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

Assaf Bar-Natan Lorenzo Najt Zachary Schutzman

February 23, 2019

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*, the *convex hull score*, and the *Reock score*. 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[citation needed] 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[citation needed] 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 Polsby-Popper

The first compactness score we analyze is the *Polsby-Popper score*, which takes the form of an *isoperimetric quotient*:

Definition 1. 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. 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.*

Lemma 2. *Let $\mathbb{S}^2$ be the unit sphere in $\mathbb{R}^3$, and let $\kappa(h) = \{(x, y, z) \in \mathbb{S}^2 : z \geq h\}$ be a cap. 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:

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

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 $\mathrm{area}_S(\kappa(h)) = 2\pi h$ and plug it in to get:

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

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.

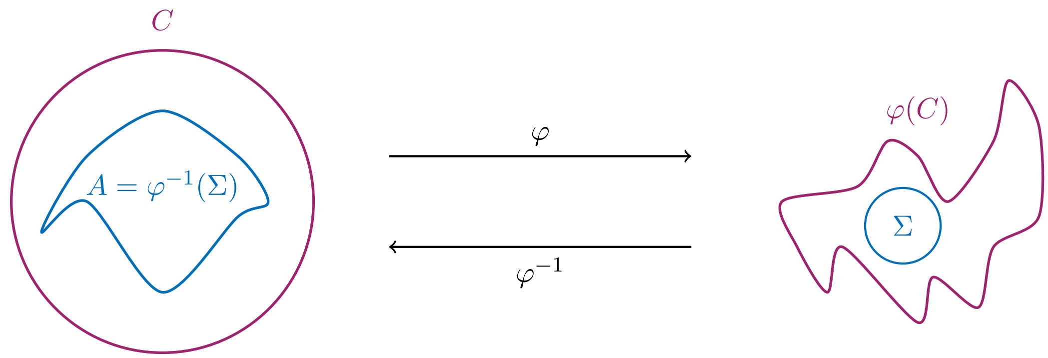

Figure 1: The construction of regions A, C, and $\hat{A}$ in the proof of Theorem 5.

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, $\mathrm{PP}_S(\hat{A}) \geq \mathrm{PP}_S(A)$. Since map projections preserve containment, $\Sigma \subsetneq \varphi(C)$ implies that $A \subsetneq C$, meaning that $\mathrm{area}(\hat{A}) = \mathrm{area}(A) \lneq \mathrm{area}(C)$. By Corollary 1, we know that $\mathrm{PP}_{\mathrm{S}}(\hat{A}) < \mathrm{PP}_{\mathrm{S}}(C)$, and combining this with the earlier inequality, we get

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

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. □

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.

3 Convex Hull

A 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 2. Let $\mathrm{conv}(\Omega)$ denote the *convex hull* of a region Ω in either the sphere or the plane, which is the smallest convex region containing Ω. We define the *convex hull score* of Ω as

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

Throughout this section, let $\varphi : \Omega \to \mathbb{R}^2$ be a map projection defined on a region $\Omega \subset \mathbb{S}^2$.

Lemma 3. *If φ preserves $\mathcal{C}_{CH}$, then the following must hold:*

1. *φ, φ^{-1} send convex sets to convex sets*
2. *φ is a geodesic map.*
3. *There exists a region $U \subset \Omega$ such that for any regions $A, B \subset U$, A and B have equal area if and only if $\varphi(A)$ and $\varphi(B)$ have equal area.*

Proof. 1. This follows because φ, φ^{-1} must preserve maximizers of CH.

2. Assume that φ sends a geodesic segment γ to a nongeodesic segment η. Then for any $\varepsilon > 0$, the ε-neighbourhood of γ $N_\varepsilon(\gamma)$ is convex, but for sufficiently small ε, $\varphi(N_\varepsilon(\gamma)$ is not, a contradiction. Showing that φ^{-1} sends geodesic segments to geodesic segments is proved analogously.

3. Let $U \subset \Omega$ be a cap, and let $A, B \subset U$ be regions of equal area. Let $X = U \setminus A$ and $Y = U \setminus B$, and note that $CH(X) = CH(Y)$. Since φ preserves $\mathcal{C}_{CH}$, we must have $CH(\varphi(X)) = CH(\varphi(Y))$. However, since $\varphi(U)$ is convex, it follows that

$$\frac{\text{Area}(\varphi(U)) - \text{Area}(\varphi(A))}{\text{Area}(\varphi(U))} = CH(Y) = CH(X) = \frac{\text{Area}(\varphi(U)) - \text{Area}(\varphi(B))}{\text{Area}(\varphi(U))}$$

meaning that $\text{Area}(\varphi(A)) = \text{Area}(\varphi(B))$. Applying the same argument to φ^{-1} proves the result.

□

Theorem 2. *There does not exist a map projection satisfying the conditions in Lemma 6*

Proof. Assume that such a map, φ, exists, and restrict it to U as above. Let $T \subset U$ be a spherical equilateral triangle with $\text{Diam}(T) < \frac{\text{Diam}(U)}{3}$, centered at the center of U. Let T_1 and T_2 be two congruent triangles meeting at a point and each sharing a face with T

TODO:picture here

Claim 1. *The image of $T \cup T_i$ is a parallelogram for any i.*

Proof. Set $i = 1$, and note that $T \cup T_1$ is a convex spherical quadrilateral. By symmetry, its geodesic diagonals D_1, D_2 on the sphere split it into four equal-area triangles.

