Order-Reversal of Compactness Scores under Map Projections [DRAFT]

Assaf Bar-Natan   Lorenzo Najt   Zachary Schutzman

December 11, 2018

Abstract

In political redistricting, the *compactness* of a district is used as a quantitative proxy for its fairness. Several well-established, yet competing, notions of geometric compactness are commonly used to evaluate the shapes of regions, including the Polsby-Popper score, the convex hull score, and the Reock score, and these scores are used to compare two or more districts or plans. In this paper, we prove mathematically that, given any *map projection* from the sphere to the plane, that there is some pair of regions whose score is reversed after the projection, for all three of these measures. Finally, we demonstrate empirically the existence of legislative districts whose scores are permuted under the choice of map projection. Our proofs use elementary techniques from geometry and analysis and should be accessible to an interested undergraduate.

1 Introduction

Striving for the *geometric compactness* of legislative districts is a traditional principle of redistricting, and, to that end, many jurisdictions have included the criterion of compactness in their legal code for drawing districts. However, there is no agreed-upon definition for what makes a district compact or not. Several competing mathematical definitions have emerged over the past two centuries, including the *Polsby-Popper score*, which measures the ratio of a district's area to the square of its perimeter, the *convex hull score*, which measures the ratio of the area of a district to the smallest convex region containing it, and the *Reock score*, which measures the ratio of the area of a district to the smallest circle containing it. Each of these measures is appealing at an intuitive level, since they each assign to a district a single scalar value between zero and one, allowing easy comparisons between proposed redistricting plans. Additionally, the mathematics underpinning each is widely understandable by the relevant parties, including lawmakers, judges, advocacy groups, and the general public. None of these measures is perfect, however. For each, it is not difficult to construct a mathematical counterexample for which a human's intuition and the score's evaluation of a shape's compactness differ, such as a circle with slightly perturbed boundary for the Polsby-Popper measure and a very long, thin rectangle for the convex hull measure. Additionally, these scores often do not agree. The long, thin rectangle has a very good convex hull score, but a very poor Polsby-Popper score. These issues are well-studied by political scientists and mathematicians alike [**? ? ?**].

In this paper, we propose a further critique of these measures, namely *sensitivity under the choice of map projection*. Each of the compactness scores named above is defined as a tool to evaluate geometric shapes in the plane, but in reality we are interested in analyzing shapes which sit on the surface of the planet Earth, which is (roughly) spherical. When a shape is assigned a compactness score, it is implicitly done with respect to some choice of map projection. We show, both mathematically and empirically, that this may have serious consequences for the evaluation of compactness. In particular, we define the analogue of the Polsby-Popper, convex hull, and Reock scores on the sphere, and demonstrate that for any choice of map projection, there are two regions, A and B, such that A is more compact than B on the sphere but B is more compact than A when projected to the plane.

2 Preliminaries, Definitions

We begin by introducing the necessary definitions and terminology, as well as a few observations about the mathematical objects of interest which will be of use later. We carefully lay out these definitions so as to align with an intuitive understanding of the concepts and to appease the astute reader who may be concerned with edge cases, geometric weirdness, and nonmeasurability.

Definition 1. A **region** Ω is an non-empty open set together with its boundary such that the region is measurable, its boundary is measurable, and it is connected.

We choose this definition so that concepts of the 'area' and 'perimeter' of a region are well-defined concepts. Throughout, we restrict our attention to the plane $\mathbb{R}^2$ (or $\mathbb{C}$ if one prefers) and the surface of the unit sphere $\mathbb{S}^2$ equipped with the standard measures and metrics. We leave the consideration of other surfaces, measures, and metrics to future work.

Definition 2. A **compactness score function** $\mathcal{C}$ is a function from the set of all regions to the positive real numbers. We adopt the convention that a region with a *higher* compactness score is *more* compact, and this naturally induces a partial order over the set of all regions, where A is at least as compact as B if and only if $\mathcal{C}(A) \geq \mathcal{C}(B)$.

The final major definition we need is that of a *map projection*. In reality, the regions we are interested in comparing sit on the surface of the Earth (i.e. a sphere), but these regions are often examined as being projected onto a flat sheet of paper or computer screen, and which means that the regions drawn in any flat map object are subject to such a projection.

Definition 3. A **map projection** φ is a local diffeomorphism from the sphere to the plane. This means that φ is continuous, φ^{-1} exists and is also continuous, and the image of a region in the sphere is a region in the plane.

Definition 4. We use the word **transformation** [of the plane/sphere] to mean to a diffeomorphism from the plane or sphere to itself. **Linear transformations** of the plane are those parametrized by invertible 2×2 matrices, and **affine linear transformations** are those transformations which are the composition of a linear transformation and a translation of the plane.

Since the image of a region under a map projection φ is also a region, we can examine the compactness score of that region both before and after applying φ, and this is the heart of the problem we address in this paper. We demonstrate, for several standard choices of compactness scores $\mathcal{C}$, that the order induced by $\mathcal{C}$ is different than the order induced by $\mathcal{C} \circ \varphi$ for *any* choice of map projection φ.

For each of the compactness scores we analyze, our proof that no map projection can preserve their order follows a similar recipe. We first use the fact that any map projection which preserves an ordering must preserve the *maximizers* of that ordering, meaning that if Ω is a region for which $\mathcal{C}(\Omega) \geq \mathcal{C}(\Sigma)$ for all regions Σ, then it must at the very least be the case that $\mathcal{C}(\varphi(\Omega)) \geq \mathcal{C}(\varphi(\Sigma))$ if φ preserves $\mathcal{C}$'s ordering.

Using this fact, we can restrict our attention to those map projections which preserve the maximizers in the induced ordering, then argue that any projection in this restricted set must permute the order of scores of some pair of regions.

3 Polsby-Popper

The first compactness score we analyze is the *Polsby-Popper score*, which takes the form of an *isoperimetric quotient*, meaning it measures how much area a region's perimeter encloses, relative to all other regions with the same perimeter.

Definition 5. The Polsby-Popper score of a region Ω is defined to be

$$\mathrm{PP}_X(\Omega) = \frac{4\pi \cdot \mathrm{area}_X(\Omega)}{\mathrm{perim}_X(\Omega)^2}$$

Where X is either the sphere S or the plane $\mathbb{R}^2$, and area_X and perim_X are the area and perimeter respectively of A in X.

The ancient Greeks were first to observed that if Ω is a region in the plane, then $4\pi \cdot \mathrm{area}(\Omega) \leq \mathrm{perim}(\Omega)^2$, with equality if and only if Ω is a circle. This seemingly obvious fact took a long time to prove, and became known as the *isoperimetric inequality* in the plane. This means that $0 \leq \mathrm{PP}_{\mathbb{R}^2}(\Omega) \leq 1$, where the Polsby-Popper score is equal to 1 only in the case of a circle. In particular, the Polsby-Popper score is scale-invariant in the plane. An isoperimetric inequality for the sphere exists, and we state it as the following lemma. For a more detailed treatment of isoperimetry in general, see [**?**], and for a proof of this inequality for the sphere, see [**?**].

Lemma 1. *If Ω is a region on the sphere with area A and perimeter P, then $P^2 \geq A^2 - 4\pi A$ with equality if and only if Ω is a spherical cap.*

A consequence of this is that among all regions on the sphere with a fixed area A, a spherical cap with area A has the shortest perimeter. However, the key difference between the isoperimetric quotient in the plane and on the sphere is that on the sphere, there is no scale invariance.

