Stereographic Projection

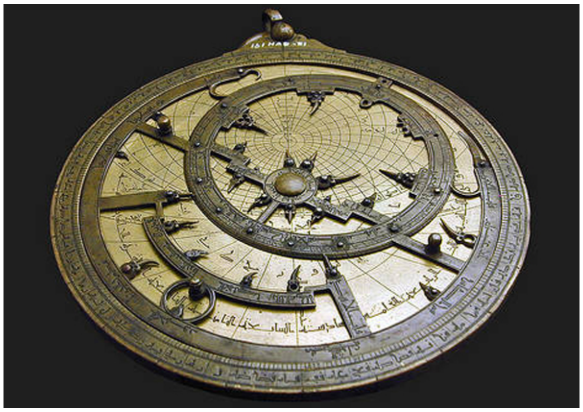
Patch Kessler, December 12, 2018

Stereographic projection is a way to flatten out the surface of a ball, such as the *celestial sphere*, which surrounds the earth and contains all the stars in the sky.

The celestial sphere gives order out of chaos. It gives a way to predict star locations at different times, as well as the time of night from the stars. Here is a physical model of the celestial sphere from around 1480.

1480-1481, http://www.artfund.org/artwork/3600/spherical-astrolabe

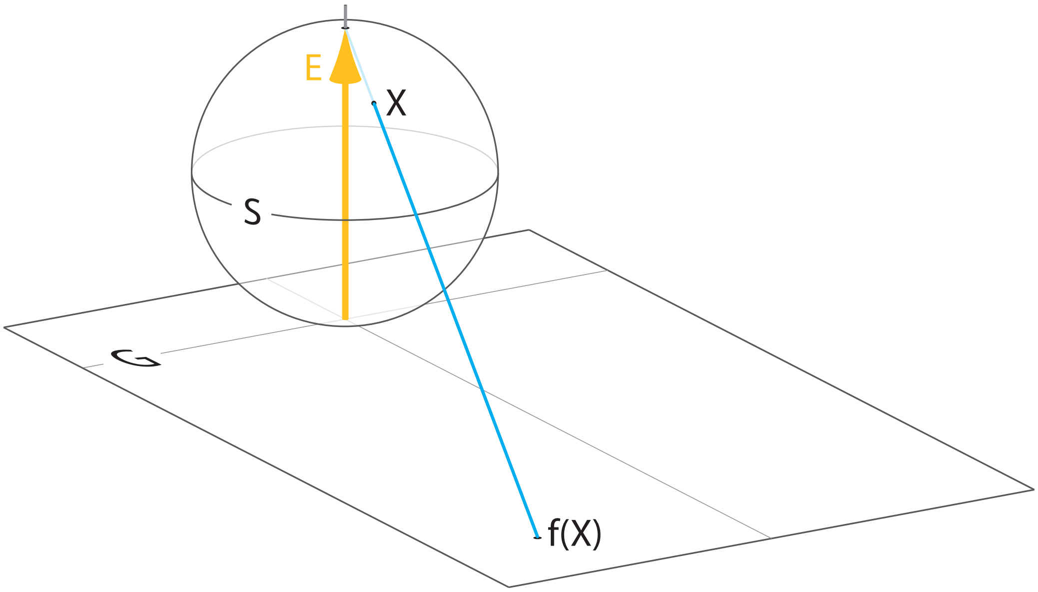
According to legend, Ptolemy (AD 90 - AD 168) thought of stereographic projection after a donkey stepped on his model of the celestial sphere, making it flat. Applying stereographic projection to the celestial sphere results in a planar device called an astrolabe.



