quite generous, as the actual convexity coefficient is approximately 0.237.

Calculation of convexity coefficients

Calculating the exact convexity coefficient can be difficult, even for relatively simple shapes. Consider, for instance, the region K in Figure 3. In order to calculate $\chi(K)$, we introduce a few related definitions.

Definition 2. Let D be a subset of $\mathbb{R}^2$, and let $P = (x, y)$ be a point in D. A point $P' \in D$ is said to be *visible* from P (with respect to D) if the line segment connecting P and P' is contained entirely within D. The *visible region* of P, denoted $V_D(P)$, is the set of all points in D that are visible from P. The area of the visible region of P is denoted $A_D(P)$. If $V_D(P) = D$, then P is said to be a *universal point*. The set of all universal points in D is called the *universal region* of D and denoted U_D.

Figure 3 shows the visible regions for three different points in K. Note that P_1 belongs to the universal region (bounded by the dashed line in the leftmost diagram), whereas P_2 and P_3 do not. Also, if D is a subset of $\mathbb{R}^2$ with area A_D, and P is a point in D, then the probability that a randomly selected point P' in D will be visible from P is given by $A_D(P)/A_D$. The convexity coefficient $\chi(D)$ is simply the probability that two randomly selected points are visible from each other. Thus,

$$\chi(D) = \iint_D \frac{A_D(x, y)}{A_D}\, dx\, dy = \frac{1}{A_D} \iint_D A_D(x, y)\, dx\, dy.$$

Consequently, to calculate $\chi(D)$, one must first find the area of the visible region for each point within D.

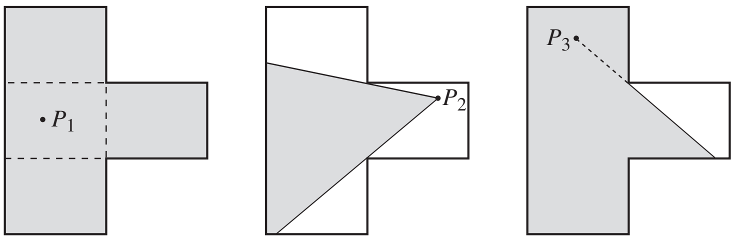

Figure 3. Visible regions of various points

For the region K in Figure 3, there are 10 distinct cases to consider, as the shape of the visible region of a point P depends on the location of P within K. The integral for each of these cases can be calculated fairly easily, and doing so yields a convexity coefficient of approximately 0.875. Now in K, the shorter sides all have the same length. This regularity simplifies the calculations significantly. For regions whose boundaries depend on unspecified parameters, finding a general formula for the convexity coefficient (in terms of the parameters) can be particularly difficult. For such regions, symmetry is an asset.

Consider, for instance, the annulus S in Figure 4, whose convexity is affected by its *puncturedness*. Because the annulus is symmetric about every line passing through the origin, the visible region for any point within the annulus has the same shape. Some relatively straightforward geometry establishes that the area of each point's visible region is given in polar coordinates by

$$A_S(y, \theta) = (R^2 - r^2)\left(\frac{\pi}{2} - \arcsin\left(\frac{r}{y}\right)\right) + R^2 \arccos\left(\frac{r}{R}\right) - r\sqrt{R^2 - r^2}.$$