Lemma 2. *Let S be the unit sphere, and let $\kappa(h)$ be the cap at height h (i.e. the region comprised of all points whose z-coordinate is greater than $1-h$, where we imagine the sphere as being embedded in $\mathbb{R}^3$, centered at the origin) Then $\mathrm{PP}(\kappa(h))$ is a monotonically increasing function of h.*

Proof. Let $r(h)$ be the radius of the circle bounding $\kappa(h)$. We compute:

$$\begin{aligned} 1 &= r(h)^2 + (1-h)^2, \text{ by right triangle trigonometry} \\ &= r(h)^2 + 1 - 2h + h^2 \end{aligned}$$

Rearranging, we get that $r(h)^2 = 2h - h^2$, which we can plug in to the standard formula for perimeter:

$$\mathrm{perim}_S(\kappa(h)) = 2\pi r(h) = 2\pi\sqrt{2h - h^2}$$

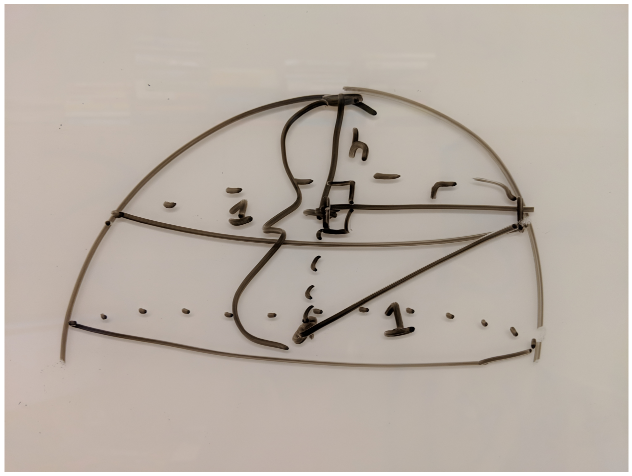


Figure 1: The height and radius of a spherical cap.

We can now use the Archimedian equal-area projection defined by $(x, y, z) \to \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}, z\right)$ to compute $\text{area}_S(\kappa(h)) = 2\pi h$ and plug it in to get:

$$\text{PP}_S(\kappa(h)) = \frac{4\pi(2\pi h)}{4\pi^2(2h - h^2)} = \frac{2}{2-h}$$

Which is a monotonically increasing function of h. $\square$

Corollary 1. *On the sphere, Polsby-Popper scores of caps are monotonically increasing with area*

Using this, we can show the main theorem of this section, that no map projection from the half-sphere to the plane can preserve the ordering of Polsby-Popper scores for all regions.

Theorem 1. *If φ is a map projection from the half-sphere to the plane, then there are two regions A and C in the half sphere such that the Polsby-Popper score of C is greater than that of A in the sphere, but the Polsby-Popper score of $\varphi(A)$ is greater than that of $\varphi(C)$ in the plane.*

Proof. Let φ be a map projection, and take C to be a cap in the half-sphere. Let Σ be a circle in the plane such that $\Sigma \subsetneq \varphi(C)$ and let $A = \varphi^{-1}(\Sigma)$ (see Figure 3).

We now use the isoperimetric inequality for the sphere and Corollary 1 to claim that A does not maximilze Polsby-Popper score in the sphere.

To see this, take $\hat{A}$ to be a cap in the sphere with area equal to that of A. By the isoperimetric inequality of the sphere, $\text{PP}_S(\hat{A}) \geq \text{PP}_S(A)$. Since map projections preserve containment, $\Sigma \subsetneq \varphi(C)$ implies that $A \subsetneq C$, meaning that $\text{area}(\hat{A}) = \text{area}(A) \lneq \text{area}(C)$. By Corollary 1, we know that $\text{PP}_\text{S}(\hat{A}) < \text{PP}_\text{S}(C)$, and combining this with the earlier inequality, we get

$$\text{PP}_\text{S}(A) \leq \text{PP}_\text{S}(\hat{A}) < \text{PP}_\text{S}(C)$$

Since $\Sigma = \varphi(A)$ maximizes the Polsby-Popper score in the plane, but A does not do so in the sphere, we have shown that φ does not preserve the maximal elements in the score ordering, and therefore it cannot preserve the ordering itself. $\square$

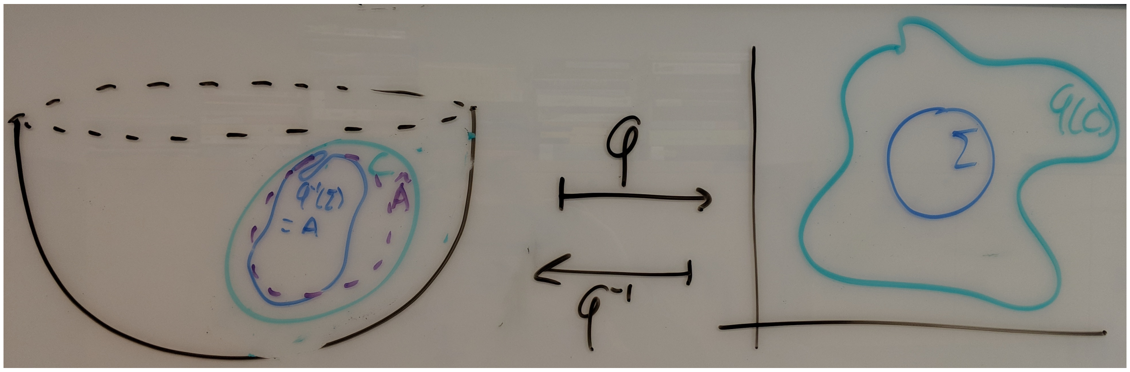


Figure 2: The construction of regions A, C, and $\hat{A}$ in the proof of Theorem 1.

The reason why every map projection fails to preserve the ordering of Polsby-Popper scores is because the score itself is constructed from the *planar* notion of isoperimetry, and there is no reason to expect this formula to move nicely back and forth between the sphere and the plane. This proof crucially exploits a scale invariance present in the plane but not the sphere. If we consider any circle in the plane, it's Polsby-Popper score is definitionally equal to one, but that is not true of every cap in the sphere. This naturally raises the question of whether being more careful, and defining a compactness score which uses the isoperimetric quotient of the surface the region is actually in will evade this problem. We show later that it does resolve the issue of scale-noninvariance, but it is still induces an ordering which is not preserved by any map projection. We address this in Section 6, as it uses on machinery we develop in the interceding sections.

4 Convex Hull

Another commonly used compactness score is the *convex hull score*. We briefly recall the definition of a convex set and then define this score function.

Definition 6. A set X in a metric space is **convex** if for any pair of points x_1 and x_2 in X, the shortest geodesic segment connecting x_1 and x_2 is entirely contained in X. In the plane, these geodesic segments are ordinary line segments. On the sphere, the geodesics are segments of great circles.