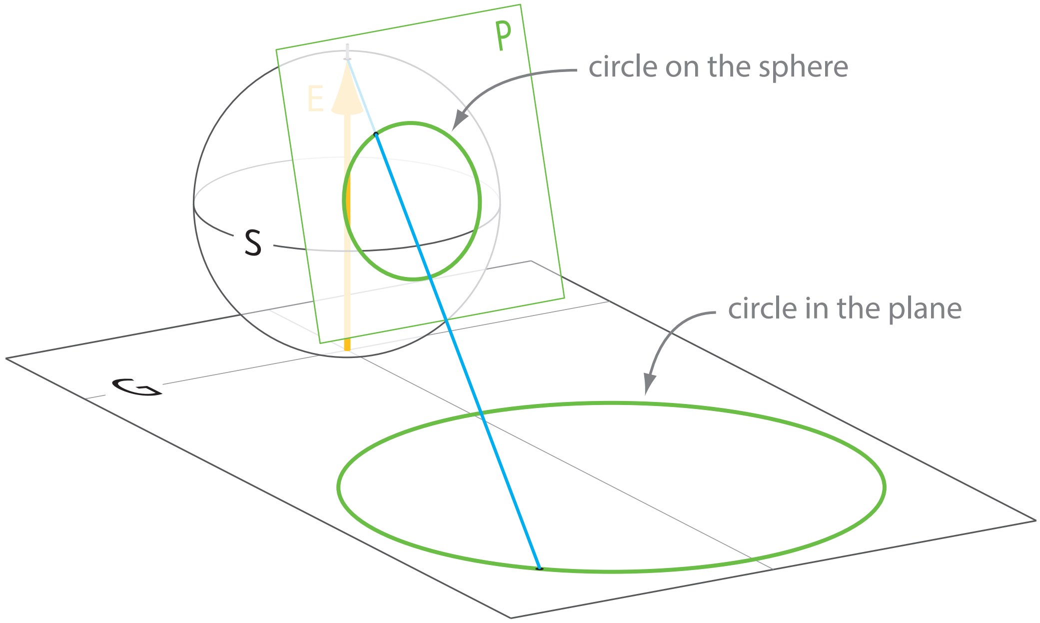
16th century islamic astrolabe

The tips on the perforated upper plate correspond to stars, while points on the lower plate correspond to viewing directions. For instance, the point on the lower plate that the circles are shrinking towards corresponds to the viewing direction directly overhead (i.e., looking straight up).

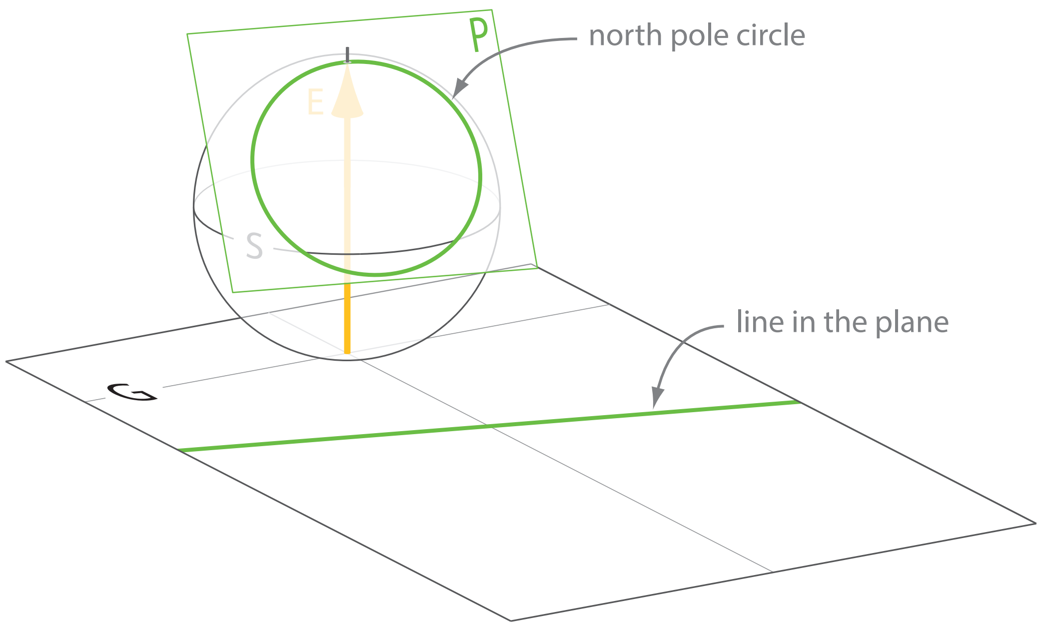
The stereographic projection of a point X on a sphere S is obtained by drawing a line from the north pole of the sphere through X, and continuing on to the plane G that the sphere is resting on. The projection of X is the point $f(X)$ where this line hits the plane.



One of the most surprising things about stereographic projection is that circles get mapped to circles.



Circles that pass through the north pole get mapped to lines, however if you think of a line as an infinite radius circle, then you can say that *circles get mapped to circles* in all cases.