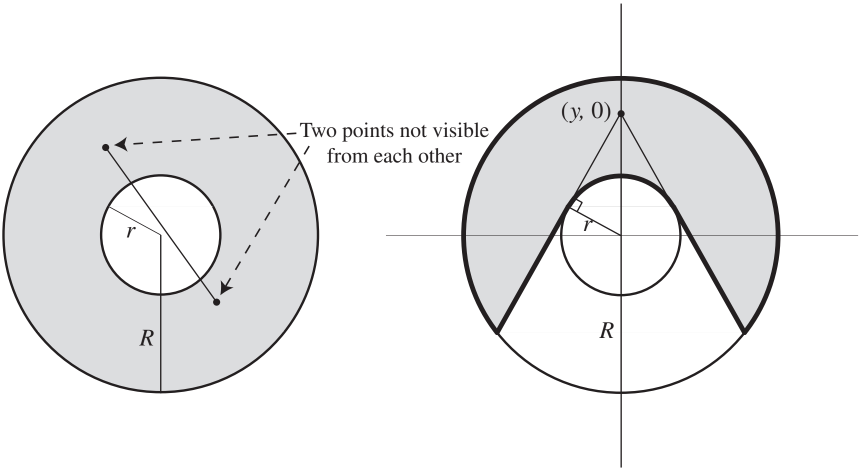

Figure 4. An annulus, and the visible region of a given point within it.

Integrating then yields

$$\chi(S) = \frac{1}{A_S^2} \int_0^{2\pi} \int_r^R A_S(y, \theta)\; y\; dy\; d\theta$$

$$= \frac{\pi R^2 - 2Rr\sqrt{1 - \left(\frac{r}{R}\right)^2} - 2r\sqrt{R^2 - r^2} + 2R^2\left(\arccos(\frac{r}{R}) - \arcsin(\frac{r}{R})\right)}{2\pi(R^2 - r^2)}.$$

Note that the value of $\chi(S)$ is determined entirely by the relative difference between the inner and outer radii. In particular, the larger r/R is, the less convex S is.

Empirical results

Calculating the exact value of the convexity coefficient of a region can be difficult, even for relatively simple shapes. For congressional districts, whose boundaries are often polygons with thousands of sides, it is practically impossible. That said, reasonably good approximations can be calculated easily and efficiently using Monte Carlo methods.

To approximate the convexity coefficient of each of the 435 U.S. congressional districts, we wrote a computer program to choose 10,000 random pairs of points within each district, and for each pair determine if the connecting line segment was contained entirely within the district. In other words, for each district we used 10,000 Bernoulli trials to approximate the probability parameter p of the corresponding binomial random variable. Using basic statistical analysis, we can say with 95% confidence that our approximations have an error of less than 0.01. Table 1 shows the average convexity coefficient for each of the 50 states and the District of Columbia.

Note that the average convexity coefficient varies significantly from state to state. The best state by this measure is Wyoming, with a perfect convexity coefficient. The worst is Maryland, with an average χ of 0.367. In these extreme cases, however, one could argue that the shape of the state itself is the primary factor in determining the convexity coefficient. In particular, Wyoming is a nearly perfect rectangle with a single congressional district, rendering any discussion of gerrymandering moot. On the other

Table 1. Average convexity coefficients by state

Alabama	0.699	Kentucky	0.789	North Dakota	0.999
Alaska	0.714	Louisiana	0.754	Ohio	0.683
Arizona	0.836	Maine	0.674	Oklahoma	0.776
Arkansas	0.813	Maryland	0.367	Oregon	0.787
California	0.646	Massachusetts	0.599	Pennsylvania	0.635
Colorado	0.771	Michigan	0.778	Rhode Island	0.631
Connecticut	0.728	Minnesota	0.868	South Carolina	0.708
DC	0.951	Mississippi	0.801	South Dakota	0.997
Delaware	0.855	Missouri	0.775	Tennessee	0.705
Florida	0.598	Montana	0.987	Texas	0.700
Georgia	0.829	Nebraska	0.894	Utah	0.838
Hawaii	0.641	Nevada	0.719	Vermont	0.949
Idaho	0.811	New Hampshire	0.709	Virginia	0.677
Illinois	0.664	New Jersey	0.584	Washington	0.746
Indiana	0.804	New Mexico	0.831	West Virginia	0.577
Iowa	0.806	New York	0.655	Wisconsin	0.842
Kansas	0.872	North Carolina	0.585	Wyoming	1.000

hand, Maryland has a highly irregular border. In fact, the convexity coefficient of the entire state is approximately 0.3133. Thus, on average, the districts within Maryland are more convex than the state itself.

