The Area of Ellipses

We are still becoming acquainted with our ellipse. The area of an ellipse is a significant quantity. Later we will appreciate the significance when we study the correlation between the area swept by a planet and time.

In this chapter, in several steps, the area of an ellipse will be shown to be equal to πab. First, the customary derivation of the equation of the ellipse will be presented in the manner of many venerable mathematics textbooks. Subsequently, the equation for the curve of the ellipse will be manipulated and logically compared to the known formula for the area of a circle. The result of the comparison will be the desired formula for the area of an ellipse. The approach is not of my original design; I find it to be advantageous since it avoids a struggle with formal and difficult calculus.

The equation of the ellipse

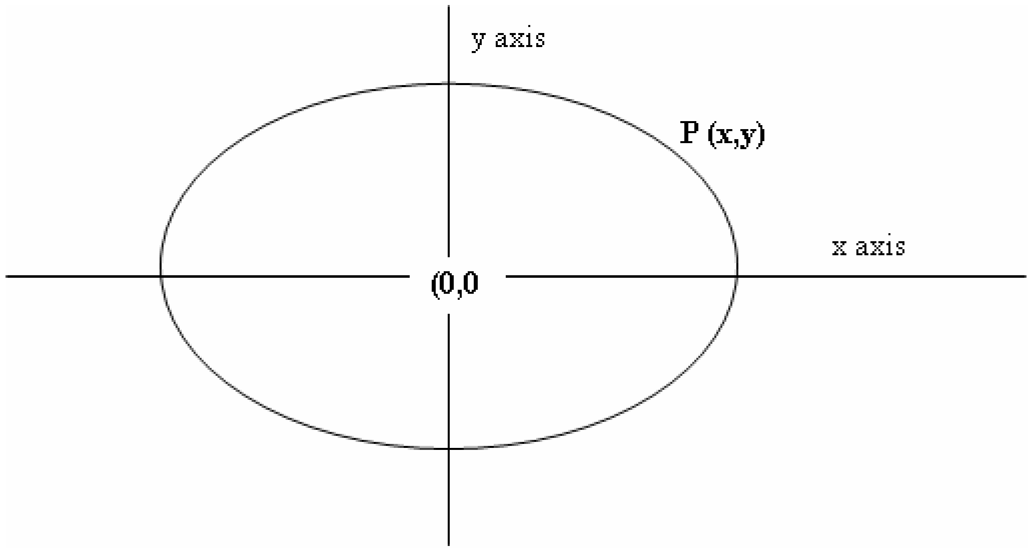
The following demonstration will prove the standard equation of the ellipse which states that:

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

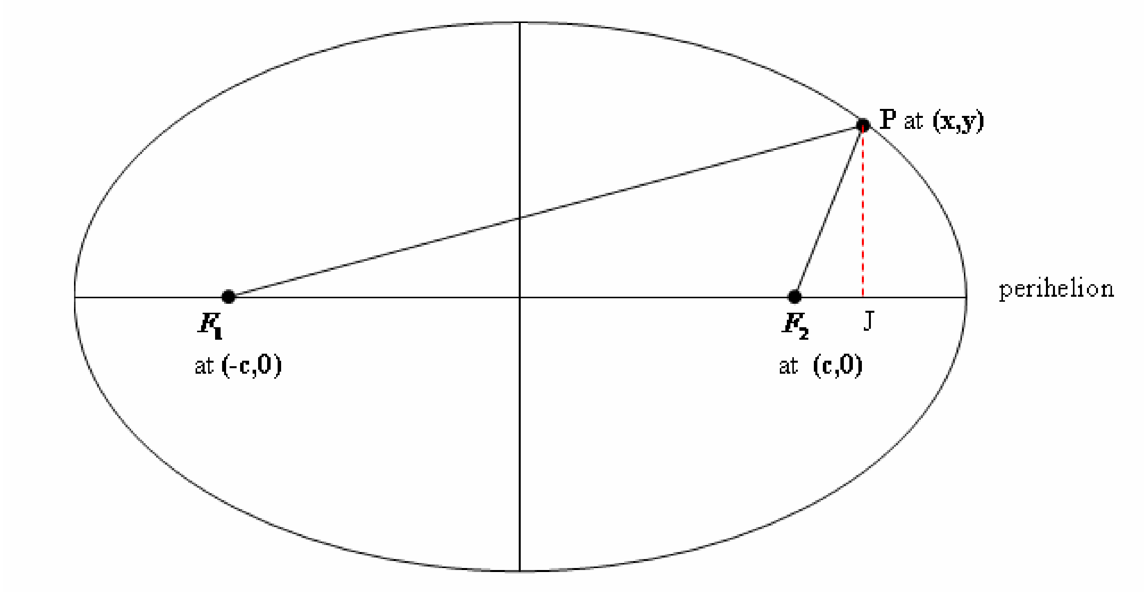
where *x* and *y* are the coordinates of a point on the ellipse, *a* is the semimajor axis, and *b* is the semiminor axis.