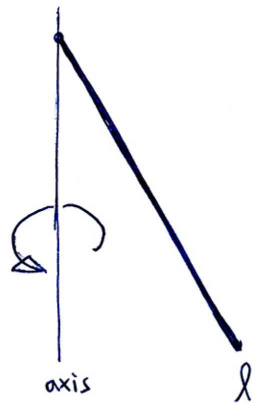


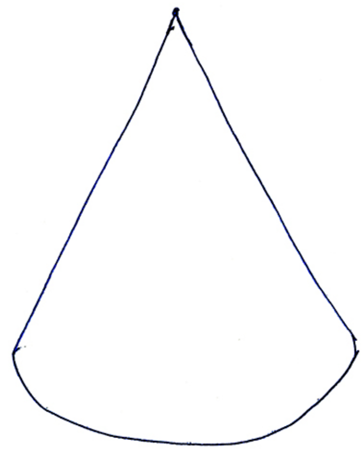
The case of a north pole circle is straightforward to prove. Simply extend the cutting plane P until it intersects the ground plane G. That's it! The stereographic projection of any point on the north pole circle comes from a line within P, and of course the intersection of P and G is a line (or infinite radius circle).

The opposite direction is also easy. Consider any line in the ground plane G. Let P be the plane that contains this line as well as the north pole of the sphere. The intersection of P and the sphere is the circle which gets mapped to the line in G.

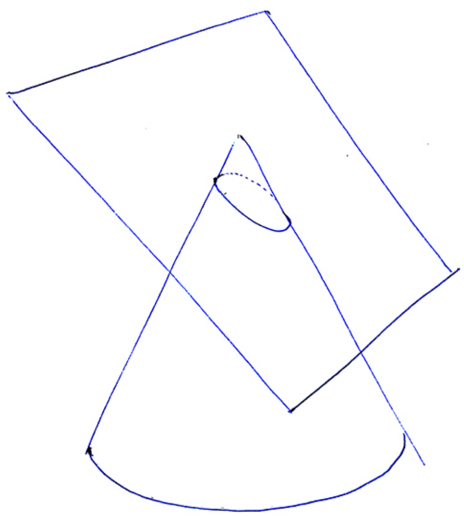
conic sections

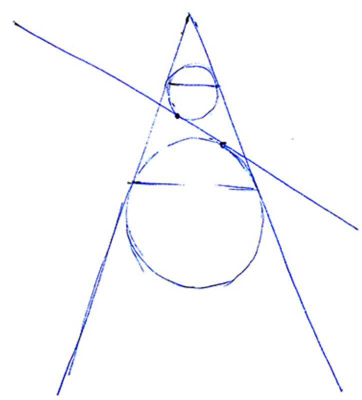
The general *circles to circles* result is surprising because you wouldn't expect a cone with a circular section in one direction to have a circular section in a different direction as well. To see how this is possible, we start with a right circular cone, created by rotating one line about another.





Several properties are obvious, like the fact that horizontal sections of the cone are circles. Other properties are less obvious, like the fact that non horizontal sections can be elliptical. I recently learned of a way to see this using so called *Dandelin*[1] spheres, which is so simple and beautiful that I can't help showing it here.





The Dandelin sphere above the plane is tangent to the plane at a point, and tangent to the cone along a circle. This sphere can be used to build two right triangles for any point x on the cone section. Both triangles have a shared hypotenus from x to the center of the sphere, followed by edges of equal length (equal to the sphere radius), followed by a right angle due to a tangency condition. It follows that these are identical right triangles. The last edge on one of the triangles goes from x to where the sphere touches the plane, while on the other it goes within the cone surface from x to where the sphere touches the cone.

Repeat this construction using a Dandelin sphere *below* the plane. For any x on the intersection curve, the sum of the distances from x to each of the sphere intersection points (the ellipse foci) is equal to the length l of the line segment on the cone that passes through x and is bounded by the two circles. These line segments are clearly all the same length.

[1]Dandelin, Germinal Pierre. 1822. See also *Measurement* by Paul Lockhart 2014, as well as Grant Sanderson's video from 2014 (3blue1brown).