The same cannot be said for other states. Illinois, for example, is relatively convex ($\chi \approx 0.968$), but many of its districts (see, for example, Figure 2) are not. On average, Illinois' congressional districts are about 30% less convex than the state. This discrepancy suggests that Illinois' low average convexity coefficient is not simply a consequence of a non-convex state boundary. Indeed, a quick glance at the congressional districts of Illinois (see Figure 5) reveals several districts that appear to be highly gerrymandered.

Figure 5. Congressional districts of Illinois

The cover of this JOURNAL shows the 10 least convex U.S. congressional districts, as measured by our convexity coefficient. Consistent with Taylor's classification [**12**], each of these districts exhibits a high degree of elongation, indentation, and/or separation. While some of these irregular features are due to irregular state boundaries, many are not. In the next section, we propose a modification to the convexity coefficient that more clearly distinguishes between naturally occurring irregularities and those that result from gerrymandering.

State boundaries and population data

Unlike our convexity coefficient, Chambers and Miller's measure takes into account non-convex state boundaries by calculating the probability that the district in question contains the shortest path *within the state* between any two randomly selected points. Note that, for non-convex states, the shortest path between two points may not be a straight line.

For example, within Maryland's sixth congressional district, all but the easternmost boundary is state boundary (shown by the darker line in Figure 6). As such, the shortest path within Maryland between the two points shown is not the straight line connecting the two points (dashed in the figure), but rather a path consisting of three distinct line segments. Because the shortest path remains within the district, Chambers and Miller's measure is not negatively affected by the fact that the line segment connecting these two points exits and re-enters the district. Our convexity coefficient, however, does penalize this departure. Consequently, our convexity coefficient for Maryland's 6th district is a low 0.470, while Chambers and Miller's measure is a much higher 0.926.

Figure 6. Two points in Maryland's 6th congressional district

In this example, Chambers and Miller's measure seems to reflect more accurately the shape of the district as it relates to gerrymandering. The eastern boundary of the district is somewhat jagged, but the indentations are relatively small. While some gerrymandering may have occurred, our convexity coefficient gives an artificially low result due to its failure to account for the non-convex state boundary.

To remedy this defect, we modify our original convexity coefficient as follows. We assume that each congressional district and state has a boundary that can be represented by a collection of polygons in $\mathbb{R}^2$, an assumption consistent with the boundary data provided by the U.S. Census Bureau [**4**]. Furthermore, given a collection of boundary polygons, standard algorithms (such as ray casting) can be used to determine if a point is on the interior, boundary, or exterior of the corresponding district or state.

Definition 3. Let D and S each be a set of interior points for a collection of polygons, with $D \subseteq S$. For any points p_1 and p_2 in D, p_2 is said to be *semi-visible* from p_1 (with respect to S) if the line segment connecting p_1 and p_2 intersects the boundary of D only at points that also belong to the boundary of S. The *boundary-adjusted convexity coefficient of* D (with respect to S), denoted $\tilde{\chi}(D)$, is the probability that two random points in D are semi-visible from each other.

With this new definition, the two points in Figure 6 are semi-visible from each other. In fact, most pairs of points in Maryland's 6th district are semi-visible from each other. The boundary-adjusted convexity coefficient (0.959) reflects this fact and is much closer to Chambers and Miller's measure (0.926) than our original convexity coefficient (0.470).

Further discrepancies between our convexity coefficient and Chambers and Miller's measure can be attributed to the fact that the latter takes into account population data. Most measures of shape compactness deal only with the geometry of the district, and not the way the population is distributed within the district. Population, however, is a salient issue in gerrymandering. A highly irregular district boundary is irrelevant if the bulk of the district's population is concentrated in a convex subset of the district's interior. Conversely, small indentations can be strong evidence of gerrymandering if they occur in highly populated areas.

There is an easy way to account for population data in both the original and the boundary-adjusted convexity coefficients: simply choose random census block locations instead of random points. Doing so ensures that points in more populated areas are chosen with higher frequency than those in less populated areas. Thus, boundary irregularities that affect more constituents are penalized more severely than those that affect fewer constituents. We let $\tilde{\chi}_p$ denote the boundary-adjusted convexity coefficient, with population data incorporated as described above.