Start from the standard definition of the ellipse which states that: An ellipse is a locus of points whose distances from two fixed points is constant. This is precisely the same procedure as taking a string of fixed length and bringing it around two tacks in taut fashion.

Below is an ellipse that has been drawn on a standard x and y axis. The center of the ellipse is at the position of the graph where both x and y are equal to zero. The planet is drawn at a position, **P**, on the ellipse for which there is a value of x and a value of y. (Do not confuse this point with what is later defined to be point *P* - the position of a planet at the top of the semilatus rectum).



In the next figure the principal lines of the ellipse are drawn but the background image of the graph is omitted in order to improve the visual clarity of the diagram. However, keep in mind that the ellipse is centered at the zero value for x and y. The two foci are labeled and are by definition at a distance c from the center of the ellipse. Therefore the x value of one focus is "positive c" and the x value of the other focus is "negative c".



By the standard definition of an ellipse, the distance PF_1 plus the distance PF_2 must sum to a constant length as the point *P* moves around the ellipse.

By algebra and logic, the two legs of the ellipse sum to *2a*. In other words the sum of the length of the two legs is equal to twice the length of the semimajor axis. . We can see this from the fact that the long leg is equal to *a+c* whereas the short leg is equal to *a-c.* Now since the string never changes in total length, the sum of the legs everywhere must also be equal to *2a*.

So:

$$PF_1 + PF_2 = 2a$$

Note in the figure above that there is a red dashed line that is dropped from point P perpendicularly to hit the x axis at point J. By inspection, the red dashed line has a height equal to the value of y at point P. The red dashed line is one side of a right triangle. The base of the triangle is the segment $\overline{F_1J}$. By inspection the length of segment $\overline{F_1J}$ is equal to the distance between the focus F_1 and the center of the ellipse which we know to be defined as the distance c, plus the distance from the center of the ellipse to point J which is equal to the value of x at point P. So the length of the segment $\overline{F_1J}$ is equal to $x+c$. Now the hypotenuse of our right triangle is the segment $\overline{PF_1}$.

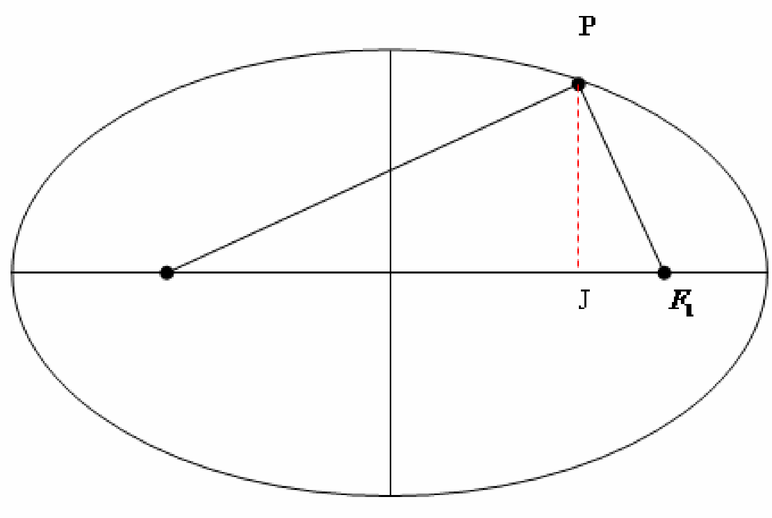
By the rule of Pythagorus for right triangles:

$$PF_1 = \sqrt{(x+c)^2 + y^2}$$

Now by inspection of the figure notice a smaller right triangle whose height is the red dashed line and whose hypotenuse is the segment $\overline{PF_2}$. The base of this small right triangle is the segment $\overline{F_2J}$ whose length is equal to the x value of position P, minus the distance c. Thus again by the rule of Pythagorus:

$$PF_2 = \sqrt{(x-c)^2 + y^2}$$

As an aside notice that if the point P moves so that it in the position above the segment between the center of the ellipse and the left focus:



The short base of the right triangle, segment $\overline{JF_1}$, is actually now equal to $(c-x)$ by inspection, but by the rule of Pythagorus and squaring the side we get the result

$(c-x)^2$ which by the rules of algebra is equivalent to $(x-c)^2$ so that the formula , $PF_2 = (x-c)^2 + y^2$, still holds true.

Knowing that $PF_1 + PF_2 = 2a$ and substituting our expressions for PF_1 and PF_2:

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2a$$

Rearrange to:

$$\sqrt{(x+c)^2 + y^2} = 2a - \sqrt{(x-c)^2 + y}$$

Squaring both sides (and first multiplying the left side fully):

$$(x+c)^2 + y^2 = x^2 + 2cx + c^2 + y^2 = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} + x^2 - cx + c^2 + y^2$$

Simplify by bringing terms that are alike together. Some terms cancel out :

$$2cx = 4a^2 - 4a\sqrt{(x-c)^2 + y^2} - 2cx$$

Move “cx” terms:

$$4cx = 4a^2 - 4a\sqrt{(x-c)^2 + y^2}$$

Divide through by four:

$$cx = a^2 - a\sqrt{(x-c)^2 + y^2}$$

Move terms and multiple both sides by minus one:

$$a^2 - cx = a\sqrt{(x-c)^2 + y^2}$$

Divide through by “a”:

$$a - \frac{c}{a}x = \sqrt{(x-c)^2 + y^2}$$

Now square both sides:

$$(a - \frac{c}{a}x)(a - \frac{c}{a}x) = (x-c)(x-c) + y^2$$

Now multiply out on both sides:

$$a^2 - \frac{ac}{a}x - \frac{ac}{a}x + \frac{c^2x^2}{a^2} = x^2 - cx - cx + c^2 + y^2$$

Simplify:

$$a^2 - cx - cx + \frac{c^2x^2}{a^2} = x^2 - cx - cx + c^2 + y^2$$

Simplify again:

$$a^2 + \frac{c^2x^2}{a^2} = x^2 + c^2 + y^2$$

Rearrange terms:

$$a^2 = x^2 + c^2 + y^2 - \frac{c^2x^2}{a^2}$$

Move terms:

$$a^2 - c^2 = x^2 + y^2 - \frac{c^2x^2}{a^2}$$

Multiply first term on right side by unity as defined by "a" squared divided by "a" squared:

$$a^2 - c^2 = \frac{a^2 x^2}{a^2} + y^2 - \frac{c^2 x^2}{a^2}$$

Rearrange:

$$a^2 - c^2 = y^2 + \frac{a^2 x^2}{a^2} - \frac{c^2 x^2}{a^2}$$

Factor on the right side:

$$a^2 - c^2 = y^2 + (a^2 - c^2)\frac{x^2}{a^2}$$

Now divide both sides by $a^2 - c^2$

$$1 = \frac{y^2}{(a^2 - c^2)} + \frac{x^2}{a^2}$$

And now, since $b^2 = a^2 - c^2$:

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

So:

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

And so, above is the standard equation of the ellipse.

