Ellipses

February 20, 2018

Ellipses are described algebraically. Four different approaches are related.

Contents

1 Introduction

1.1 Motivation

Kepler orbits are ellipses, so a knowledge of ellipses is important for studying planetary motion. This module introduces the mathematics of ellipses from an algebraic point of view. Calculus or advanced mathematics is not needed.

1.2 Axes and foci

An ellipse is shown in **Figure 1a**. It is centered at the origin O. Let the *major axis* AA' (length $2a$) lie along the x axis and the *minor axis* BB' (length $2b$) lie on the y axis. An ellipse is characterized by two *foci* F and F' along the major axis. The *eccentricity* e (by convention ≥ 0) is defined as

$$e = \frac{|FF'|}{|AA'|}$$

Thus the coordinates of F and F' are $\mp c = \mp ea$. In other words, the points A, F, O, F', A' lie on the x axis with coordinates $-a, -ea, 0, ea, a$.

1.3 Four ways to define an ellipse

There are four ways to define an ellipse.

1. An ellipse is the set of points P such that $PF + PF'$ is a constant. (If the two foci coincide, the ellipse reduces to a circle.)

2. An ellipse is the set of points (x, y) such that $x^2/a^2 + y^2/b^2 = 1$, which we refer to as the canonical formula. (The case $b = a$ reduces to a circle.)

3. An ellipse is the intersection of the surface of a cone with a plane. (The case where the plane is parallel to the base reduces to a circle.)

4. An ellipse is the set of points P with polar coordinates (r, ϕ) measured from one focus (not the center), satisfying $r = a/(1+e\cos\phi)$. (The case $e = 0$ reduces to a circle.)

2 Sum of distances from two foci

2.1 Basic definition

We take the following to be the basic definition (**Figure 1b**): Consider all points P such that

$PF+PF'$ is a constant, say L. Take P to be A. Then

$$L = AF+AF' = AF+A'F = AA' = 2a$$

2.2 Minor axis

Now consider the point B. Consider the right-angle triangle BFO. Obviously the hypotenuse is $BF = BF' = a$. The base is $OF = ea$. Hence the height b is given by $b^2 = a^2 - (ea)^2$,

$$\boxed{b = a\sqrt{1-e^2}} \tag{1}$$

3 Canonical formula

3.1 Derivation of canonical formula

The condition in Section 2 can be written as

$$\sqrt{Q} + \sqrt{Q'} = 2a \tag{2}$$

where

$$\begin{aligned} Q = (PF)^2 &= (x+c)^2 + y^2 \\ Q' = (PF')^2 &= (x-c)^2 + y^2 \end{aligned} \tag{3}$$

To get rid of the square roots in (2) requires squaring twice:

$$\begin{aligned} Q + 2\sqrt{QQ'} + Q' &= 4a^2 \\ 4a^2 - (Q+Q') &= 2\sqrt{QQ'} \\ 16a^4 - 8a^2(Q+Q') + (Q+Q')^2 &= 4QQ' \\ 16a^4 - 8a^2(Q+Q') + (Q-Q')^2 &= 0 \end{aligned}$$

We have

$$\begin{aligned} Q+Q' &= 2(x^2+y^2+c^2) \\ Q-Q' &= 4cx \end{aligned}$$

Note that in $(Q-Q')^2$, the terms quartic in the coordinates cancel, and we are left with a quadratic. Specifically

$$\begin{aligned} 16a^4 - 16a^2(x^2+y^2+c^2) + 16c^2x^2 &= 0 \\ x^2+y^2+c^2 - \frac{c^2}{a^2}x^2 &= a^2 \\ \left(1 - \frac{c^2}{a^2}\right)x^2 + y^2 &= a^2 - c^2 \\ \frac{b^2}{a^2}x^2 + y^2 &= b^2 \end{aligned}$$

which leads to the canonical formula

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1} \tag{4}$$

Although the preceding calculation looks a bit complicated, once you realize that (a) the formula is quadratic, and (b) there are no cross terms xy, then it *must* be reducing to a form such as (4).