Table 2 compares $\tilde{\chi}_p$ to Chambers and Miller's measure (denoted CM in the table) for each congressional district in Connecticut, Maryland, and New Hampshire. The original and boundary-adjusted convexity coefficients, without population data, are provided as well. For most districts, $\tilde{\chi}_p$ and Chambers and Miller's measure are quite similar. In fact, of the 105 possible pairwise comparisons among these 15 districts, the two measures give the same relative (i.e., ordinal) ranking 93% of the time. The two measures yield different relative rankings for only 7 district pairs. This is significant, since our mechanism for boundary adjustment is simpler and potentially more efficient than that employed by Chambers and Miller. In particular, it does not require the computation of a shortest path between each pair of points.

Table 2. Convexity coefficients vs. Chambers and Miller's measure

District	χ	$\tilde{\chi}$	$\tilde{\chi}_p$	CM	$\tilde{\chi}_p - \text{CM}$
Connecticut 1	0.417	0.451	0.642	0.609	0.033
Connecticut 2	0.922	0.928	0.876	0.860	0.016
Connecticut 3	0.744	0.807	0.913	0.891	0.022
Connecticut 4	0.829	0.903	0.984	0.977	0.007
Connecticut 5	0.718	0.719	0.509	0.481	0.028
Maryland 1	0.136	0.974	0.908	0.549	0.359
Maryland 2	0.073	0.497	0.397	0.294	0.103
Maryland 3	0.156	0.371	0.325	0.140	0.185
Maryland 4	0.320	0.338	0.389	0.366	0.023
Maryland 5	0.313	0.869	0.542	0.517	0.025
Maryland 6	0.470	0.959	0.924	0.926	-0.002
Maryland 7	0.745	0.760	0.769	0.732	0.037
Maryland 8	0.742	0.810	0.678	0.657	0.021
New Hampshire 1	0.750	0.784	0.809	0.801	0.008
New Hampshire 2	0.671	0.733	0.572	0.561	0.011
Average	0.534	0.727	0.682	0.624	0.058

The difference between the two measures is largest for Maryland's 1st, 2nd, and 3rd congressional districts. The discrepancies may be due to the fact that all three of these districts are disconnected, which poses a potential problem for calculating the shortest path between two points. Chambers and Miller do not specifically describe their convention for dealing with disconnected districts, but it seems reasonable to infer that the mechanism they employ penalizes disconnection more harshly than ours.

It is also worth noting that when $\tilde{\chi}_p$ is calculated for the districts on the cover, all 10 districts experience some improvement from their original convexity coefficients, but only three ultimately exceed a value of 0.5. For Maryland's first congressional district, $\tilde{\chi}_p$ is approximately 0.958, a reflection of the fact that most of the lack of convexity in this district can be attributed to islands and indentations in its western coastal boundary. Likewise, Florida's 18th district and California's 46th district see improvements to 0.688 and 0.675, respectively. These improvements can be attributed to the fact that the original convexity coefficient penalizes the separation caused by islands within the districts, whereas the boundary-adjusted coefficient does not.

Of the six states in Table 1 with the lowest average convexity coefficients (all under 0.6), only Maryland and West Virginia see a significant improvement when state boundaries and population data are taken into account. Maryland's average improves to 0.617, and West Virginia's to 0.756 (from 0.367 and 0.578, respectively). Of the remaining four states, two (Massachusetts and New Jersey) see an increase of less than 0.05, and two (Florida and North Carolina) actually experience a *decrease* (to 0.586 and 0.559, respectively). From this data, we can conclude that the lack of convexity exhibited by the districts within these four states is not due to irregular state boundaries or even irregular district boundaries that affect relatively few constituents. Indeed, in the case of Florida and North Carolina, the fact that $\tilde{\chi}_p < \chi$ suggests that the irregular district boundaries within these states occur in more populated areas and therefore can be attributed at least in part to gerrymandering.

Finally, as one might expect from Figure 5, accounting for state boundaries and population data does not yield a significant improvement in the average convexity coefficient within the state of Illinois. In fact, these factors only increase the convexity coefficient from 0.664 to 0.677.