Definition 7. Let $\text{conv}(\Omega)$ denote the *convex hull* of a region Ω in either the sphere or the plane, which is the smallest convex region containing Ω. Then we define the *convex hull score* of Ω as

$$\text{CH}(\Omega) = \frac{\text{area}(\Omega)}{\text{area}(\text{conv}(\Omega))}.$$

The convex hull score of a region Ω will be equal to one if and only if Ω is itself convex. We note that the convex hull and Polsby-Popper scores agree when Ω is a circle, provide a similar score when Ω is roughly circluar, and differ greatly if Ω is a highly non-circular but convex region, such as a long, thin rectangle.

Our strategy to demonstrate that any map projection cannot preserve the orders of convex hull scores will be similar to the previous section. We first argue that any order-preserving map must, in particular, preserve the maximizers in the ordering, meaning convex regions on the sphere are mapped to convex regions in the plane. Then, we use this condition to restrict our attention to the map projections which preserve these maximizers, which we can then use to construct our counterexample.

An overview of the structure of the argument is as follows:

1. Any map projection which preserves the ordering of convex hull scores must, in particular, preserve the maximizers in that ordering. This means that such a projection must send convex sets on the sphere to convex sets in the plane.

2. One projection which is convexity-preserving is the gnomonic projection.

3. Affine linear transformations of the plane are also convexity-preserving.

4. *Any* map projection which is convexity-preserving must be the composition of the gnomonic projection and an affine linear transformation.

5. The gnomonic projection does not preserve the ordering of convex hull scores, and any affine linear transformation *does*.

6. Therefore, there is no composition of the gnomonic projection and an affine linear projection which together preserve the ordering of convex hull scores.

Since the convex hull score of Ω is one if and only if Ω is itself a convex region, the maximizers of the induced score ordering are exactly the convex sets in the surface of interest. We can observe that if there is some map projection which preserves the ordering of convex hull scores, it must at the very least map convex sets in the sphere to convex sets in the plane (and vice versa), since convex sets are the maximizers of the convex hull score.

As it turns out, the family of map projections which send every convex set on the lower half-sphere to a convex set in the plane is exactly the set of projections which can be written as a composition of the *gnomonic projection* from the lower half-sphere to the plane followed by an *affine linear transformation*.

Definition 8. The **gnomonic projection** is the map projection from the lower half-sphere to the plane defined by placing the south pole of the sphere tangent to the plane at the origin and sending a point p on the lower half-sphere to the unique point q in the plane which lies on the line in $\mathbb{R}^3$ which passes through the center of the sphere, p and q.

Intuitively, the image of a region on the sphere under the gnomonic projection is the shadow in the plane of that region which is created when we put the bottom of the sphere at the origin and turn on a light source at the center of the sphere.

We can observe that the gnomonic projection sends caps centered at the south pole to circles in the plane centered at the origin, and sends any segment of a great circle passing through the south pole to a line segment in the plane passing through the origin. What may be somewhat surprising is that it sends *every* convex set on the sphere to a convex set in the plane.

Lemma 3. *The gnomonic projection from the lower half-sphere to the plane preserves convex sets. That is, if X is a convex set on the sphere, its image under the gnomonic projection will be a convex set in the plane.*

Proof. To prove this, it suffices to show that the one-dimensional convex sets are preserved, i.e. that any segment of a great circle on the sphere is sent to a line segment in the plane.

A great circle on the sphere is uniquely identified by the intersection of a 'slicing' plane passing through the center of the sphere with the surface of the sphere. Then, for any point p on the great circle, the line passing through the sphere's center and p is contained within this plane. But this is exactly the construction used to find the image of p under the projection, so if p is mapped to the point q, then q is also on this line and therefore in the plane, and q must lie at the intersection of the slicing plane and the projection plane. Since the intersection of two planes is a line, any choice of p on the great circle is sent to a q on this line of intersection, so the image of any segment of the great circle is a segment of a line. □

Using this, we can complete the construction of the main tool of this section – that any map projection which preserves convex sets can be written as the composition of the gnomonic projection and an affine linear transformation of the plane.

Lemma 4. *The map projections from the half-sphere to the plane which map every convex region to a convex region are exactly those which can be written as the composition of the gnomonic projection followed by an affine linear transformation of the plane.*

Proof. By the previous lemma, the gnomonic projection does indeed preserve convex sets. If a map projection preserves convex sets, it must in particular preserve the one-dimensional convex sets, meaning it maps segments of great circles on the sphere to line segments in the plane. Take ψ and φ to convexity-preserving map projections, and, without loss of generality take φ to be the standard gnomonic projection described above. We consider the composition $\psi\varphi^{-1}$, which is a transformation of the plane which preserves convex sets, so in particular it sends line segments to line segments. Since these transformations also preserve containment, we can additionally observe that $\psi\varphi^{-1}$ preserves *collinearity* in the plane, which means it is an affine linear transformation of the plane. Since we can write ψ as $\psi\varphi^{-1}\varphi$, we have successfully written our convexity-preserving map as the composition of the gnomonic projection and an affine linear transformation of the plane. □

This lemma lets us prove our theorem in two steps. We can first show that the gnomonic projection does not preserve the ordering of convex hull scores, then argue that this misordering cannot be corrected by any affine linear transformation of the plane. This first part we will prove by explicitly constructing two regions whose convex hull scores are permuted under the gnomonic projection, and to facilitate this construction, we make the following observation:

Lemma 5. *Let φ be the gnomonic projection and C_θ be a spherical cap centered at the south pole parametrized by the angle θ formed between the central axis of the sphere and a line segment between the center of the sphere and the boundary of the cap (Figure 3). Then $\varphi(C_\theta)$ is a disk in the plane, centered at the origin, and has radius* $\tan(\theta)$*.*

Proof. A (literal) sketch of this proof can be seen in Figure 3.

Since φ projects from the center of the sphere and the sphere's south pole is mapped to the origin in the plane, the image of C_θ is totally radially symmetric about the origin, and is therefore a circle. To see that its radius is $\tan(\theta)$, place the south pole of the sphere tangent to the plane at the origin. By construction, for any point p on the boundary of C_θ, there is a unique line passing through the center of the sphere, p, and the point $\varphi(p)$ on the boundary of the disk in the plane.

