Area of an Ellipse—C.E. Mungan, Summer 2013

Here I review a noncalculus derivation that the area of an ellipse of semi-major axis a and semi-minor axis b is πab. I use two ingredients: the standard Cartesian form for the equation of an ellipse and the area of a unit circle. Motivations for these two ingredients are in the preliminary section and can be skipped if your students are already comfortable with them.

Two preliminaries

Choosing the origin of two rectangular coordinates to lie at the center of an ellipse, with the x and y axes coinciding with its semi-major and semi-minor axes, the equation of an ellipse can be written as

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

One way to motivate the correctness of this formula is to solve it for y to get

$$y = \pm b\sqrt{1 - \left(\frac{x}{a}\right)^2}. \tag{2}$$

Arbitrarily choose $a = 1$ (which really defines one unit of length), so that we must have $0 < b \leq 1$ and $-1 \leq x \leq 1$. Open Excel and create a column of x values varying from –1 to 1 in increments of 0.01 with the header labeled "x." Calculate y in the second column as

$$y = b\sqrt{1 - x^2} \tag{3}$$

using an absolute reference to column 1 for the value of x. Then compute $-y$ in the third column. Label the header of both columns 2 and 3 with a starting value of b of "1" and use an absolute reference to that header to get the value of b. Copy and paste columns 2 and 3 to columns 4 and 5, and change the header of the two new columns to "0.7." Repeat for b header values of "0.3" and "0.01." The resulting four pairs of curves are plotted below.
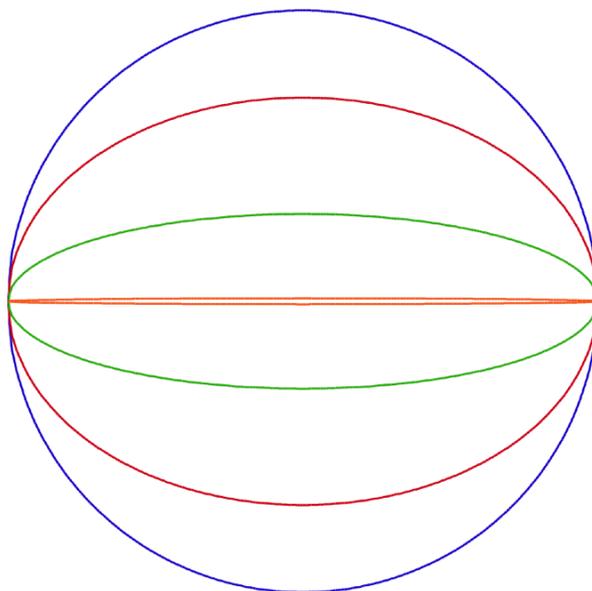
Another way to make such plots is to let $\cos\theta = x/a$ and $\sin\theta = y/b$, which are consistent with each other because Eq. (1) then becomes the Pythagorean relation $\cos^2\theta + \sin^2\theta = 1$. (The physical meaning of θ is that it is the azimuthal angle around the rescaled unit circle. It is NOT the azimuthal angle ϕ around the ellipse. They are related by $b\tan\theta = a\tan\phi$.) In Excel, let θ vary from 0 to 2π in column 1, and compute $x = \cos\theta$ in column 2 and $y = b\sin\theta$ in subsequent columns. An advantage of this angular method of plotting is that one does not need to separately compute y and $-y$.

We see that $a = b = 1$ plotted in blue corresponds to a unit circle, as expected from Eq. (1), which is an ellipse that is not eccentric, i.e., has eccentricity $e = 0$ (whose value we can remember because a circle has the shape of a zero written roundly by hand). On the other hand, the most eccentric ellipse is $e \to 1$ when $b \to 0$, which is approximated by the orange plot above (and again we can remember this value of e because a line segment has the shape of the number one in a sans serif font).

Second, we need the area of a circle. The area of a closed figure is given by tiling it with small rectangles of length Δx and height Δy, and summing over their areas $\Delta x \Delta y$. However, it is difficult to match the circular boundary with rectangles. It makes better sense to switch to plane polar coordinates. We can chop a circle of radius R into pizza slices of radial length R and small angular width θ. Orient a slice vertically, pointed vertex upward, centered above the circular arc at its base. We see that this slice is a triangle of height R and base length $R\theta$. The base becomes straight in the limit of sufficiently small θ. Thus its area is $(R)(R\theta)/2$. Adding up the area of all slices around the circle of total angle $\theta = 2\pi$, we thus get

$$A_{\text{unit circle}} = (1)(1 \cdot 2\pi)/2 = \pi \tag{4}$$

for $R = 1$.

An alternate proof of this result is to note that a unit circle lies between inscribed and circumscribed n-sided polygons, so that

$$A_{\text{inscribed}} \le A_{\text{unit circle}} \le A_{\text{circumscribed}} \tag{5}$$

which becomes

$$n\tan\theta_n \cos^2\theta_n \le A_{\text{unit circle}} \le n\tan\theta_n \tag{6}$$

where $\theta_n = \pi/n$. (These formulae for the inscribed and circumscribed areas can be derived by drawing perpendiculars from the center of the circle to the middle of each side of the polygons, and then summing the areas of the $2n$ right triangles.) In the limit as $n \to \infty$, one has $\tan\theta_n \to \theta_n$ and $\cos\theta_n \to 1$, so that Eq. (6) becomes $\pi \le A_{\text{unit circle}} \le \pi$ as desired.

Derivation of the area of an ellipse

Consider an ellipse. Its area is given by tiling it with small rectangles and summing over their areas,

$$A_{\text{ellipse}} = \sum \Delta x \Delta y. \tag{7}$$

Now rescale the two coordinates by defining two new variables $X = x/a$ and $Y = y/b$. Solving these two relations for x and y, and substituting them into Eq. (7) gives

$$A_{\text{ellipse}} = ab \sum \Delta X \Delta Y. \tag{8}$$

Meanwhile, if we substitute the two new variables into Eq. (1) we get

$$X^2 + Y^2 = 1.$$ (9)

But that is just the formula for a unit circle, whose area is

$$\sum \Delta X \Delta Y = \pi$$ (10)

according to Eq. (4). Substituting Eq. (10) into (8) gives the desired result $A_{\text{ellipse}} = \pi ab$. As a check, note that this result reduces to the formula for the area of a circle when $a = b = R$.