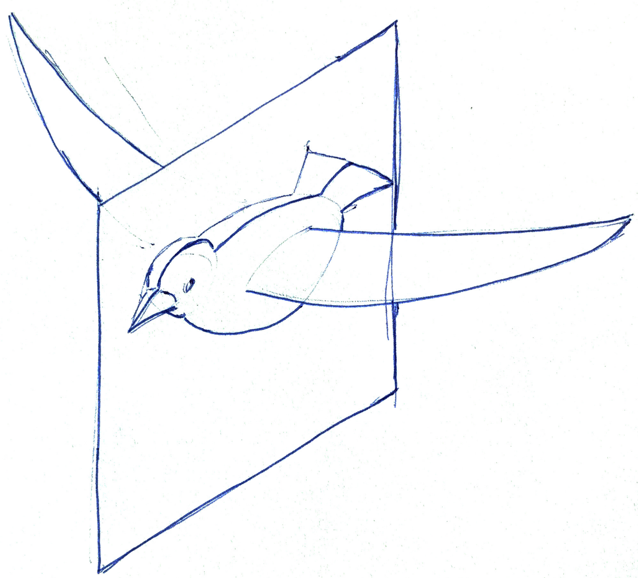
If you begin by defining an ellipse to be the locus of points in the plane for which the sum of the distances to two *focus* points is fixed, then Dandelin spheres show that conic sections can be elliptical. Some basic algebra then shows that these same conic sections satisfy

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \tag{1}$$

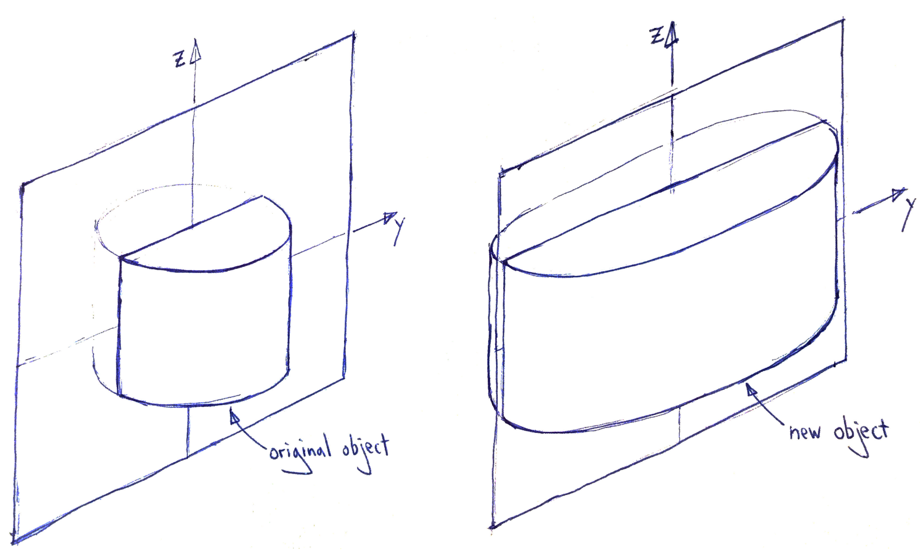
This establishes that ellipses can by described parametrically by $x = a\cos(\theta)$ and $y = b\sin(\theta)$, which in turn makes it clear that scaling an ellipse in the x direction can turn it into a circle.

symmetry

An object is symmetric about a plane if the reflection of any point from the object about the plane is also part of the object.