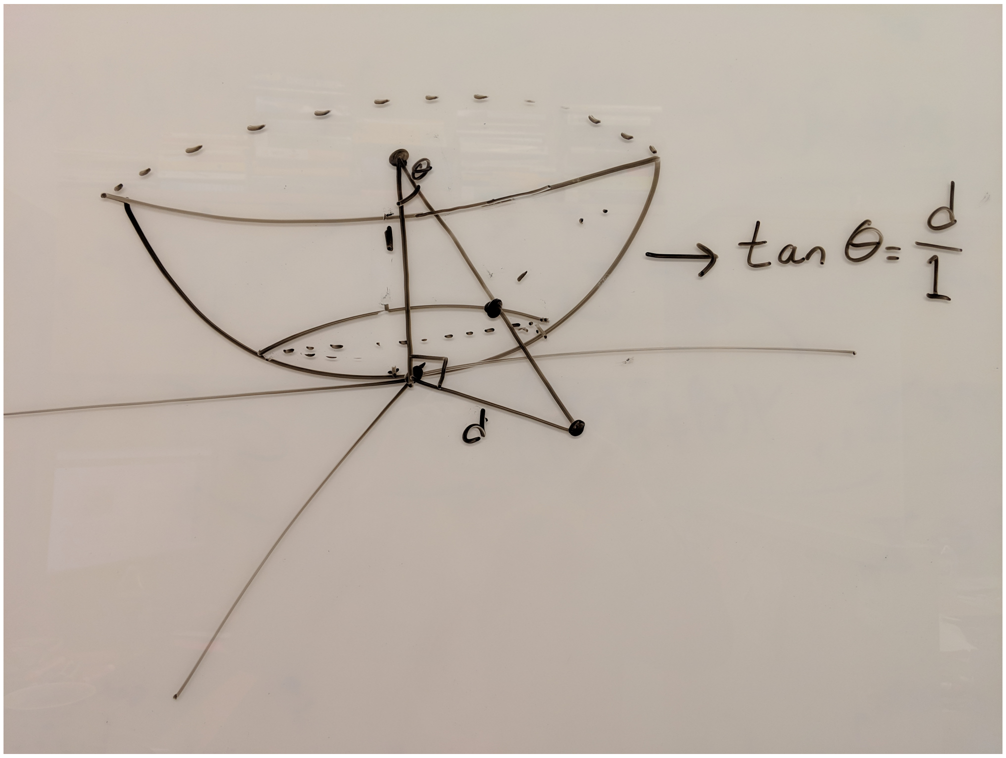


Figure 3: The image of a cap with polar angle θ is a circle of radius $\tan\theta$.

By definition, this line meets the central axis of the sphere at an angle of θ, and the central axis of the sphere meets the plane orthogonally, so the center of the sphere, the origin, and the point $\varphi(p)$ form a right triangle with angle θ. Since we know that the distance between the center of the sphere and the origin is 1, we can write the distance between the origin and $\varphi(p)$ as $\tan(\theta)$.

□

Using this, we can perform the construction of two regions whose convex hull scores are permuted by the gnomonic projection.

Lemma 6. *There exist two regions on the sphere, A and B, such that* $\mathrm{CH}(A) > \mathrm{CH}(B)$ *in the sphere, but, under the gnomonic projection φ,* $\mathrm{CH}(\varphi(A)) < \mathrm{CH}(\varphi(B))$.

Proof. Let A be the region on the sphere defined by taking a cap centered at the south pole parametrized by the angle α_2 and removing the cap centered at the south pole parametrized by the angle α_1, with $\alpha_1 < \alpha_2$. Similarly, let B be the region defined by angles β_1 and β_2. The convex hull score of this kind of region on the sphere is $1 - \frac{\text{area of inner cap}}{\text{area of outer cap}}$. The projection under φ of this kind of region is a disk in the plane with a smaller, cocentric disk deleted. The convex hull score of this kind of region in the plane is $1 - \frac{\text{area of inner disk}}{\text{area of outer disk}}$.

In the previous section, we parametrized a cap on the sphere by its *height*, but we can instead parametrize it by its *polar angle*, which is the angle formed between the polar axis and a line segment connecting the center of the sphere with the boundary of the cap. Using the notation of h for the height of a cap and r for its radius as before, and letting θ be the cap's polar angle, we can use trigonometry to rewrite the area of the cap at height h, $\kappa(h)$. Since $(1-h) = \cos(\theta)$, the area of the cap at height h is $2\pi(1-\cos(\theta))$. Therefore, if we have two cocentric caps on the disk parametrized by angles $\theta_1 < \theta_2$, the convex hull score of this region is

$$1 - \frac{2\pi(1-\cos(\theta_1))}{2\pi(1-\cos(\theta_2))} = 1 - \frac{1-\cos(\theta_1)}{1-\cos(\theta_2)}.$$

In the plane, since the image of a cap centered at the south pole parametrized by angle θ is a circle of radius $\tan^2(\theta)$, and the convex hull score of the image of a pair of cocentric caps parametrized by angles $\theta_1 < \theta_2$ is

$$1 - \frac{\tan^2(\theta_1)}{\tan^2(\theta_2)}.$$

Next, we observe that as we take a cap parametrized by an angle close to $\pi/2$ and increase that angle, the area of the cap grows much more slowly than the area of the disk defined by the cap's projection under φ. This can be shown formally using calculus[1] and intuitively recognized by considering that near $\theta = 0$, $1-\cos(\theta)$ and $\tan^2(\theta)$ are almost equal, but as θ grows and approaches 1, $\tan^2(\theta)$ is much larger than $1-\cos(\theta)$. What this means in the context of the gnomonic projection is that the areas of regions near the south pole of the sphere are not distorted much, but the areas of regions far from the south pole become very large under the projection, and increasingly so the further from the south pole they are.

We will use this observation to construct our example of two regions whose convex hull scores are permuted under φ, since for all such regions constructed as above for a fixed convex hull score on the sphere, those parameterized by larger angles will have comparably worse scores in the plane. We will take two regions whose spherical convex hull scores are near .4, but perturbed slightly such that the one parametrized by larger angles has a slightly lower score. The distortion of areas under the projection will more than compensate for this slight difference.

Let A be defined by the angles (in degrees) $\alpha_1 = 46°$ and $\alpha_2 = 60°$ and B be defined by $\beta_1 = 20°$ and $\beta_2 = 26°$. Then we have $\mathrm{CH}(A) \approx .39$ and $\mathrm{CH}(B) \approx .41$, but $\mathrm{CH}(\varphi(A)) \approx .64$ and $\mathrm{CH}(\varphi(B)) \approx .45$, which is our example of two regions whose scores' order are permuted by φ.

□

Now that we've shown that the gnomonic projection alone cannot preserve convex hull scores, to complete the argument, we must now show that there is no affine linear transformation of the plane which can correct for this.

Lemma 7. *Let L be an affine linear transformation of the plane. Then L preserves the convex hull scores of all figures in the plane.*

Proof. Since affine linear transformations map lines to lines, they preserve convexity, and since they also preserve the containment of regions in other regions, if Ω is a region in the plane, then it must be the case that the convex hull of Ω is mapped to the convex hull of $L(\Omega)$ by L. In other words, the image of the convex hull of Ω is the convex hull of the image of Ω. Finally, since affine linear transformations preserve the ratio of the areas of any two regions, L must, in particular, preserve the ratio of the areas of Ω and its convex hull, so the convex hull score is preserved.

□

[1] The derivative of $1-\cos(\theta)$ is $\sin(\theta)$ and the derivative of $\tan^2(\theta)$ is $2\tan(\theta)\sec^2(\theta)$. The quantity $2\tan(\theta)\sec^2(\theta) - \sin(\theta)$ is positive for $0 < \theta < \pi/2$. We can also observe that as θ approaches $\pi/2$, $\tan^2(\theta)$ grows without bound, but $1-\cos(\theta)$ approaches 1.

We finally have the tools to prove the main result of this section, which is a direct consequence of the four previous lemmas.

Theorem 2. *There is no map projection from the half-sphere to the plane which preserves the ordering of convex hull scores for all regions.*

Proof. Since such a projection must map convex regions on the sphere to convex regions in the plane, it must be a composition of the gnomonic projection followed by an affine linear transformation of the plane. Since affine linear transformations of the plane preserve convex hull scores and therefore preserve their orders, a convex hull score order-preserving projection from the sphere to the plane cannot exist if the gnomonic projection does not preserve their orders. By the counterexample we constructed, it does not, and therefore there is no projection from the sphere to the plane which preserves the ordering of convex hull scores for all regions.