3.2 Area of ellipse

Start with a circle

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

and scale the y coordinate by a ratio b/a. Obviously the result is (4). This shows that the area of an ellipse is

$$\boxed{\mathcal{A} = \pi ab} \tag{5}$$

4 Conic section

4.1 General conic section

Historically, especially before the advent of coordinate geometry, ellipses (as well as parabolas and hyperbolas) were described as *conic sections*, i.e., the intersection of a cone with a plane. Start with horizontal circles

$$x^2 + y^2 = \rho^2$$

where the radius is $\rho = R$ for the base at $z = 0$ and decreases linearly to zero at the apex at $z = H$ (**Figure 2a**):

$$\rho = R\,(1 - z/H)$$

Thus the surface of the cone is

$$x^2 + y^2 = R^2\,(1 - z/H)^2$$

This is a quadratic, and importantly will remain a quadratic in all the subsequent manipulations. The surface makes an angle with the z axis given by

$$\beta = \arctan\,(R/H)$$

Instead of intersecting this with an inclined plane, we tilt the cone by an angle α about the x axis, and then intersect it with the plane $z = 0$ (**Figure 2b**). This is achieved algebraically by

$$\begin{aligned} y &\mapsto Cy - Sz = Cy \\ z &\mapsto Cz + Sy = Sy \\ (C, S) &= (\cos\alpha, \sin\alpha) \end{aligned}$$

giving

$$\begin{aligned} x^2 + (Cy)^2 &= R^2(1 - Sy/H)^2 \\ x^2 + Ay^2 + By &= D \qquad (6) \end{aligned}$$

where

$$A = C^2 - S^2\frac{R^2}{H^2} \qquad (7)$$

4.2 Three cases

There are three cases depending on whether A is positive, zero or negative. This in turns depends on whether $\tan\alpha = S/C$ is smaller than, equal to or larger than H/R, or in other words, whether $\alpha + \beta$ is smaller than, equal to or larger than $\pi/2$. The upper edge of the conic surface VV' is below, at or above the horizontal in the three cases (**Figure 3**).

Ellipse
If $A > 0$, it is possible to get rid of the term linear in y by a translation

$$y \mapsto y - y_0$$

The term y^2 is unchanged by this shift, and (6) becomes

$$x^2 + Ay^2 = D'$$

for some D', and this can obviously be brought to the standard form (4) for some choice of a and b.

Hyperbola
If $A < 0$, the term linear in y can again be eliminated, resulting in

$$x^2 - |A|y^2 = D' \qquad (8)$$

Supposing $D' > 0$, this can be case into the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \qquad (9)$$

which is the equation for a pair of hyperbolas that cross the x axis at $x = \pm a$ and asymptotically approach straight lines $y = \pm(b/a)\,x$.

Parabola
In the critical case $A = 0$, (6) becomes

$$y = \frac{D}{B} - \frac{1}{B}x^2 \qquad (10)$$

which is a parabola.

Note on the algebra
It is somewhat messy to write out the parameters of the ellipse (say major and minor axes) in terms of the parameters of the cone; but all we need for our present purpose is that a section of a cone (for $\alpha + \beta < \pi/2$) is equivalent to an ellipse with *some* value of a and b.

5 Measured from one focus

Kepler's first law (here taken as an empirical statement) states that each planet moves in an ellipse with the sun at one focus. So for planetary motion, it is best to use one focus, say F', as the origin, and describe a point P by polar coordinates (r, ϕ) measured from F' (**Figure 4**). The rectangular coordinates referred to the center are then

$$\begin{aligned} x &= r\cos\phi + c = r\cos\phi + ea \\ y &= r\sin\phi \end{aligned}$$

If this is substituted into (4), we get a quadratic equation for r, which can be solved in terms of a, e.