Straight-line divisions

As we have demonstrated, convexity-based measures of shape compactness can be effective for identifying irregularly shaped and potentially gerrymandered congressional districts. But what if one's goal is not to analyze, but rather to prescribe, a districting plan?

At least one answer to this question is apparent: Any district defined by a sequence of straight-line divisions (that is, by intersecting of one or more half-planes and possibly the state boundary) necessarily yields a boundary-adjusted convexity coefficient of 1. As such, methods like the shortest splitline algorithm proposed by the Center for Range Voting [**5**] are ideal from the standpoint of convexity, other practical critiques notwithstanding. Using straight lines to define district boundaries is natural and intuitive. In fact, one might conjecture that straight-line divisions always yield districts that are more convex than the states they belong to. We believe, however, that this is not the case. Our reasoning is based on the following theorem, which follows easily from Definitions 1 and 2.

Theorem. *Let S be a polygonal region, and let D_1 and D_2 be disjoint subsets of S such that $D_1 \cup D_2 = S$. Let $\chi_1 = \chi(D_1)$, $\chi_2 = \chi(D_2)$, and let χ_{12} denote the proba-*

bility that a random point in D_1 is visible from a random point in D_2. Furthermore let A_1 and A_2 denote the areas of D_1 and D_2 respectively, so that the area of S is equal to $A = A_1 + A_2$. Then

$$\chi(S) = \left(\frac{A_1}{A}\right)^2 \chi_1 + 2\left(\frac{A_1 A_2}{A^2}\right)\chi_{12} + \left(\frac{A_2}{A}\right)^2 \chi_2.$$

Furthermore, if $A_1 = A_2 = \dfrac{A}{2}$, then

$$\chi(S) = \frac{1}{4}(\chi_1 + \chi_2) + \frac{1}{2}\chi_{12}.$$

A consequence of this theorem is that the weighted average of χ_1 and χ_2, with weights $\left(\frac{A_1}{A}\right)^2$ and $\left(\frac{A_2}{A}\right)^2$, can be maximized by minimizing χ_{12}. In other words, to maximize the average convexity coefficient resulting from a single, straight line division of a state, the dividing line should create two districts whose points are minimally visible from each other. Consider, for example, the bowtie-shaped state in Figure 7. A single, vertical cut through the middle of the state yields two districts, each having a convexity coefficient of 1, a significant improvement from the entire state's convexity coefficient of just over 0.5. For this division, χ_{12} is close to zero, thus illustrating the fact that a minimal value of χ_{12} yields a maximal value of χ_1 and χ_2.

Surprisingly, such improvement is not always possible. Let S denote the star-shaped region in Figure 8. Each point on the x-axis defines a unique line that divides S into two equal-area districts. Furthermore, all equal-area dividing lines can be identified with a

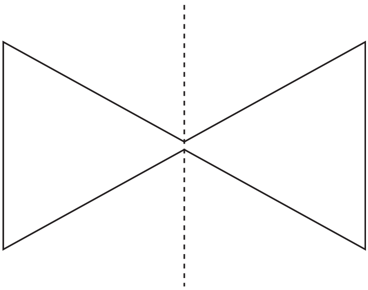

Figure 7. A bowtie-shaped state

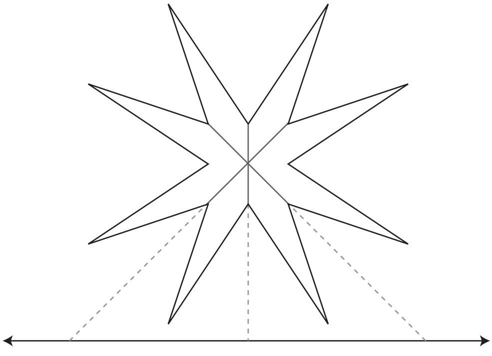

Figure 8. A star-shaped region, with three equal-area straight line divisions

point on the x-axis, with the exception of the horizontal dividing line coinciding with the x-axis, which by symmetry yields the same division as the vertical line coinciding with the y-axis.