□

5 Reock

Let $\mathrm{circ}(\Omega)$ denote the *smallest bounding circle* (smallest bounding *cap* on the sphere) of a region Ω. Then the *Reock score* of Ω is

$$\mathrm{Reock}(\Omega) = \frac{\mathrm{area}(\Omega)}{\mathrm{area}(\mathrm{circ}(\Omega))}.$$

In words, it is the ratio of the area of the region to the area of the smallest circle containing that region, and it is equal to one if and only if Ω is a circle. We can observe that $\mathrm{CH}(\Omega) \geq \mathrm{Reock}(\Omega)$, since the convex hull of a region is never larger than the smallest bounding circle. This relation holds with equality if and only if the convex hull of Ω is a circle.

The similarity of this measure to the convex hull score suggest a similar tack for proving that no map projection can preserve the ordering of Reock scores, and indeed the structure of the proof is identical. We proceed by demonstrating the following:

1. Any map projection which preserves the ordering of convex hull scores must, in particular, preserve the maximizers in that ordering. This means that such a projection must send caps on the sphere to circles in the plane.

2. One projection which is convexity-preserving is the stereographic projection.

3. Affine linear transformations of the plane which can be written as the composition of a scaling, a rotation, a reflection, and a translation preserve circles, i.e. a *scaled isometry.*

4. The set of scaled isometries is exactly the set of circle-preserving transformations of the plane.

5. Any circle-preserving map projection must therefore be the composition of the stereographic projection and a scaled isometry.

6. The stereographic projection does not preserve the ordering of Reock scores, and any scaled isometry *does.*

7. Therefore, there is no composition of the stereographic projection and a scaled isometry which together preserve the ordering of Reock scores.

Since $\text{Reock}(\Omega) = 1$ if and only if Ω is a circle, we can observe that any candidate for a map projection which preserves the ordering of Reock scores must, at the very least, preserve the maximizeres in the ordering, meaning it sends caps on the sphere to circles in the plane. The *stereographic projection* is one such circle-preserving map projection, and we will show that the family of map projections which are circle-preserving are exactly those which can be written as the composition of the stereographic projection and a *scaled isometry* of the plane.

Definition 9. The **stereograpic projection** from the sphere to the plane is defined by placing the plane tangent to the sphere at the south pole, and for any point p on the sphere other than the north pole, p is sent to the unique point q in the plane which is on the line in 3-space passing through the north pole and p.

If (x, y, z) is a point on the sphere, then the action of the stereographic projection is

$$(x, y, z) \mapsto \left(\frac{x}{1-z}, \frac{y}{1-z} \right),$$

and its inverse, for (u, v) in the plane:

$$(u, v) \mapsto \left(\frac{2u}{1+u^2+v^2}, \frac{2v}{1+u^2+v^2}, \frac{u^2+v^2-1}{1+u^2+v^2} \right).$$

What is somewhat surprising is that this projection sends every cap on the sphere which doesn't pass through the north pole to a circle in the plane. There are several proofs of this fact, and we present a rather straightforward algebraic one here.

Lemma 8. *The stereographic projection sends every cap which does not pass through the north pole to a circle in the plane.*

Proof. We proceed algebraically. Since a cap on the sphere can be identified by the plane in $\mathbb{R}^3$ which intersects the sphere along its boundary, we can parametrize such a cap by writing the equation for its corresponding plane, $ax + by + cz = d$, restricting (x, y, z) to be points on the sphere. The image in the plane of this cap is some set of (u, v) points, and we can explicitly write this by substituting in for x, y, and z the corresponding values for the inverse stereographic projection. We write $\mathcal{W} = u^2 + v^2$ for the ease of presentation.

$$\begin{aligned}
d &= ax + by + cz \\
d &= a\left(\frac{2u}{1+u^2+v^2}\right) + b\left(\frac{2v}{1+u^2+v^2}\right) + c\left(\frac{u^2+v^2-1}{1+u^2+v^2}\right), \text{ by substitution} \\
d &= a\left(\frac{2u}{1+\mathcal{W}}\right) + b\left(\frac{2v}{1+\mathcal{W}}\right) + c\left(\frac{\mathcal{W}-1}{1+\mathcal{W}}\right), \text{ by change-of-variable } \mathcal{W} = u^2+v^2 \\
d\,(1+\mathcal{W}) &= 2au + 2bv + c\,(\mathcal{W}-1)\,, \text{ multiplying through by } 1+\mathcal{W} \\
0 &= (c-d)\mathcal{W} + 2au + 2bv - c - d, \text{ by rearrangement} \\
0 &= (c-d)(u^2+v^2) + 2au + 2bv - c - d, \text{ by change-of-variable } \mathcal{W} = u^2+v^2
\end{aligned}$$

This last line is the equation of a circle in the plane if $c \neq d$ and a line otherwise. Since $c = d$ if and only if the point $(x, y, z) = (0, 0, 1)$ (i.e. the north pole) is on the plane, and we assumed that this is not the case, we have shown that the image under the stereographic projection of every cap which does not pass through the north pole is a circle in the plane.

□

We now have in-hand a circle-preserving map projection, and we now turn to demonstrating that the stereographic projection is essentially the unique map projection with this property. We can observe that if we perform stereographic projection and compose it with a transformation of the plane which sends every circle to a circle, then this composition is a circle-preserving map projection. The next lemma demonstrates that the class of transformations of the plane which preserve circles is actually quite narrow.

Lemma 9. *If T is a transformation of the plane which sends every circle in the plane to another circle, then T must be a linear transformation which is the composition of a scaling, a rotation, a reflection, and a translation, i.e. a scaled isometry.*

Proof. We first argue that if we restrict our attention to *linear* transformations, the scaled isometries are the only transformations which preserve circles.

If T is a scaled isometry with scaling factor α, then, by definition, for any points x and y, the distance between $T(x)$ and $T(y)$ is α times the distance between x and y. Therefore, if we choose a circle Y of radius r and let x be the center, then for any $y \in Y$, $d(T(x), T(y)) = \alpha r$, so $T(Y)$ is a circle of radius r.

Next, if T is a linear transformation which is *not* a scaled isometry, then we can find three points x, y, and z such that $d(T(x), T(y)) = \alpha d(x, y)$ but $d(T(x), T(z)) \neq \alpha d(x, z)$. Furthermore, since linear transformations preserve ratios of lengths of collinear segments, and we have two segments whose ratios of lengths are not preserved, these three points cannot be collinear, so we can consider the circle centered at x and passing through y and z. But, since the ratio of lengths isn't preserved, T distorts this circle, so T is not circle-preserving.

We next need to rule out the existence of a non-linear transformation T which is circle preserving. We proceed by contradiction. Suppose that T is non-linear and circle-preserving. Then, since T is non-linear, there is some line ℓ such that the image $T(\ell)$ is not a line, so we can find three points x, y, and z which are collinear, but $T(x)$, $T(y)$, and $T(z)$ are not. Therefore, we can find a unique circle C_T passing through $T(x)$, $T(y)$, and $T(z)$. Let $T(a)$ be some other point on this circle, and we consider the preimage of this point, a. If a is not on the line ℓ, then we can construct two circles, one passing through a, x, and y and one passing through a, y, and z. These two circles intersect at only two points, a and y, but both of these circles are sent to C_T by T, which is impossible if T is injective. Therefore, a must lie on the line ℓ.