For dimensional reasons, the relationship must take the form

$$A(r/a)^2 + 2B(r/a) + C = 0$$

where A, B, C are dimensionless and therefore can only depend on e, ϕ. To work this out, we note that

$$a^{-2}(r\cos\phi + ea)^2 + b^{-2}(r\sin\phi)^2 = 1$$

Multiplying by $b^2 = (1 - e^2)a^2$ gives

$$\begin{aligned} &(1 - e^2)(r^2\cos^2\phi + 2ear\cos\phi + e^2a^2) \\ &+ r^2(1 - \cos^2\phi) = (1 - e^2)a^2 \end{aligned}$$

From the coefficients of $r^2, 2ra, a^2$, we read off

$$\begin{aligned} A &= (1 - e^2)\cos^2\phi + 1 - \cos^2\phi = 1 - e^2\cos\phi \\ B &= (1 - e^2)e\cos\phi \\ C &= (1 - e^2)e^2 - (1 - e^2) = -(1 - e^2)^2 \end{aligned}$$

It turns out to be more convenient to work with the reciprocal:

$$C(1/r)^2 + 2B(a/r) + A = 0$$

The discriminant is

$$\begin{aligned} D &= B^2 - AC \\ &= (1-e^2)e^2\cos^2\phi + (1-e^2\cos^2\phi)(1-e^2)^2 \\ &= (1-e^2)^2 \end{aligned}$$

Hence

$$\begin{aligned} \frac{a}{r} &= \frac{-B \pm \sqrt{D}}{C} \\ &= \frac{-(1-e^2)e\cos\phi \pm (1-e^2)}{-(1-e^2)^2} \\ &= \frac{e\cos\phi \mp 1}{1-e^2} \end{aligned}$$

Since $0 \le e < 1$ for an ellipse, the minus sign should be discarded, and finally