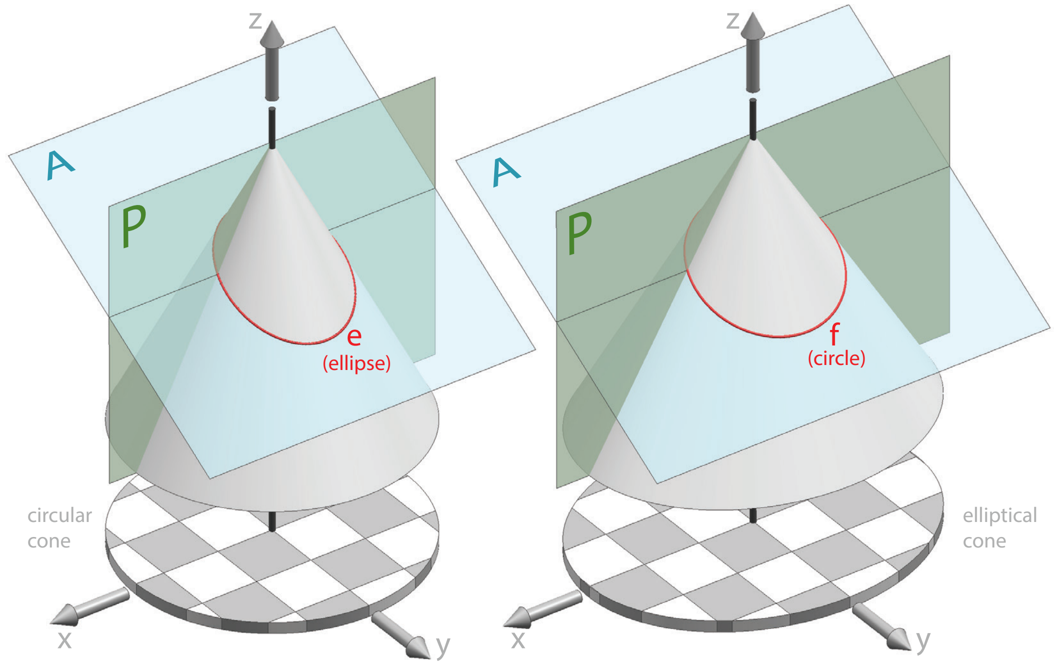


Symmetry is preserved by any distortion which doesn't change the distance of points to the plane. For instance, suppose a collection of points is symmetric about the yz plane, and we create a new object by stretching in the y direction. This new object is still symmetric with respect to the plane.



stretching a cone

Consider a right circular cone cut by a plane A to create an ellipse e. Let P be the plane which contains the cone axis, and which intersects A in a horizontal line. Set up a coordinate system so that the cone axis is the z-axis, and P is the xz plane.



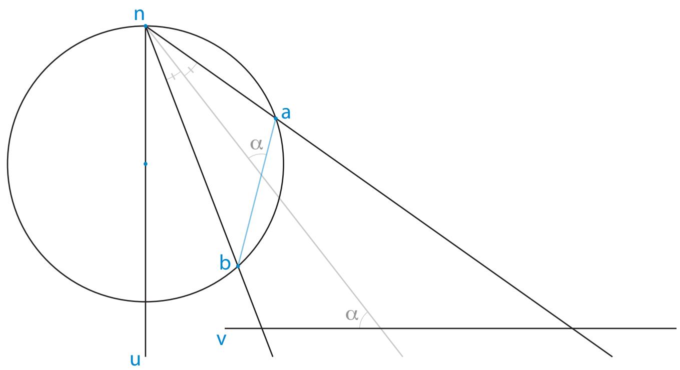
Replacing the x coordinate with αx transforms the ellipse e into a circle f. The circular cone is transformed into an elliptical cone, and this transformed cone *remains symmetric* with respect to P. In particular, reflecting the plane A and the circle f through the plane P gives a different plane A' which intersects the cone at a different circle b'. Thus there are two ways to slice the transformed cone which result in circular sections.

We've established a pleasing duality. Slicing a right circular cone gives an ellipse, while slicing a right elliptical cone gives a circle. The circular cone can be sliced in any direction to create an ellipse, however the elliptical cone can only be sliced in two directions to create a circle.

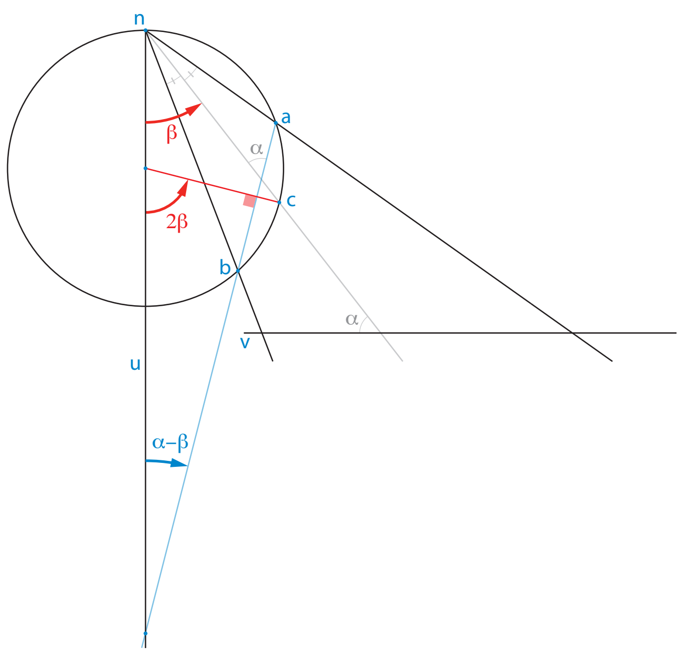
Parallel sections generate the same shape, but at different scales, and so it is clear that the elliptical cone is also a *sheared* right circular cone. So there are at least two ways to get the elliptical cone- by stretching and also by shearing a right circular cone.

back to stereographic projection

Now that we understand elliptical cones, consider the stereographic projection of a circle viewed from the side (i.e., look at the plane defined by the cone axis and the center of the sphere). Let n denote the north pole of the sphere, and let a and b denote the highest and lowest points (respectively) of the circular section.



