Determine the law of forces when the trajectory is an ellipse: The Newton-Kepler problem

Nestor Caticha

September 12, 2017

Abstract

Newton's derivation of the inverse of the distance square law, the same but otherwise. Aulas para a diversão, ou talvez não, dos alunos de Física 1 de Ciências Moleculares 2017

1 Introduction

This note deals with Proposition XI, Problem VI from Newton's Principia that is (see figure 1):

> If a body revolves in an ellipsis; it is required to find the law of the centripetal force tending to the focus of the ellipsis.

Reading Newton's Principia is a little beyond the reach of a typical first year student. Actually Feynman had his problems too. Thanks to Goodstein and Goodstein we have an account of Feynman's struggle to deal with the proof that an elliptical orbit would result from a central force directed to the focus of the ellipse and decaying with the square of the distance. He blamed our lack of familiarity with Apollonious. Reading Apollonious is still more difficult than Newton. It is a fact that the modern training that students (and their teachers) have is quite different with what was expected at the time of Newton and it is certainly inferior with regard to the methods of Euclid and Apollonious. A biographer of Newton claims that the effort to use geometry instead of the newly developed tools of calculus was guided by his desire to present the results in a geometric language, the *lingua franca* of the contemporary scholars. However Newton's mastery of such methods is beyond other scientists, such as Hooke, whom despite claiming to know that the force was central and decayed with the square of the distance -obtained in the particular case of circular orbits - was unable to determine that in general the trajectories would be conics. Chandrasekhar has written a guide to the common reader. It is really helpful despite having gaps in the explanations that were the source of the trouble that Feynman found. In this note the proof of Newton is simplified so that a first year student can understand every detail. It is not easy but it is within the reach of a dedicated student. In addition we prove the converse, that an inverse square law would give rise to a conic using Feynman's method, which he introduces because Newton "perpetually uses (for me) completely obscure properties of the conic sections." It is a beautiful way to say that, he couldn't follow Newton....but, had he traveled backwards in time and met Newton, he could have shown a geometric way that was even more elegant than Newton's.

Only a few concepts are needed so that a beginner can follow the proofs here. First we are going to use the rules to sum vectors. In addition we need, the first two of Newton's laws of Dynamics. Actually these are according to Newton, due to Galileo, although no one was able to learn them by reading Galileo. Finally Galileo's theorem (as Newton called it), that under a constant acceleration A acting during a time Δt, the change in position with respect to the zero acceleration case is $\Delta x = A(\Delta t)^2/2$. The big IC. Of course some properties of the conics have to be used but these are given below.

2 Some properties of an ellipse

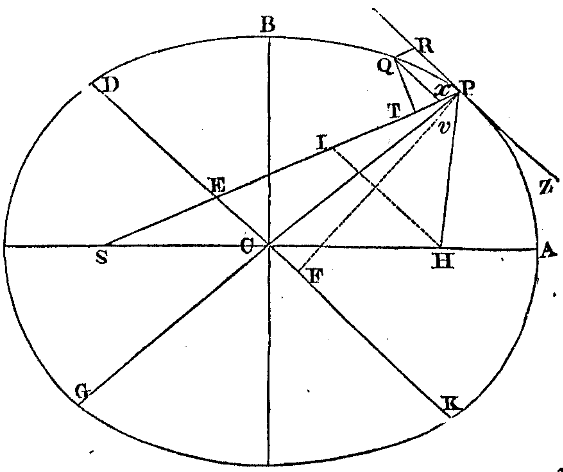
force, now tending to a centre infinitely remote, will become equable. Which is *Galileo's* theorem. And if the parabolic section of the cone (by changing the inclination of the cutting plane to the cone) degenerates into an hyperbola, the body will move in the perimeter of this hyperbola, having its centripetal force changed into a centrifugal force. And in like manner as in the circle, or in the ellipsis, if the forces are directed to the centre of the figure placed in the abscissa, those forces by increasing or diminishing the ordinates in any given ratio, or even by changing the angle of the inclination of the ordinates to the abscissa, are always augmented or diminished in the ratio of the distances from the centre; provided the periodic times remain equal; so also in all figures whatsoever, if the ordinates are augmented or diminished in any given ratio, or their inclination is any way changed, the periodic time remaining the same, the forces directed to any centre placed in the abscissa are in the several ordinates augmented or diminished in the ratio of the distances from the centre.

SECTION III.

Of the motion of bodies in eccentric conic sections.