$$\boxed{r = \frac{r_0}{1 + e\cos\phi}} \tag{11}$$

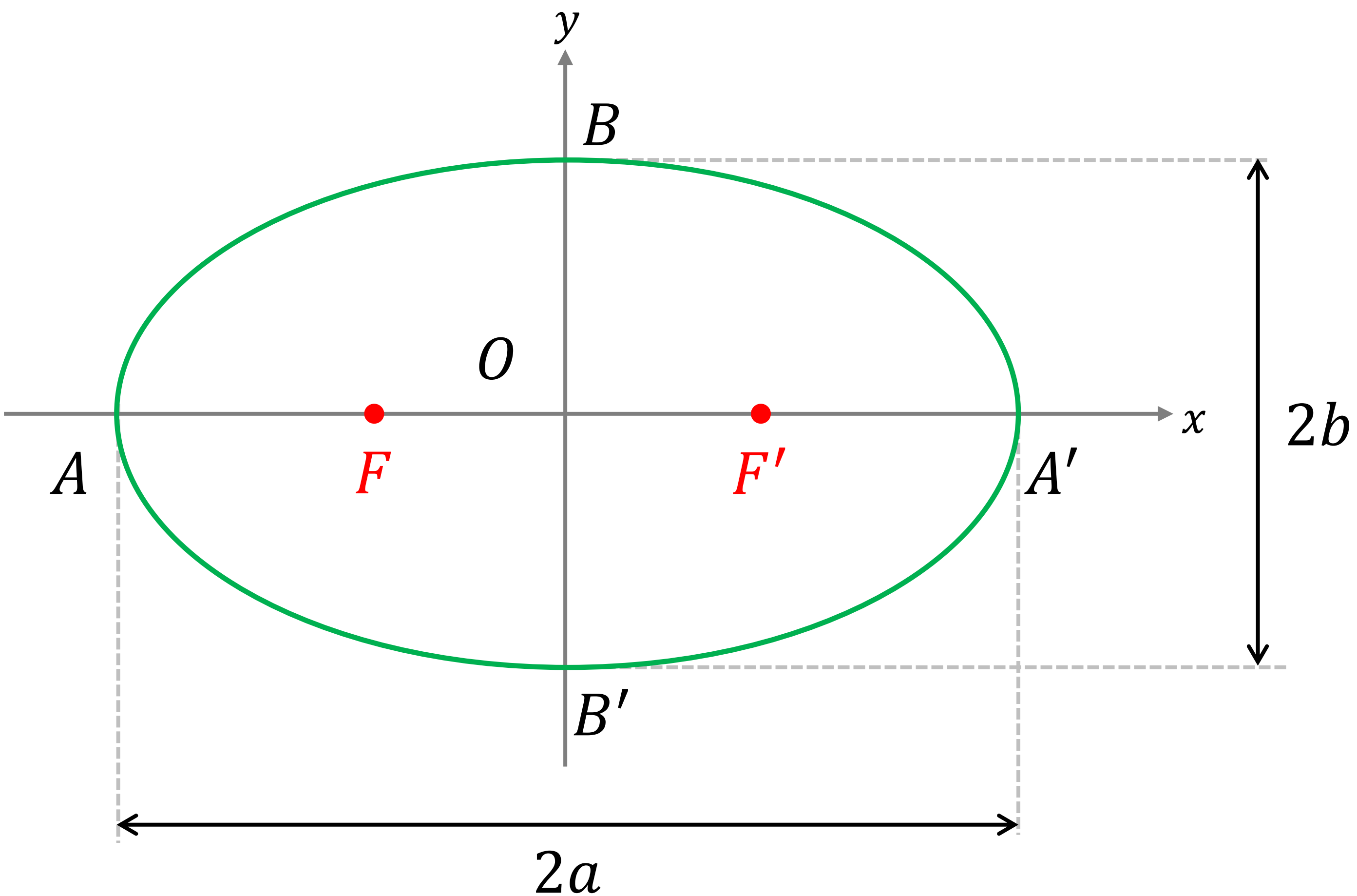
where

$$r_0 = a(1-e^2)$$

Sometimes the equation for an ellipse is written as

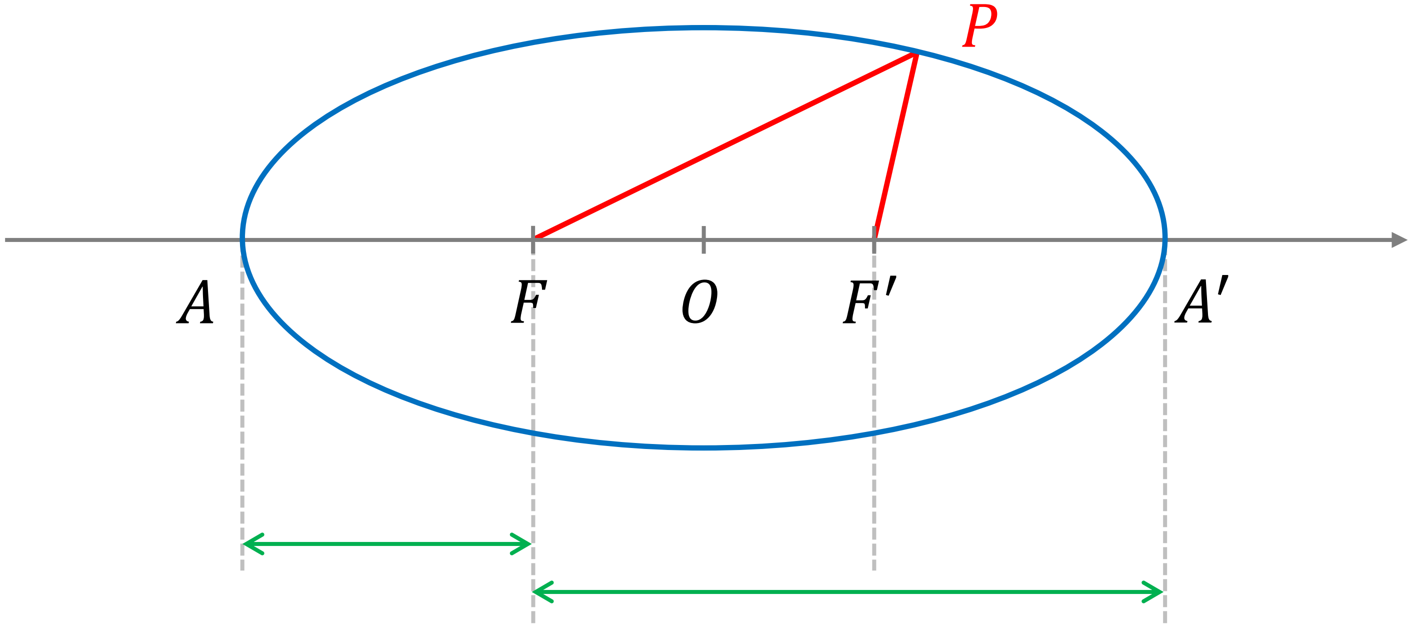
$$r = \frac{r_0}{1 - e\cos\phi}$$

still with $e \ge 0$. This is obtained from (11) by $\phi \mapsto \phi + \pi$, in other words, orienting the x axis in the opposite direction, or equivalently, measuring from the other focus.