The value of y along the ellipse

Now solve the above equation for the term " y " for any point along the ellipse. This will be useful later in this chapter.

$$\frac{y^2}{b^2} = 1 - \frac{x^2}{a^2}$$

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

Now take the square root:

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

Above is the formula for the value of y for any position along the ellipse.

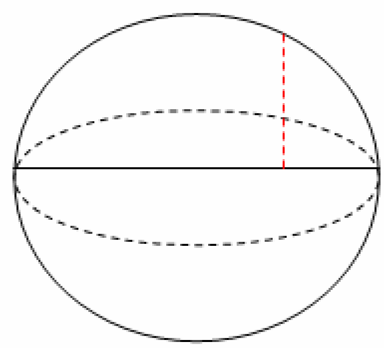
Now for a circle $a = b$ since the semimajor axis and the semiminor axis are both equal to the radius, a, of the circle. So for a circle the expression within parentheses above becomes equal to unity and drops out:

$$y = \sqrt{a^2 - x^2}$$

Above is the formula for the value of " y " for any position along a circle.

<u>The auxiliary circle of an ellipse</u>

The auxiliary circle of the ellipse is a helpful concept. It is the circle that encompasses the ellipse. The diameter of the auxiliary circle is equal to the length of the major axis of the ellipse:



Notice the red dashed vertical line as it intersects the ellipse and the circle. The value of y where it intersects the ellipse can be compared to the value of y where it intersects the circle by using the formulae derived above.

By inspection of the formulae for the y value on the ellipse and for the circle we see that the y value of the ellipse is $\frac{b}{a}$ times the y value of the circle.

Next, this relationship will be used to derive the formula for the area of an ellipse.