PROPOSITION XI. PROBLEM VI.

If a body revolves in an ellipsis; it is required to find the law of the centripetal force tending to the focus of the ellipsis.



Let S be the focus of the ellipsis. Draw SP cutting the diameter DK of the ellipsis in E, and the ordinate Qv in x; and complete the parallelogram QxPR. It is evident that EP is equal to the greater semi-axis AC: for drawing HI from the other focus H of the ellipsis parallel to EC, because CS, CH are equal, ES, EI will be also equal; so that EP is the half sum of PS, PI, that is (because of the parallels HI, PR, and the equal angles IPR, HPZ), of PS, PH, which taken together are equal to the whole axis 2AC. Draw QT perpendicular to SP, and putting L for the principal latus rectum of the ellipsis (or for

$\frac{2BC^2}{AC}$), we shall have L $\times$ QR to L $\times$ Pv as QR to Pv, that is, as PE or AC to PC; and L $\times$ Pv to GvP as L to Gv; and GvP to Qv^2 as PC^2 to CD^2; and by (Corol. 2, Lem. VII) the points Q and P coinciding, Qv^2 is to Qx^2 in the ratio of equality; and Qx^2 or Qv^2 is to QT^2 as EP^2 to PF^2, that is, as CA^2 to PF^2, or (by Lem. XII) as CD^2 to CB^2. And compounding all those ratios together, we shall have L $\times$ QR to QT^2 as AC $\times$ L $\times$ PC^2 $\times$ CD^2, or 2CB^2 $\times$ PC^2 $\times$ CD^2 to PC $\times$ Gv $\times$ CD^2 $\times$ CB^2, or as 2PC to Gv. But the points Q and P coinciding, 2PC and Gv are equal. And therefore the quantities L $\times$ QR and QT^2, proportional to these, will be also equal. Let those equals be drawn into $\frac{SP^2}{QR}$, and L $\times$ SP^2 will become equal to $\frac{SP^2 \times QT^2}{QR}$. And therefore (by Corol. 1 and 5, Prop. VI) the centripetal force is reciprocally as L $\times$ SP^2, that is, reciprocally in the duplicate ratio of the distance SP. Q.E.I.

The same otherwise.

Since the force tending to the centre of the ellipsis, by which the body P may revolve in that ellipsis, is (by Corol. 1, Prop. X.) as the distance CP of the body from the centre C of the ellipsis; let CE be drawn parallel to the tangent PR of the ellipsis; and the force by which the same body P may revolve about any other point S of the ellipsis, if CE and PS intersect in E, will be as $\frac{PE^3}{SP^2}$ (by Cor. 3, Prop. VII.); that is, if the point S is the focus of the ellipsis, and therefore PE be given as SP^2 reciprocally. Q.E.I.

With the same brevity with which we reduced the fifth Problem to the parabola, and hyperbola, we might do the like here: but because of the dignity of the Problem and its use in what follows, I shall confirm the other cases by particular demonstrations.

PROPOSITION XII. PROBLEM VII.

Suppose a body to move in an hyperbola; it is required to find the law of the centripetal force tending to the focus of that figure.

Let CA, CB be the semi-axes of the hyperbola; PG, KD other conjugate diameters; PF a perpendicular to the diameter KD; and Qv an ordinate to the diameter GP. Draw SP cutting the diameter DK in E, and the ordinate Qv in x, and complete the parallelogram QRPx. It is evident that EP is equal to the semi-transverse axis AC; for drawing HI, from the other focus H of the hyperbola, parallel to EC, because CS, CH are equal, ES, EI will be also equal; so that EP is the half difference

We use the figure from Newton's Principia shown here. C is the center of the ellipse, A and B are two of its vertices. The semiaxes are $a = CA$ and $b = CB$. The focii are S and H. Define the vectors $\vec{r}_P = \vec{CP}$, $\vec{r}_S = \vec{SP}$ and $\vec{r}_H = \vec{HP}$ for any point P. Call $r_S = |\vec{r}_S|$, $r_H = |\vec{r}_H|$ and $r_P = |\vec{r}_P|$ their modules. What makes the generic point P be at the ellipse is that $r_S + r_H$ is a constant independent of P, which for these particular ellipse is $2a$. Note that of all these distances the important one, that will remain at the end of the analysis is r_S the distance of the point mass to the center of force at S.