Figure 9 shows the approximate values of χ_1, χ_2, χ_{12}, and $\chi(S)$ for the dividing lines corresponding to various x-values. Notice that χ_{12} is always greater than $\chi(S)$, which implies, by the preceding theorem, that the weighted average of χ_1 and χ_2 must always be less than $\chi(S)$. In fact, for each value of x, the corresponding equal-area dividing line yields values of χ_1 and χ_2 that are both less than $\chi(S)$.

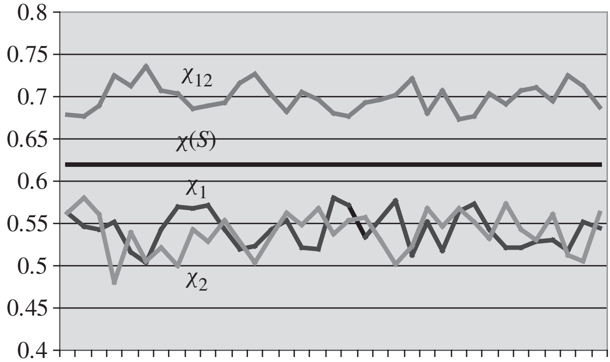

Figure 9. Convexity coefficients for equal-area divisions of a star-shaped region

These data suggest the following conjecture:

> There exists a polygonal region S such that, for every possible straight-line division of S into two equal-area subregions, D_1 and D_2, both $\chi(D_1)$ and $\chi(D_2)$ are less than $\chi(S)$.

We have not been able to prove this conjecture, but the above example seems to support it. If it is true, then apart from the theoretical curiosity of the result, one implication is that comparison of district convexity coefficients to those of their containing states may not be an accurate way to assess whether gerrymandering has occurred. Rather, more sophisticated methods, such as those described in the previous section, should be used to distinguish natural irregularities in state boundaries from evidence of true gerrymandering.

Finally, it should be noted that, just as a low convexity coefficient does not necessarily imply gerrymandering, a high convexity coefficient is not conclusive evidence of its absence. Figure 10 shows that, with some clever choices and the right population distribution, even perfectly rectangular districts can be manipulated to favor one party over another. Vickrey [**13**] gives a similar example. Nonetheless, it seems reasonable to conclude that gerrymandering is considerably more difficult when districts are required to be convex (or nearly so).

Applications and further reading

There are a number of theoretical questions to ask about convexity coefficients; for instance, what types of transformations preserve the convexity coefficient of a given region? There are also numerous practical questions; for instance, how consistent is

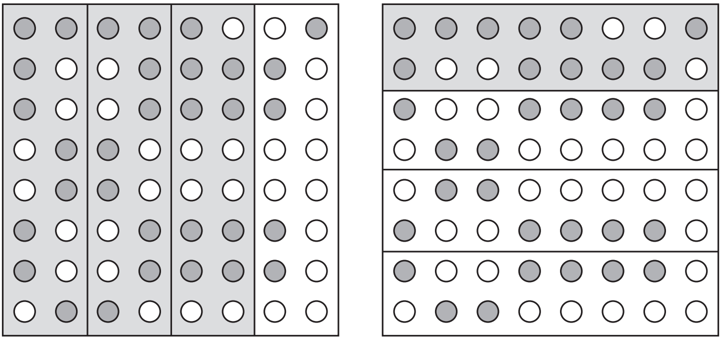

Figure 10. Gerrymandering with convex (rectangular) districts

the convexity coefficient with the way humans intuitively classify shapes according to their regularity? If a group of individuals were asked to rank a set of shapes according to certain attributes associated with geometric regularity, would their rankings be consistent with those induced by the convexity coefficients of the shapes?

In addition to gerrymandering, convexity coefficients would appear to apply to other areas such as fair division, image recognition, and machine vision. For instance, preliminary investigations show significant differences in the convexity coefficients of various letters in the English alphabet. It may be possible, therefore, to use convexity coefficients to help decipher blurred or distorted text, an essential component of applications such as breaking internet captchas [**2**].