$$e = \frac{FF'}{AA'}$$

Figure 1a



$$PF + PF' = L$$

Figure 1b

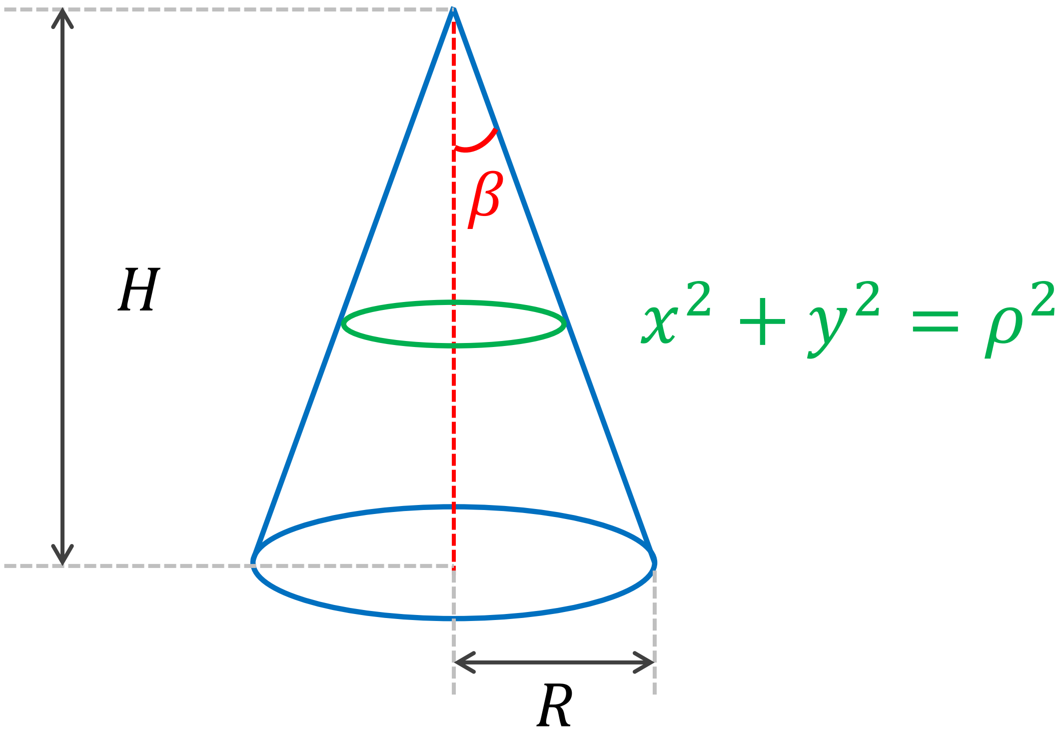


Figure 2a