The eccentricity ϵ is defined as the ratio of the distance of a focus to the center and the semimajor axis a:

$$\epsilon := \frac{SC}{CA} = \sqrt{1 - \frac{b^2}{a^2}} \tag{1}$$

Using the normal cartesian coordinates, the points on the ellipse satisfy

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

An interesting parametrization that will be useful here is obtained by writing

$$x = a\cos\psi, \tag{3}$$
$$y = b\sin\psi, \tag{4}$$

ψ is not an obvious angle in the figure. It is called the eccentric anomaly. Draw two concentric circles, with center on C of radii a and b. Draw a line which makes an angle ψ with the x axis. It crosses the large circle at an x coordinate and the small circle at a y coordinate that satisfy equations 3 and 4 respectively. As ψ goes from 0 to 2π point P describes the ellipse. The nice property of this parametrization is that if ψ increases by $\pi/2$ another point, D in the figure is determined. Any line that goes through the center is called a diameter of the ellipse and if their eccentric anomalies differ by $\pi/2$ the diameters are called conjugated. Since

$$\frac{d\cos\psi}{d\psi} = -\sin\psi = \cos(\psi + \pi/2), \tag{5}$$
$$\frac{d\sin\psi}{d\psi} = \cos\psi = \sin(\psi + \pi/2), \tag{6}$$

one concludes that the conjugated diameter CD is parallel to the line RZ tangent to the ellipse at P. [1]

Point P is the position of a point mass (a planet?) at a certain time and point Q the position at an interval Δt later. Since Q is also at the ellipse we can write:

$$\vec{r}_P = a\vec{i}\cos\psi + b\vec{j}\sin\psi, \tag{7}$$
$$\vec{r}_Q = a\vec{i}\cos(\psi + \Delta\psi) + b\vec{j}\sin(\psi + \Delta\psi), \tag{8}$$

With the notation of Newton in the figure, points $PRQx$ in the figure form a parallelogram by construction, hence $|Qx| = |PR|$ and $|Px| = |QR| := x_P$. The distance x_P is the change in position that the planet (Q) has with respect to what it would be (R) if the planet moved under the action of no force (first law). Galileo's theorem

$$x_P = \frac{1}{2}\frac{F_c}{m}(\Delta t)^2 \tag{9}$$

holds the promise that, if we are able to obtain the change in position x_P from the geometrical properties of the figure, we might be able to obtain the centripetal force directed from point P to the focus S (where the sun is), provided we are able to deal with the time Δt it took the planet to move to Q. This problem will be dealt geometrically with the result that for a central force equal areas are swept in equal times. For this we need another angle θ, which now is quite natural to define. It is the angle that the vector $\vec{r}_s$ makes with the x-axis. It is not the usual polar angle, defined as the angle that the vector $\vec{r}_P$, from the center of the ellipse to the planet, makes with

[1] Note that Newton (or the translator) didn't bother to mention at this point that the diameters CD and CP are conjugated, although, in previous figures it is mentioned, but some letters do not maintain the same meaning across all figures.