By repeating this argument, any point on the circle C_T must be the image of some point on ℓ. By continuity, we can see that every point on ℓ is sent to a point on C_T, but the homeomorphic image of a line cannot be a circle, so such a non-linear T cannot exist.

Thus, a transformation of the plane is circle-preserving if and only if it is a scaled isometry.

□

We can put these two pieces together to prove the main tool of this section – that any map projection from the sphere to the plane which preserves circles is the composition of the stereographic projection and a scaled isometry.

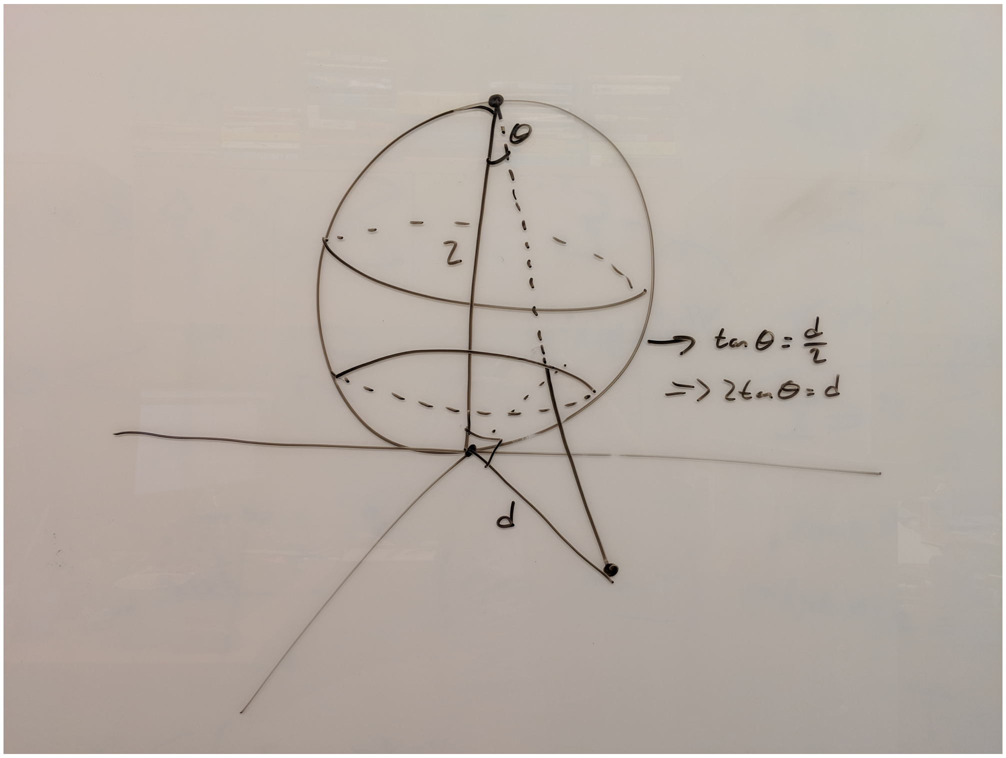


Figure 4: The image of a cap with angle θ between the polar axis and the line of projection is a circle of radius $2\tan\theta$.

Lemma 10. *The map projections from the sphere to the plane which send every cap to a circle are exactly those which can be written as the composition of the stereographic projection followed by a scaled isometry of the plane.*

Proof. Let φ and ψ be two map projections which preserve circles, and without loss of generality let φ be the standard stereographic projection. Then the composition $T = \psi \circ \varphi^{-1}$ is a transformation of the plane which preserves circles, so by the previous lemma, T is a scaled isometry. Then, we can write $\psi = T^{-1}\varphi$, which is the composition of the stereographic projection and a scaled isometry. $\square$

As in the previous section, we can construct an example of two regions whose Reock scores are permuted by the stereographic projection. In fact, we can use *exactly* the same pair of regions as in the convex hull settings, as is made clear by the following lemma.

Lemma 11. *Let φ be the stereographic projection and C_θ be a spherical cap centered at the south pole parametrized by the angle $\theta < \pi/4$ formed between the central axis of the sphere and the line of projection passing through the north pole and the boundary of the cap (Figure 4). Then $\phi(C_\theta)$ is a disk in the plane, centered at the origin, and has radius $2\tan\theta$*

Proof. Since φ projects from the north pole of the sphere and the sphere's south pole is mapped to the origin in the plane, the image of C_θ is totally radially symmetric about the origin, and is therefore a circle. To see that its radius is $2\tan(\theta)$, place the south pole of the sphere tangent to the plane at the origin. By construction, for any point p on the boundary of C_θ, there is a unique line passing through the north pole of the sphere, p, and the point $\varphi(p)$ on the boundary of the disk in the plane. By definition, this line meets the central axis of the sphere at an angle of θ, and the central axis of the sphere meets the plane orthogonally, so the center of the sphere, the origin, and the point $\varphi(p)$ form a right triangle with angle θ. Since we know that the distance between the north pole of the sphere and the origin is 2, we can write the distance between the origin and $\varphi(p)$ as $2\tan(\theta)$.

□

Since the properties of the stereographic projection near the origin is structurally so similar to that of the gnomonic projection, the exact same pair of regions serves as the counterexample in this setting as well. When we examine the Reock scores of the projected regions in the plane, the factors of two cancel out.

We now have a concrete example of the stereographic projection permuting scores, by Lemma 7 in the previous section, since scaled isometries are a special kind of affine linear transformations, they too preserve ratios of areas and therefore Reock scores. We can conclude this section by summarizing the proof of the main result.

Theorem 3. *There is no map projetion from the sphere to the plane which preserves the ordering of Reock scores.*

Proof. Since such a projection must send caps on the sphere to circles in the plane, it must be the composition of the stereographic projection with a scaled isometry of the plane. Since scaled isometries preserve Reock scores and therefore their ordering, a Reock score order-preserving projection from the sphere to the plane cannot exist if the stereographic projection does not preserve the ordering. By the constructed counterexample, it does not, and therefore there is no such Reock score ordering map projection. □

6 Isoperimetric Quotient

We remarked at the end of Section 3 that a more principled version of an isoperimetric quotient may avoid the failure of the ordinary Polsby-Popper score to induce an ordering which can be preserved by some map projection.

Definition 10. We define the **isoperimetric quotient score** of a region Ω to be

$$\mathrm{IPQ}(\Omega) = \begin{cases} \frac{4\pi \ \mathrm{area}(\Omega)}{\mathrm{perim}(\Omega)^2} \text{ in the plane,} \\ \\ \frac{\mathrm{area}(\Omega)^2 - 4\pi \ \mathrm{area}(\Omega)}{\mathrm{perim}(\Omega)^2} \text{ on the sphere} \end{cases}$$

This score properly resolves the issue of scale-noninvariance, and the isoperimetric quotient scores of circles in the plane *and* caps on the sphere are one, regardless of their size.