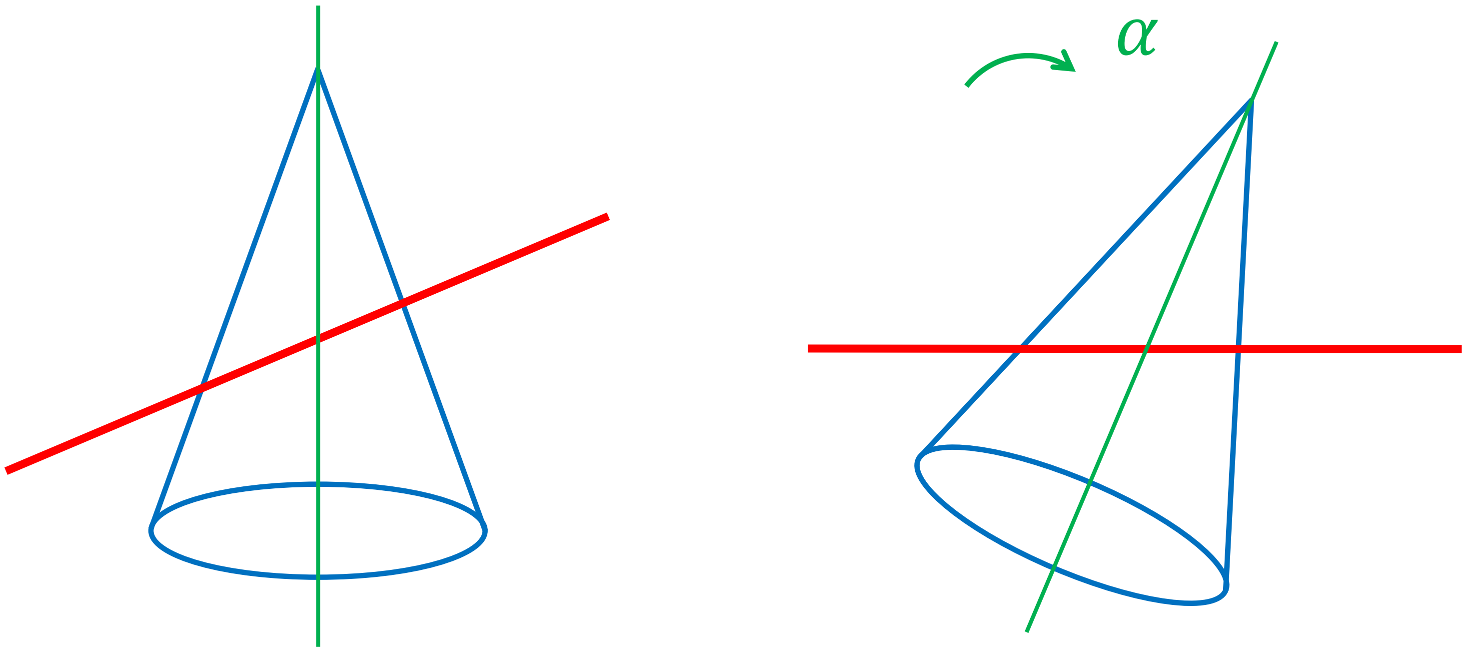


Figure 2b

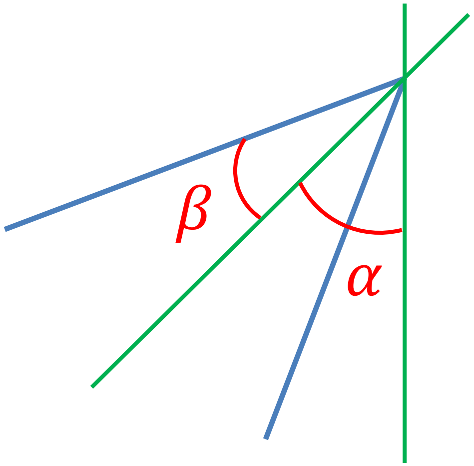


Figure 3

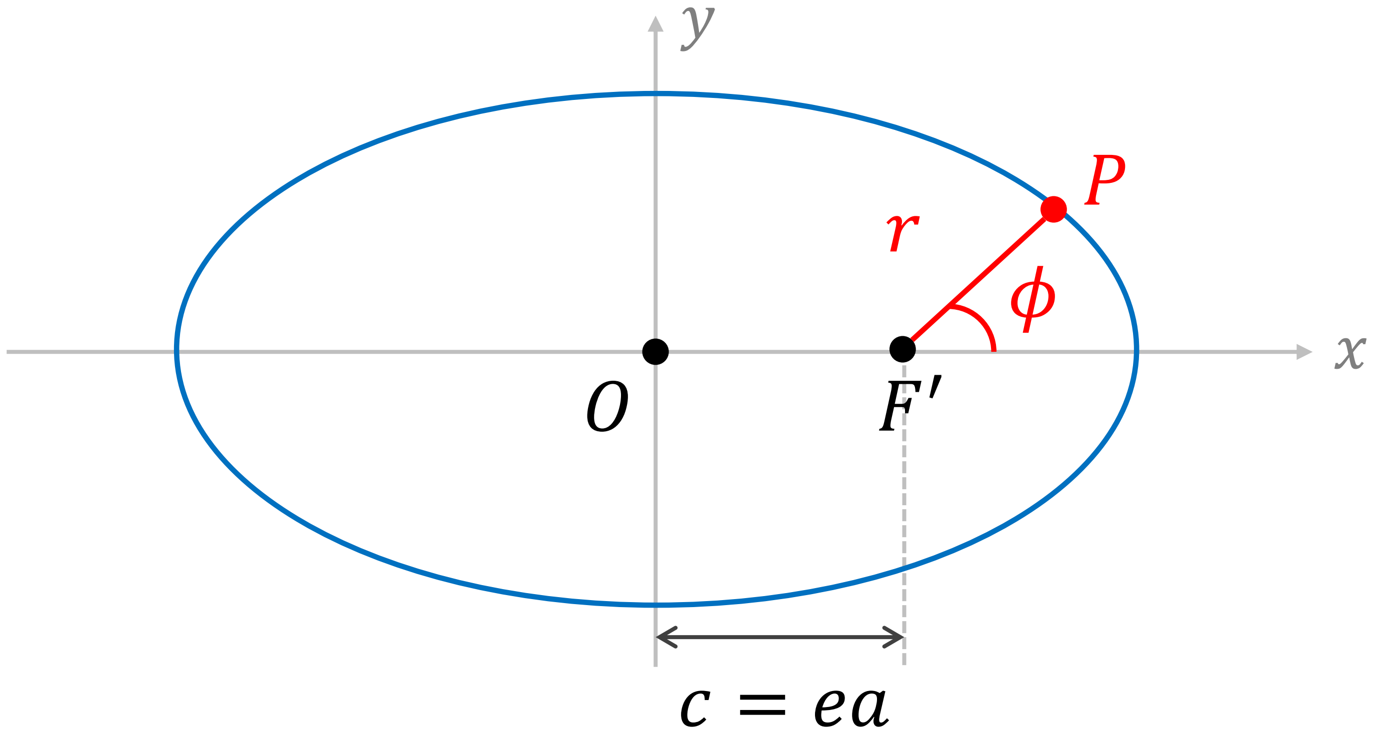


Figure 4