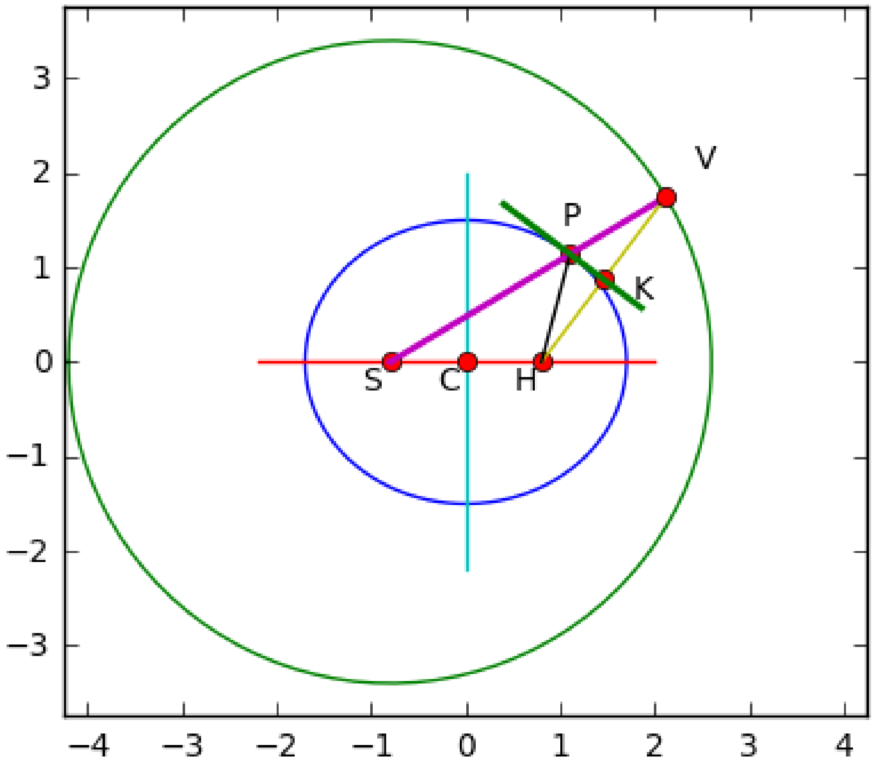


Figure 3: A similar diagram to Newton's. Here we have extended the line SP to point V by an amount equal to the length of HP. Since $|SP| + |HP| = 2a$ is a constant as the point P moves around, the point V describes a circle.

the x-axis. Of course they coincide for the particular case when the ellipse is a circle. The relation between the two angles we have introduced is simple to obtain, in terms of the coordinates x and y and therefore:

$$\begin{aligned} r_s \cos\theta &= |SC| + x = a\epsilon + a\cos\psi, \\ r_s \sin\theta &= y = b\sin\psi. \end{aligned} \tag{10}$$

Summing the squares of the left sides we obtain

$$r_s = a(1 + \epsilon\cos\psi), \tag{11}$$

Dividing the left sides we obtain:

$$\tan\theta = \frac{b}{a}\frac{\sin\psi}{\epsilon + \cos\psi} \tag{12}$$

and small variations of θ are related to small variations of ψ by

$$\frac{\Delta\theta}{\cos^2\theta} = \frac{b}{a}\Delta\psi\frac{1 + \epsilon\cos\psi}{(\epsilon + \cos\psi)^2}. \tag{13}$$

The reader who wants to very this result should use that the derivative of $\frac{u}{v}$ is $\frac{u'v-uv'}{v^2}$

3 The geometry using vectors

From figure 4 one can see that the following relations between the vectors:

$$\vec{r}_Q = \vec{CS} + \vec{SX} + \vec{XQ} \tag{14}$$

$$\vec{r}_Q = -a\epsilon\vec{i} + \xi_1\vec{r}_S + \xi_2\vec{r}_D, \tag{15}$$

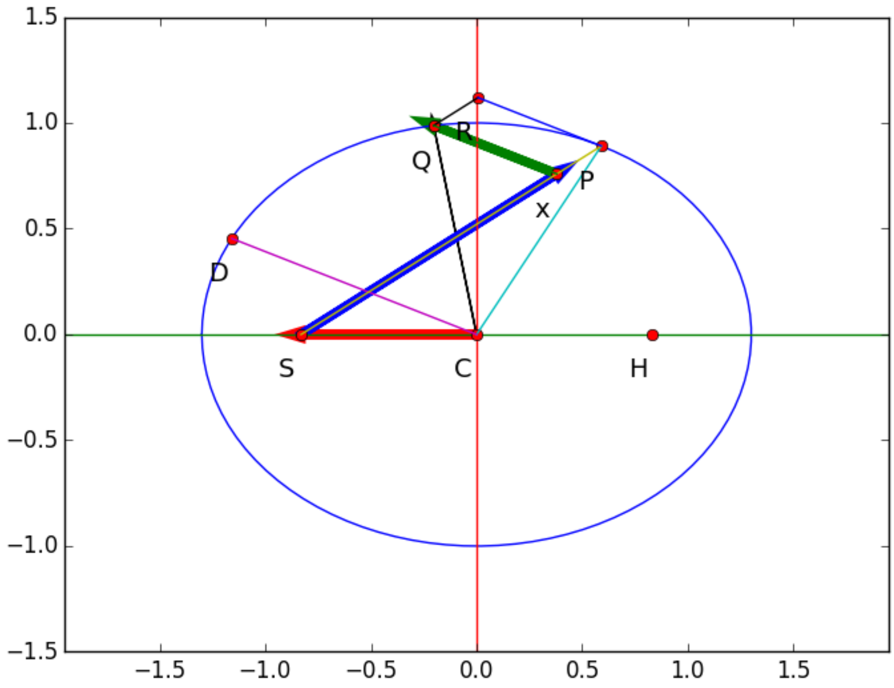


Figure 4: The vectors $\vec{SC} + \vec{Sx} + \vec{xQ}$

are valid, where we introduced the unknown quantities ξ_1 and ξ_2, since the vectors $\vec{SX}$ and $\vec{r}_S$ are parallel and so are $\vec{XQ}$ and $\vec{r}_D$. These yet unknown quantities can't be arbitrarily chosen, since equation 8 has to be satisfied, otherwise point Q is not on the ellipse. From equations 8 and 15 we obtain two equations, for the x and y components respectively, for the two unknowns:

$$\text{x coord:} \qquad \xi_1(\epsilon + \cos\psi) - \xi_2 \sin\psi = \epsilon + \cos(\psi + \Delta\psi) \tag{16}$$

$$\text{y coord:} \qquad \xi_1 \sin\psi + \xi_2 \cos\psi = \sin(\psi + \Delta\psi). \tag{17}$$

Cramer's rule leads to

$$\xi_1 = \frac{\cos\Delta\psi + \epsilon\cos\psi}{1 + \epsilon\cos\psi} \tag{18}$$

$$\xi_2 = \epsilon\frac{\sin(\psi + \Delta\psi) - \sin\psi}{1 + \epsilon\cos\psi}. \tag{19}$$

Note that as $Q \to P$, $\Delta\psi \to 0$ and $\xi_1 \to 1$ and $\xi_2 \to 0$ as it should. Actually we need to be a little more careful in analyzing the behavior of ξ_1 for small $\Delta\psi$:

$$\xi_1 = \frac{1 - \frac{\Delta\psi^2}{2} + \ldots + \epsilon\cos\psi}{1 + \epsilon\cos\psi} = 1 - \frac{\Delta\psi^2/2}{1 + \epsilon\cos\psi} \tag{20}$$

and using equation 11

$$\xi_1 = 1 - \frac{1}{2}\Delta\psi^2 \frac{a}{r_s} \tag{21}$$

The interest in this development derives from the relation between the fallen distance x_P, the distance from the planet to the sun, r_P and ξ_1:

$$x_P = (1 - \xi_1) r_s \tag{22}$$

4 Putting all together

We now have almost all the ingredients:

- (i) $F_c = \frac{2m}{\Delta t^2}(1 - \xi_1) r_s$.

For small values of $\Delta\psi$

- $1 - \xi_1 \approx \frac{1}{2}\Delta\psi^2 \frac{a}{r_s}$

The relation between $\Delta\psi$ and $\Delta\theta$ can be simplified

- From equation 13

$$\Delta\psi = \Delta\theta \frac{a}{b} \frac{1}{\cos^2\theta} (\frac{r_s \cos\theta}{a})^2 \frac{1}{1 + \epsilon\cos\psi}$$

and using equation 11

$$\Delta\psi = \Delta\theta \frac{r_s}{b} \tag{23}$$

Putting it all together with Galileo's result, Newton could obtain the centripetal force

$$\begin{aligned} F_c &= \frac{m}{\Delta t^2} a \Delta\psi^2 = ma(\frac{\Delta\psi}{\Delta t})^2 \\ &= \left(\frac{\Delta\theta}{\Delta t}\right)^2 \frac{mar_s^2}{b^2} \end{aligned} \tag{24}$$

which is not the whole history since we have to investigate how the angular velocity, obtained at the limit $Q \to P$ of $\frac{\Delta\theta}{\Delta t}$.

But this is obtained from a previous result of Newton, that for a central force equal areas are swept in equal times:

$$\frac{\Delta A}{\Delta t} = \frac{1}{2} r_s^2 \frac{\Delta\theta}{\Delta t} = K \tag{25}$$

and K has, for a particular planet a fixed value. Of course, the fact that this was empirically found by Kepler suggested to Hooke and Newton that the gravitational force was central, a result that would have surprised Kepler and Copernicus, who thought that the gravitational force acted transversely, i.e. in the direction of the motion of the planet. Then $\frac{\Delta\theta}{\Delta t} = \frac{2K}{r_s^2}$ and together with equation 24 gives the final result for the magnitude of the centripetal force:

$$\begin{aligned} F_c &= \frac{4maK^2}{b^2} \frac{r_s^2}{r_s^4} \\ F_c &= \frac{C}{r_s^2}. \end{aligned} \tag{26}$$

Hooke could not prove this result for the the ellipse, although he could for the circular orbit. In the controversy between Hooke and Newton, that followed the $1/r^2$ law, Newton's claim to priority was supported by the fact that no other had the geometrical ability to obtain this result for the conics. It also follows for the hyperbole or parabola, but we will not pursue now, and the reader is invited to make the small changes needed for the proof.