We'll prove order non-preservation in the same way as before, and we built all of the machinery needed to do it in the previous section. Since any order-preserving map projection needs to preserve the maximizers and caps and circles maximize the isoperimetric quotient score, we can restrict our attention to map projections which send caps to circles. We showed in the previous section that these are exactly the projections which are the composition of the stereographic projection and a

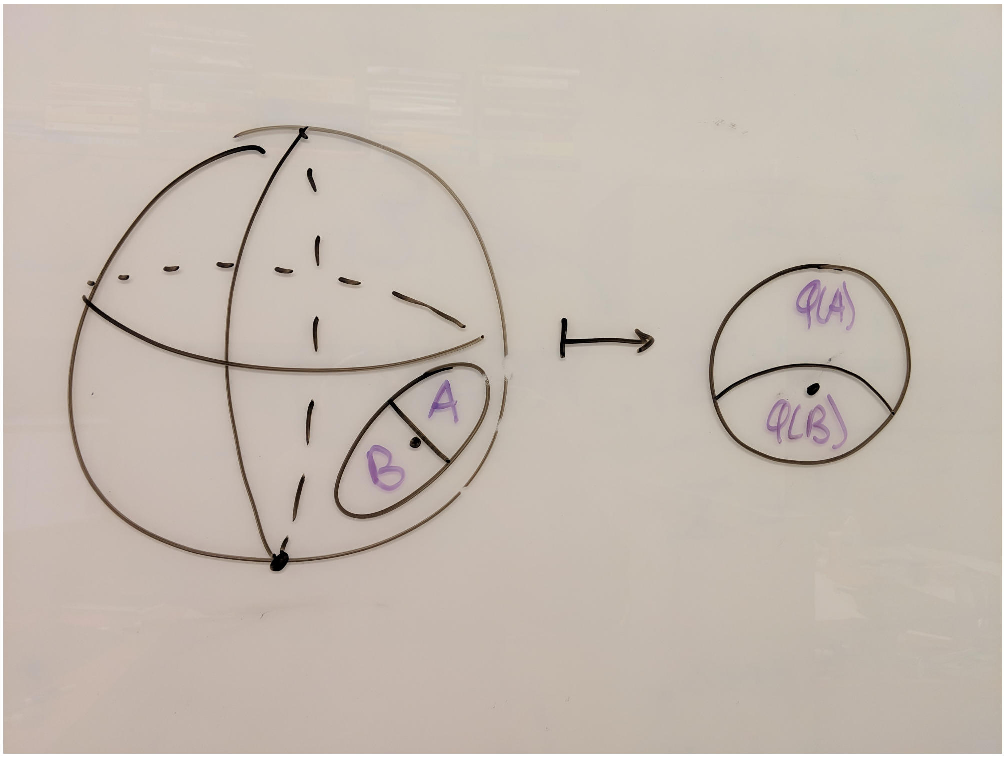


Figure 5: The image of a cap away from the south pole under the stereographic projection.

scaled isometry of the plane. Since this score is scale-invariant, it is preserved by scaled isometries, so we only need to demonstrate the existence of regions on the sphere whose scores are permuted by the stereographic projection.

Theorem 4. *There exist two regions on the sphere A and B such that* $\mathrm{IPQ}(A) > \mathrm{IPQ}(B)$ *but under the stereographic projection φ,* $\mathrm{IPQ}(\varphi(A))) < \mathrm{IPQ}(\varphi(B))$.

Proof. Let C be a cap with its center slightly away from the south pole, and divide it into a 'top' and 'bottom' half with the plane that meets the great circle connecting the poles and cap's center orthogonally. This divides the cap into two pieces with equal area and perimeter, and therefore equal isoperimetric quotient scores. Shift this division line down so that the top part of the circle has a slightly higher than the bottom. We let the top part be A and the bottom be B. The stereographic projection sends caps to circles, but does not preserve the *centers* of such caps away from the south pole, and will send the slightly perturbed division line to a nearly circular arc (Figure 5).

If we had used the original division of the cap into exact halves, then in the plane, the isoperimetric quotient score of $\varphi(A)$ would be strictly less than that of $\varphi(B)$. Since we can choose our perturbation to be small enough to not reverse this order, we have constructed our example of a pair of regions whose isoperimetric score order is not preserved by the stereographic projection.

□

The remainder of the argument follows from our previous results. Lemma 7 says that no scaled isometry can correct for the permutation of score ordering, so there does not exist a map projection which preserves the ordering over regions induced by the isoperimetric quotient score.

7 Empirical Results

The proofs in the previous sections were done with respect to the *unit sphere*. Although mathematically the results remain equally true for a sphere of radius $\mathcal{R}$, a natural question to raise is whether they remain *practically* true in this setting. If we view the idealized Earth as a sphere with $\mathcal{R} = 4000$ miles and observe that the regions we are often interested in measuring are relatively small compared to the total surface area, then it could be the case that any map projection still distorts the scores of the regions, but not by enough to change their ordering. Here, we demonstrate that this is not the case, and that in practice the Polsby-Popper, convex hull, and Reock scores of the real Congresional districts are reordered by the map projections used in practice.[2]

Rather than computing the scores on the sphere, we compare the scores a pair of familiar map projections. Since for any pair of map projections, if the induced score ordering under them is different, then at least one of the two projections must have permuted the score ordering. The projections we compare are the *Mercator* projection and the *Latitude-Longitude* projection. The Mercator projection is a cylindrical projection constructed by placing the sphere in the middle of a cylinder and projecting each point on the sphere outward onto the surface of the cylinder along the line passing through the point and the center of the sphere (see Figure 6).

This cylinder is then 'unrolled' to a flat, planar map. The Latitude-Longitude projection simply uses $(\mathrm{lon}, \mathrm{lat})$ pairs as (x, y) coordinates in the plane. The Mercator and its variants, including the Universal Transverse Mercator and several state plane projections, is among the most commonly used projections. The U.S. Census Bureau provides shapefiles for geographic entities in the United States, including Congressional districts, in a Latitude-Longitude coordinate reference system.[3]

We first compare the three scores under the two projections for all[4] districts in the country. Each point represents one district. The horizontal axis is the ordinal position with respect to the ordering induced by the score under the Latitude-Longitude projection and the vertical axis under the Mercator projection. If both projections were to induce the same score ordering, every point would lie on the 45 degree line up the diagonal.

[2]The code to compute the various compactness scores is based on Lee Hachadoorian's *compactr* project [**?**].

[3]We use the U.S. Census Bureau's shapefile for the Congressional districts for the 115th Congress. These are the districts in place for the 2016 election cycle. Our Mercator projection is the WGS84 Spherical Mercator (EPSG:3857). Our Latitude-Longitude projection is the NAD83 North American Latitude-Longitude (EPSG:4269).

[4]An astute reader will notice these plots have 434, points. We omit the Alaska At-Large Congressional district since it crosses the International Date Line, which makes the convex hull and bounding circle hard to compute in a Latitude-Longitude projection.