The lines from n to the circle form a sheared circular cone. From our previous work we know that the section of this cone by the plane perpendicular to the page through v is also a circle. We just need to check that v is perpendicular to u. We do this by drawing a line (in red) from the center of the sphere to the point c, where the sphere intersects the cone axis.



Let β denote the angle of the cone axis to the vertical. Notice that the angle at the bottom of the image is $\alpha - \beta$. Next, note that the angle of the red line to the vertical is 2β, and that the red line is the perpendicular bisector of ab. It follows that $\alpha + \beta = \pi/2$, and from this we know that u and v are perpendicular.

Higher Dimensions
Stereographic projection works in higher dimensions as well, as we now show by working in $\mathbb{R}^n$. Let $\mathbf{E}$ be a unit vector in $\mathbb{R}^n$, let the *ground plane* G be the $n-1$ dimensional subspace of $\mathbb{R}^n$ perpendicular to $\mathbf{E}$, and let S be the $n-1$ dimensional sphere of radius $\frac{1}{2}$ centered at $\frac{1}{2}\mathbf{E}$.

$$S = \{\mathbf{x} \in \mathbb{R}^n \text{ such that } \mathbf{x}\cdot\mathbf{x} = \mathbf{x}\cdot\mathbf{E}\} \tag{2}$$

Stereographic projection is a mapping f from $S - \mathbf{E}$ to G, given by

$$f(\mathbf{x}) = \mathbf{E} + \frac{\mathbf{x} - \mathbf{E}}{1 - \mathbf{x}\cdot\mathbf{x}}. \tag{3}$$

Theorem 1
The stereographic projection of a north pole circle in S is a line in G.

Proof
A north pole circle is the intersection of the sphere S with a plane P that passes through the *north pole* $\mathbf{E}$. $\mathbf{x} \in S$ means that $\mathbf{x}\cdot\mathbf{x} = \mathbf{x}\cdot\mathbf{E}$, while $\mathbf{x} \in P$ means that $\mathbf{x}\cdot\mathbf{e} = \mathbf{E}\cdot\mathbf{e}$, where $\mathbf{e}$ is the unit normal to P. From these two conditions it is straightforward to show that

$$(\mathbf{x} - \mathbf{b})\cdot(\mathbf{x} - \mathbf{E}) = 0, \tag{4}$$

where $\mathbf{b} = \mathbf{e}(\mathbf{e}\cdot\mathbf{E})$ is the point on the circle closest to the ground plane G. That is, it follows from $\mathbf{x} \in S$ and $\mathbf{x} \in P$ that $\mathbf{x}$ is on the circle with antipodal points $\mathbf{E}$ and $\mathbf{b}$. Thus, even in $\mathbb{R}^n$, the intersection of a plane and a sphere is a circle! To show that $f(\mathbf{x})$ is on a line in the ground plane G, we need

$$(f(\mathbf{x}) - \alpha\mathbf{a})\cdot\mathbf{a} = 0. \tag{5}$$

for some scalar α and some vector $\mathbf{a}$ in G. Motivated by the case in 3D, we choose

$$\alpha = \frac{\mathbf{E}\cdot\mathbf{e}}{1 - (\mathbf{E}\cdot\mathbf{e})^2} \quad \text{and} \quad \mathbf{a} = \mathbf{e} - \mathbf{E}(\mathbf{E}\cdot\mathbf{e}) \tag{6}$$

It's then simple arithmetic to check that (5) is satisfied. □

Theorem 2
The stereographic projection of a non north pole circle in S is a circle in G.