We are certainly not the first to study convexity, or to apply it to shape characterization problems. Chambers and Miller's "measure of bizarreness" is similar to that of Žunić and Rosin [**14**]. Schneider [**10**] uses Minkowski addition to measure convexity, devising a measure similar to that of Arrow and Hahn [**3**]. For a more detailed treatment of Minkowski addition, see Schneider [**11**]. For a general introduction to the geometry of convex sets, see Coppel [**7**].

Acknowledgments. This research was supported by the National Science Foundation under Grant No. DMS-0451254, which funds a Research Experiences for Undergraduates (REU) program at Grand Valley State University. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation (NSF).

Summary. Convexity-based measures of shape compactness provide an effective way to identify irregularities in congressional district boundaries. A low convexity coefficient may suggest that a district has been gerrymandered, or it may simply reflect irregularities in the corresponding state boundary. Furthermore, the distribution of population within a district can either amplify or lessen the effects of boundary irregularities. As such, it is essential to take both population data and state boundaries into account when using convexity coefficients to detect gerrymandering. Our boundary-adjusted convexity coefficient provides an efficient way to do so, and it yields results similar to the shortest-path approach taken by Chambers and Miller.

References

1. *The American Heritage Dictionary of the English Language*, 4th ed., Houghton Mifflin, Boston, 2004.
2. E. Aboufadel, J. Olsen, and J. Windle, Breaking the holiday inn priority club CAPTCHA, *Coll. Math. J.* **36** (2005) 101–108. doi:10.2307/30044832

3. K. J. Arrow and F. H. Hahn, *General Competitive Analysis*, Holden Day, San Francisco, 1971.
4. U.S. Census Bureau, Cartographic boundary files; available at http://www.census.gov/geo/www/cob/cd110.html.
5. Center for Range Voting, Gerrymandering and a cure for it—the shortest splitline algorithm (executive summary); available at http://rangevoting.org/GerryExec.html.
6. C. P. Chambers and A. D. Miller, A measure of bizarreness (to appear).
7. W. A. Coppel, *Foundations of Convex Geometry*, Australian Mathematical Society Lecture Series, vol. 12, Cambridge University Press, Cambridge UK, 1998.
8. A. M. Maceachren, Compactness of geographic shape: Comparison and evaluation of methods, *Geografiska Annaler Series B, Human Geography* **67** (1985) 53–67. doi:10.2307/490799
9. B. Obama, *The Audacity of Hope: Thoughts on Reclaiming the American Dream*, Random House, New York, 2006.
10. R. Schneider, A measure of convexity for compact sets, *Pacific J. Math.* **58** (1975) 617–625.
11. ———, *Convex Bodies: The Brunn-Minkowski Theory*, Cambridge University Press, Cambridge, 1993.
12. P. J. Taylor, A new measure for evaluating electoral district patterns, *The American Political Science Review* **67** (1973) 947–950. doi:10.2307/1958636
13. W. Vickrey, On the prevention of gerrymandering, *Political Science Quarterly* **76** (1961) 105–110. doi:10.2307/2145973
14. J. Žunić and P. L. Rosin, A convexity measurement for polygons, in *Proceedings of the 13th British Machine Vision Conference*, 2002, 173–182.

Charles Dodgson intended to publish a book on his voting theory. He never did. He was frustrated by his inability to devise a truly satisfactory system. “A really scientific method for arriving at the result which is, on the whole, most satisfactory to a body of electors, seems to be still a desideratum,” he conceded in December 1877.

—from *Gaming the Vote: Why Elections Aren’t Fair*
by William Poundstone, reviewed on page 339.

Ralph: When she put two potatoes on the table, the big one and the small one, you immediately took the big one without asking me what I wanted.

Norton: What would you have done?

Ralph: I would have taken the small one, of course.

Norton: *(in disbelief)* You would?

Ralph: Yes, I would!

Norton: So, what are you complaining about? You *got* the little one!

—from *The Honeymooners*

suggested by Julius Barbanel