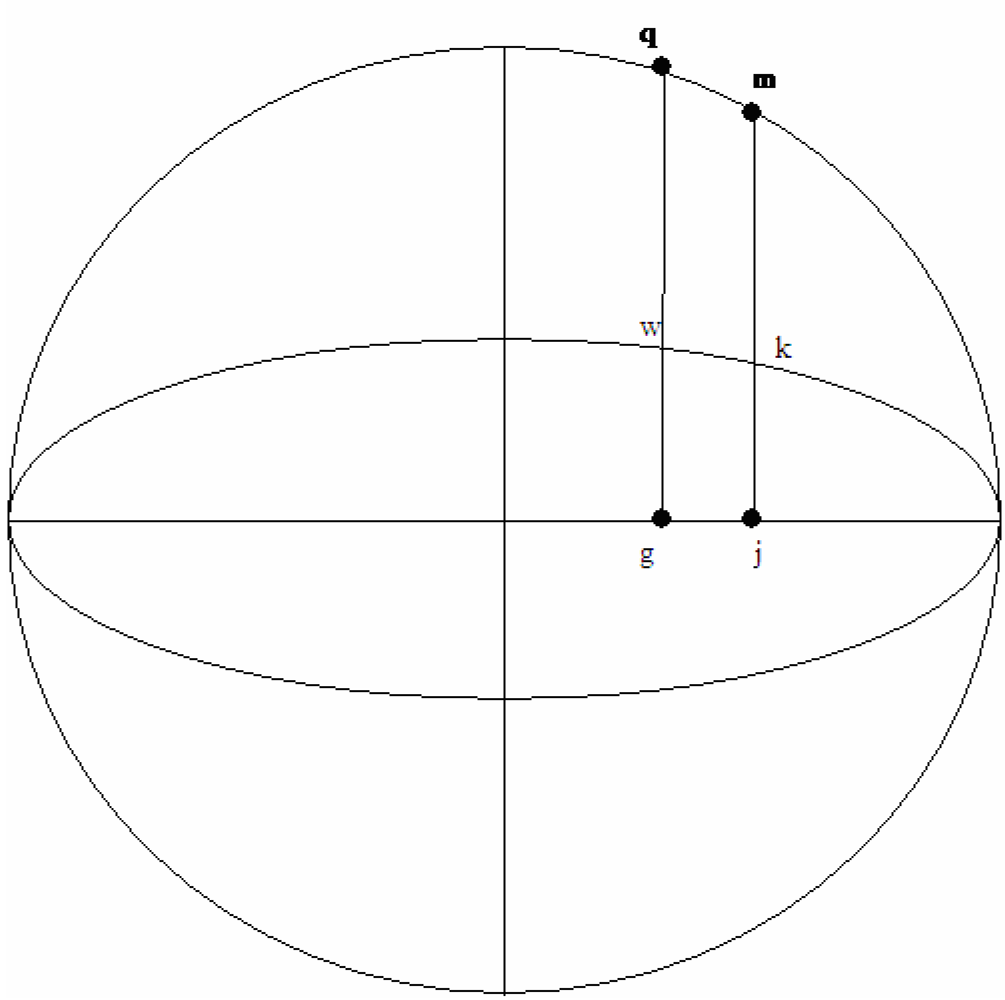
A quadrant of the auxiliary circle

By inspection it is evident that the area of the circle is four times the area of one of its quadrants. The same holds true for the ellipse. a formula will be deduced that compares the area of a quadrant of an ellipse to the area of a quadrant of its auxiliary circle. It is valid by logic that whatever is true for the quadrants will be true for the whole ellipse and circle.

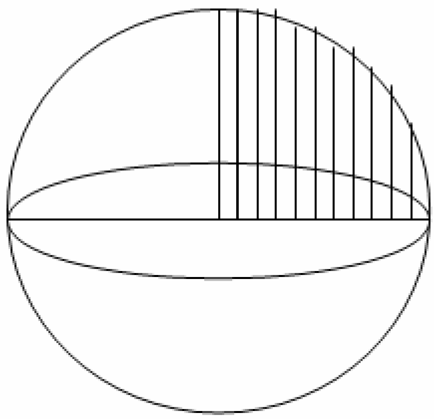
Slices of the ellipse and the circle

Each quadrant may be conceived as being sliced into narrow vertical rectangular pieces. One such representative rectangular slice is shown in the figure. The slice has a boundary marked by the points g,I,w,and k for the ellipse and the boundary points g,I, q, and m for the circle. Notice that for the circle and for the ellipse the rectangular slice in both cases has the same base bounded by the points g and i.

The height of the rectangular slice is greater for the circle than it is for the ellipse. Our formulae for the values of *y* reveal how to quantify the height of the rectangular slice in each case. The height of the rectangular slice for the ellipse is simply the value of y where the slice hits the ellipse and similarly is the value of y where the slice hits the circle. The ratio of the *y* value for an ellipse is $\frac{b}{a}$ times the *y* value of the circle. Therefore the area of the representative slice will be regulated by this ratio. The area of the rectangular slice of the ellipse will be $\frac{b}{a}$ times the area of the rectangular slice of the circle since the two slices share a base and only differ in height by the ratio of $\frac{b}{a}$.



Use the representative slice to reveal the formula for the area of an ellipse. Take as many slices of our quadrant as one can imagine and the relationships regarding area for these many representative slices will always hold true:



As an almost infinite number of slices is taken, the top of each rectangular slice becomes more and more like a straight line since it is less and less curved as the slice gets thinner. So each rectangular slice does become truly more rectangular and thus less mathematically objectionable.

As the ellipse within the auxiliary circle is sliced, it is evident that the circle and ellipse are sliced into the same great number of rectangular slices. And the evaluation of the slices has revealed that the area of each individual slice of ellipse is equal to $\frac{b}{a}$ times the area of the slice of circle. And so, when all these slices are added up we see that the following statement is true:

The area of the quadrant of the ellipse is $\frac{b}{a}$ times the area of the quadrant of the auxiliary circle whose radius is equal to the semimajor axis, a, of the ellipse. The area of a quadrant of a circle of radius a is equal to $\frac{\pi a^2}{4}$. So the area of the ellipse is $\frac{\pi a^2}{4}\left(\frac{b}{a}\right) = \frac{\pi ab}{4}$.

This has all been demonstrated by examining a quadrant of the circle and ellipse. It is evident that the logic of the demonstration holds for the whole of the ellipse and circle. Thus we have our formula for the area of an ellipse.

The area of an ellipse is equal to πab.