Proof
A *non* north pole circle (henceforth simply a *circle*) is the intersection of the sphere S with a plane P that does not pass through the point at $\mathbf{E}$. As before, it is straightforward to show that points $\mathbf{x}$ in this intersection also satisfy

$$(\mathbf{x} - \mathbf{a})\cdot(\mathbf{x} - \mathbf{b}) = 0, \tag{7}$$

where $\mathbf{a}$ and $\mathbf{b}$ are two special *antipodal* points on the circle. If $\mathbf{E}$ and $\mathbf{e}$ are not aligned, $\mathbf{a}$ is the furthest point from the ground plane G, and $\mathbf{b}$ is the closest point to G. Otherwise, $\mathbf{a}$ and $\mathbf{b}$ are any two antipodal points on the circle. Because P is perpendicular to $\text{span}(\mathbf{a}, \mathbf{b})$, any $\mathbf{x}$ on the circle can be written as

$$\mathbf{x} = \alpha\mathbf{a} + \beta\mathbf{b} + \mathbf{u}, \tag{8}$$

where $\mathbf{u}$ is perpendicular to $\mathbf{a}$ and $\mathbf{b}$. Projecting $\mathbf{x}$ onto $\text{span}(\mathbf{a}, \mathbf{b})$ results in a point on the line segment from $\mathbf{a}$ to $\mathbf{b}$, given by $\mathbf{a} + \gamma(\mathbf{b} - \mathbf{a})$ for some $\gamma \in [0, 1]$. From this it is clear that

$$\alpha + \beta = 1. \tag{9}$$

With these preliminaries established, we turn to our main result. The expression

$$F = (f(\mathbf{x}) - f(\mathbf{a}))\cdot(f(\mathbf{x}) - f(\mathbf{b})) \tag{10}$$

being zero is equivalent to $f(\mathbf{x})$ being on a circle in G with antipodal points $f(\mathbf{a})$ and $f(\mathbf{b})$.

Expanding F, we find that

$$F = \frac{AB - (x - a + A)(x - b + B)}{(1 - a)(1 - b)(1 - x)}, \tag{11}$$

where $a = \mathbf{a}\cdot\mathbf{E}$, $b = \mathbf{b}\cdot\mathbf{E}$, $x = \mathbf{x}\cdot\mathbf{E}$, $A = \mathbf{x}\cdot(\mathbf{a} - \mathbf{x})$, and $B = \mathbf{x}\cdot(\mathbf{b} - \mathbf{x})$. We find that the numerator is zero if and only if H is zero, where

$$H = (\mathbf{x}\cdot\mathbf{x})(\mathbf{a}\cdot\mathbf{b}) + (\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b}) - (\mathbf{x}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b}) - (\mathbf{x}\cdot\mathbf{b})(\mathbf{a}\cdot\mathbf{a}). \tag{12}$$

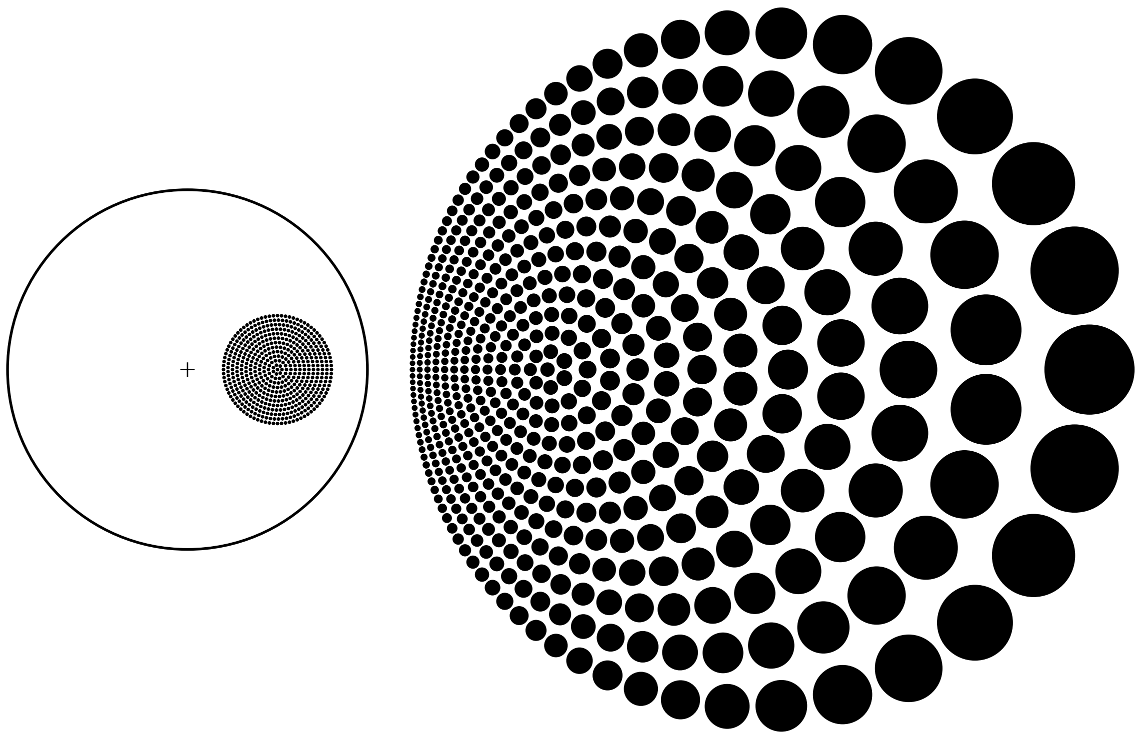
Inserting (8) for $\mathbf{x}$, and using the fact that $\mathbf{E}$ is in the span of $\mathbf{a}$ and $\mathbf{b}$, we find that