TODO: picture here

Since φ is a geodesic map, it must send $T \cup T_1$ to an Euclidean quadrilateral Q whose diagonals are $\varphi(D_1)$ and $\varphi(D_2)$. Since φ sends equal area regions to equal area regions, it follows that the diagonals of Q split it into four equal area triangles, meaning that Q must be a parallelogram. □

By the claim, the image of $T \cup T_1 \cup T_2$ must consist of two parallelograms, one of whose edges is the diagonal of the other. In other words, $\varphi(T \cup T_1 \cup T_2)$ is a quadrilateral trapezoid whose boundary consists of four Euclidean geodesic segments.

However, since all the angles of T_i and T must be at least $\frac{\pi}{3}$, it follows that the configuration $T \cup T_1 \cup T_2$ must consist of at least five geodesic segments, and hence φ cannot be a geodesic map. □

4 Reock

4.1 The Stereographic Projection

Before analyzing Reock scores, we preform a preliminary discussion about the stereographic projection.

Definition 3. The **stereograpic projection** from the north pole N is defined by placing a copy $\mathbb{R}^2$ tangent to the sphere at the south pole, and sending any $p \in \mathbb{S}^2 \setminus \{N\}$ to the unique point q in $\mathbb{R}^2$ such that $q \in \overline{Np}$.

Remark 1. It is a classical result that this projection sends every cap on the sphere which doesn't pass through the north pole to a circle in the plane.

Lemma 4. *The stereographic projection sends every cap which does not pass through the north pole to a circle in the plane.*

Lemma 5. *Let $f : \Omega \to V$ be a bijection between two planar regions sending generalised circles to generalised circles. Then f is a Möbius transformation composed with a reflection.*

This is a slight rewording of a celebrated result by Carathéodory[**?**]:

Theorem 3. *Every arbitrary bijection of a disc D to a bounded set D' by which circles lying completely in D are transformed into circles lying in D' must be of the form M or $\bar{M}$, where M is a Möbius transformation.*

The proof of the lemma then follows:

Proof. By Carathéodory's theorem, we know that on any open ball in Ω, $f = M$ or $f = \bar{M}$, where M is a Möbius transformaion. Thus, f or $\bar{f}$ are holomorphic, so by applying the identity theorem, it follows that $f = M$ or $f = \bar{M}$ on the entirety of Ω. □

This gives us the following characterization:

Theorem 4. *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 Möbius transoformation and a reflection.*

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 $M^{-1} = \psi \circ \varphi^{-1}$ is a transformation of the plane which preserves circles, so by the previous lemma, M^{-1} is a Möbius transformation composed with a reflection. Then, we can write $\psi = M\varphi$, which is the composition of the stereographic projection and a scaled isometry. □

4.2 The Reock Compactness Score

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

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

Since $\text{Reock}(\Omega) = 1$ if and only if Ω is a circle, we can observe that any candidate φ 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.

By Theorem 8, it follows that we can assume that $\varphi = R \circ M \circ \psi$, where R is a reflection, M is a Möbius transformation, and ψ is the standard stereographic projection. However, we note that reflections do not change any compactness scores, so we only need to deal with $M \circ \psi$.

We now treat the Möbius transformation composition in the following manner:

Claim 2. *There exists a rigid transformation of the sphere, T_M, such that $M\psi = T_M\psi$.*

This follows by writing T_M

In other words, to show that $M\psi$ does not preserve Reock score rankings in some region U, it suffices to show that ψ does not preserve Reock score rankings in the transformed region $T(U)$.

Theorem 5. *Let A be a region on the sphere. Then there exist two regions $\kappa'_N, \kappa'_S \subset A$ such that the Reock scores of κ'_N and κ'_S are equal on the sphere, but under the stereographic projection φ, the Reock score of κ'_S is strictly greater than that of κ'_N.*

Proof. First, since A is a region, we can find some cap κ inside of A such that the center of κ is not the south pole of the sphere, and the north pole is exterior to κ. This cap is bisected by a line of longitude, in particular the great circle which passes through the two poles and the center of κ. This line of longitude meets κ at exactly two points, and we call p_N the point closer to the north pole and p_S the point closer to the south pole.

Choose some r strictly less the radius of κ and let κ'_S be the region constructed by deleting a cap of radius r from the interior of κ tangent at p_S. Construct κ'_N analogously by deleting a cap of radius r tangent at p_n. Observe that the Reock scores of κ'_N and κ'_S are identical, since each has the boundary of κ as its smallest bounding cap and both figures have the same area.

We now consider the images of κ'_N and κ'_S under the stereographic projection φ. Since φ preserves points of tangency, containment, and sends every cap away from the north pole to a circle, the images $\varphi(\kappa'_N)$ and $\varphi(\kappa'_S)$ are both regions which are disks with a smaller disk, tangent to a point on the circumference, deleted. Furthermore, since the boundary of κ was the smallest bounding cap of both regions on the sphere, the image of the boundary of κ under φ is the smallest bounding circle of the images of these regions in the plane.

We can now observe that these two regions in the plane do not have the same Reock score. Both have the same bounding circle, but $\varphi(\kappa'_N)$ is strictly smaller than $\varphi(\kappa'_S)$. This is because the stereographic projection distorts areas in a way such that figures further from the south pole have their areas magnified more than the same region closer to the south pole. If the regions in question are sufficiently small, letting θ denote the polar angle at which we consider a small cap on the sphere, a cap of radius r will be sent to a circle with radius roughly $r/\cos^2\theta$[Zach: check this] , and a straightforward examination of this as a function of θ shows that it grows faster as θ increases.

Since $\varphi(\kappa'_N)$ fills a smaller fraction of the bounding circle than $\varphi(\kappa'_N)$ does, its Reock score is strictly worse, and the stereographic projection does not preserve Reock scores. □

Piecing this together yields:

Theorem 6. *There is no map projetion from the sphere to the plane which preserves the ordering of Reock scores.*