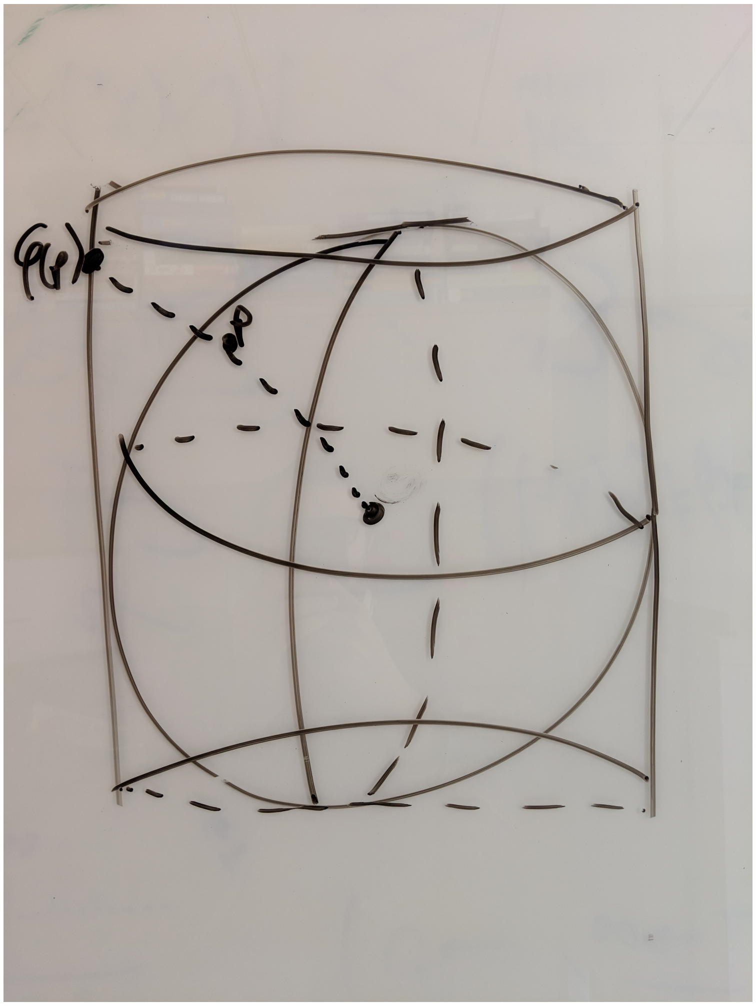
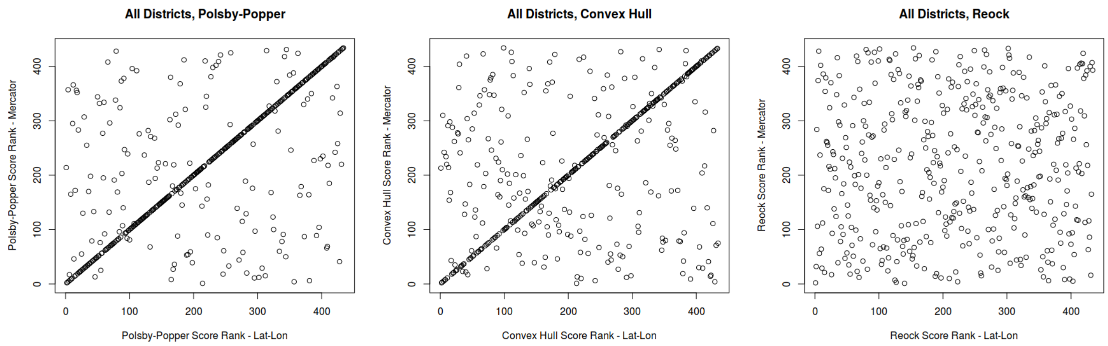


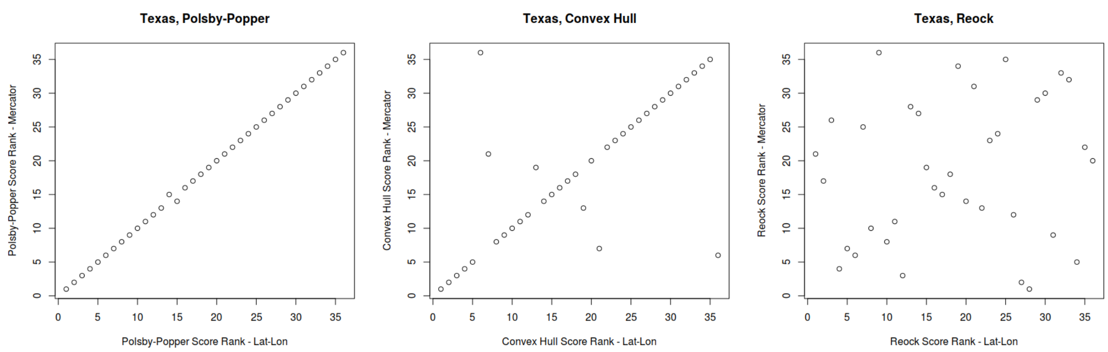
Figure 6: A Mercator projection from the sphere to the surface of the cylinder.



(a) Polsby-Popper (b) Convex Hull (c) Reock

Figure 7: The permutation of compactness scores for all Congressional districts.

While it is an illuminating exercise to demonstrate this effect across all districts, districting plans are proposed and evaluated on a state-by-state basis. In the following plots, we restrict our attention to Texas, and observe a similar phenomenon.



(a) Polsby-Popper (b) Convex Hull (c) Reock

Figure 8: The permutation of compactness scores for Texas' Congressional districts.

We observe overall that the orderings of under the Polsby-Popper score and convex hull score are relatively undisturbed compared to that of the Reock score. This is because the *values* of those scores do not change by too much under different projections. Intuitively, this is because while both projections distort shapes, they do so in a way that does not affect either of these scores by too much, since in the case of the Polsby-Popper score, the perimeter and area of the regions are changed in similar ways, and in the case of the convex hull score, the area of the convex hull of a region is distorted in the same way as the region itself, so the ratio of the areas is similar across projections.

However, the Reock score specifically requires constructing a minimum bounding circle, and this circle can be dramatically different under different map projections. The Mercator projection sends caps on the sphere which aren't too close to the poles to regions in the plane which are very nearly circular, while the Latitude-Longitude projection visibly stretches all but the smallest circles on the equator. For regions away from the equator, such as U.S. Congressional districts, this difference is significant enough to permute the ordering of Reock scores.

8 Conclusion and Future Directions

We demonstrate here the failure of a standard panel of compactness measures to provide a consistent score ordering over all regions. While compactness scores are not used critically in a *legal* context, they appear frequently in the popular discourse about redistricting issues and frame the perception of the 'fairness' of a plan. For example, a 2014 Washington Post piece [**?**] describes an algorithm which generates highly compact districts because it ignores all of the social and demographic data which are crucial to the process. The equating of 'solving' gerrymandering with generating highly compact districts presupposes that the mathematics used to evaluate the geometric features of districts are unbiased and unmanipulable, and we demonstrate here that this is certainly not the case.

The results here suggest that more nuanced models of compactness are worth exploring. Graph-theoretic (i.e. 'discrete') compactness measures as in [**? ?**] are computed without reference to any particular metric embedding, so they definitionally cannot be affected by the choice of map projection. Multiscale measures of compactness as in [**?**] provide a higher resolution view of the geometry of regions. The authors there leave open the question of to what degree a region's total variation profile is unique, and a strongly positive result to that problem would suggest that you can't "hide" any of the geometry of a region using the choice of map projection.

This work opens several promising avenues for further investigation. We prove strong results for the most common compactness scores, but the question remains what the most general mathematical result in this domain might be, such as giving a set of necessary and sufficient conditions for a compactness score to not induce a permuted order for some choice of map projection. Generalizing these results to other surfaces and non-standard measures (such as weighting by population) may also be an avenue for examination.