$$H = (1 - \alpha - \beta)(\mathbf{a}\cdot\mathbf{a})(\mathbf{b}\cdot\mathbf{b}), \tag{13}$$

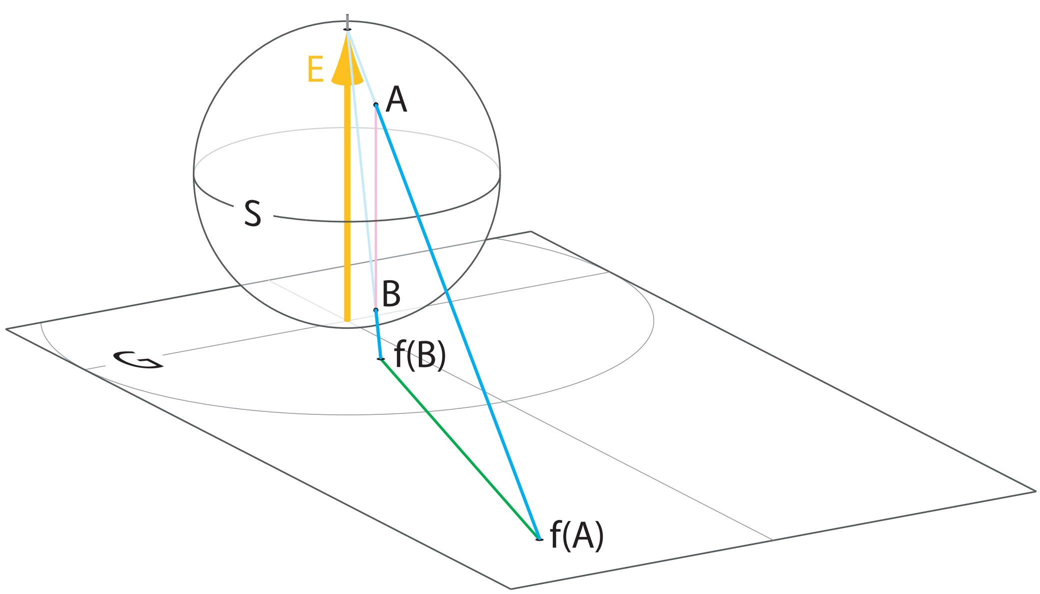
which we know is zero from (9). $\square$

Inversion

Inversion maps any non-zero $\mathbf{x} \in \mathbb{R}^n$ to $\mathbf{e}/\|\mathbf{x}\|$, where $\mathbf{e}$ is the unit vector pointing from the origin to $\mathbf{x}$. Everything inside the unit circle is mapped to outside the unit circle, and vice versa. Although things get expanded and contracted, circles are mapped to circles.



The preservation of circles reflects a link to stereographic projection wherein points A and B on the sphere are reflections of each other about the equatorial plane iff $f(A)$ and $f(B)$ are inversions of each other!



To prove this, write any point $\mathbf{A} \neq \mathbf{0}$ on the sphere as $\mathbf{a} + \alpha\mathbf{E}$, where $\mathbf{a} \cdot \mathbf{E} = 0$. The reflection of $\mathbf{A}$ about the equatorial plane is $\mathbf{B} = \mathbf{a} + (1-\alpha)\mathbf{E}$, and the reflection of $\mathbf{B}$ is $\mathbf{A}$. From $\mathbf{A} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{E}$, we have $\mathbf{a} \cdot \mathbf{a} = \alpha(1-\alpha)$, which makes it easy to see that $\mathbf{f}(\mathbf{B}) = \mathbf{a}/\alpha$ and $\mathbf{f}(\mathbf{A}) = \mathbf{a}/(1-\alpha)$ are inversions of each other.