5 Feynman's solution for the inverse problem

Given that the gravitational force on the planet due to the sun is attractive and decays inversely with the square of the distance to the sun, we want to prove that the orbit is in general a conic and in particular, for the case considered here, an ellipse.

The method is quite simple once a few ideas are introduced. Extend the line SP in figure 3 by a distance equal to PH to a point V. This is essentially the first time we mention the other focus, which is usually thought to be of no interest, but was shown by Feynman's method to have a quite interesting role. Note that the sum of distances from P to the focii is fixed:

$$2a = |SP| + |PH| = |SP| + |PV| = |SV| \tag{27}$$

hence, as point P moves around the ellipse, point V describes a circle of radius $2a$. Keep this circle in mind. We have to find a circle in the dynamics. This is the central point in Feynman's method, and he mentions this key idea was due to Mr. Fano who was working with Rutherford scattering. Divide the trajectory into N equal intervals $\Delta\theta = \frac{2\pi}{N}$ and eventually we will take $N \to \infty$ The law of areas for a central force again plays an fundamental role, but now we write it stressing that the angular intervals to be considered are the same, and so the time intervals can't be equal

$$\frac{\Delta A_n}{\Delta t_n} = \frac{1}{2} r_s(n)^2 \frac{\Delta\theta}{\Delta t_n} = K. \tag{28}$$

For equal angles the time intervals scale like the r_s^2

$$\Delta t_n = \frac{1}{2K} r_s(n)^2 \Delta\theta. \tag{29}$$

From Newton's second law, for a central force that decays with the square of distance we have, for the n^{th} interval

$$m\frac{\vec{\Delta v}}{\Delta t_n} = -\frac{C_1}{r_s(n)^2}\frac{\vec{SP}}{|SP|}, \tag{30}$$

where we have indicated the direction $-\frac{\vec{SP}}{|SP|}$ towards the sun. Note that the changes in velocities at each point θ_n has a constant magnitude, since

$$|\vec{\Delta v}| = \Delta t_n \frac{C_2}{r_s(n)^2} = C_3 \Delta\theta \tag{31}$$

where the constant C_3 is easy to calculate but unnecessary. Furthermore the direction is radial and hence the change in direction from one angle to the next is just the constant $\Delta\theta$. If we draw the changes $\Delta\vec{v}$, as in the top of figure 5, they all point in the direction of the sun. But we can move the $\Delta\vec{v}$ vectors so that they look as in the bottom of figure 5. If we plot the velocity vectors $\vec{v}_n$ from a common center, then they radiate with different lengths, but the vectors $\Delta\vec{v}_n$ move around in a perfect circle whose center doesn't coincide with the radiating center of the velocities, which change magnitude as shown in figure 6 (top). You may now realize that the new center from where the velocities radiate is related to the other focus of the ellipse... but we have to work a little bit more to see that.

It turns out to be easier, if we measure θ from the x axis, and start describing the motion when the direction is perpendicular to the axis. So the center from where the velocities radiate is at the position in the vertical line which contains the circle of velocities as shown for the N points in figure 6(top). The angle θ as the planet moves is exactly the same as the angle measured from such vertical with respect to the center of the circle. Note that all changes in velocity change orientation by the same $\Delta\theta$. The Feynman trick is to rotate the circle by an angle of $\pi/2$ so that the the direction of a velocity vector rotates to a perpendicular orientation to what is the velocity of the planet, shown in figure 6(bottom). From this we have to reconstruct the position of the planet. There are three conditions that have to be satisfied:

1. The position of the planet makes an angle θ with respect to the $x-$ axis.
2. The velocity of the planet is perpendicular to the line H$\tilde{\text{V}}$, obtained from the velocity $\vec{v}(\theta)$ by a rotation of $\pi/2$.
3. The trajectories are tangent to the velocity of the planet.

We show again this rotation for just one point in the trajectory in figures 7 and 9. Compare this last one to figure 3. You might complain that it is difficult to compare a diagram where the lengths in the picture represent distances (figure 3 and one where the lengths in the picture represent velocities and their variations. We are interested in the relative size of the several parameters in the orbit, and we can draw them so that the outer circles coincides on paper. The three conditions above are satisfied if the planet's position is chosen at the intersection of the line SV (condition 1) and a perpendicular to the line HV (condition 2). But which perpendicular? The midpoint of line HV is the only choice possible in order to satisfy the third.

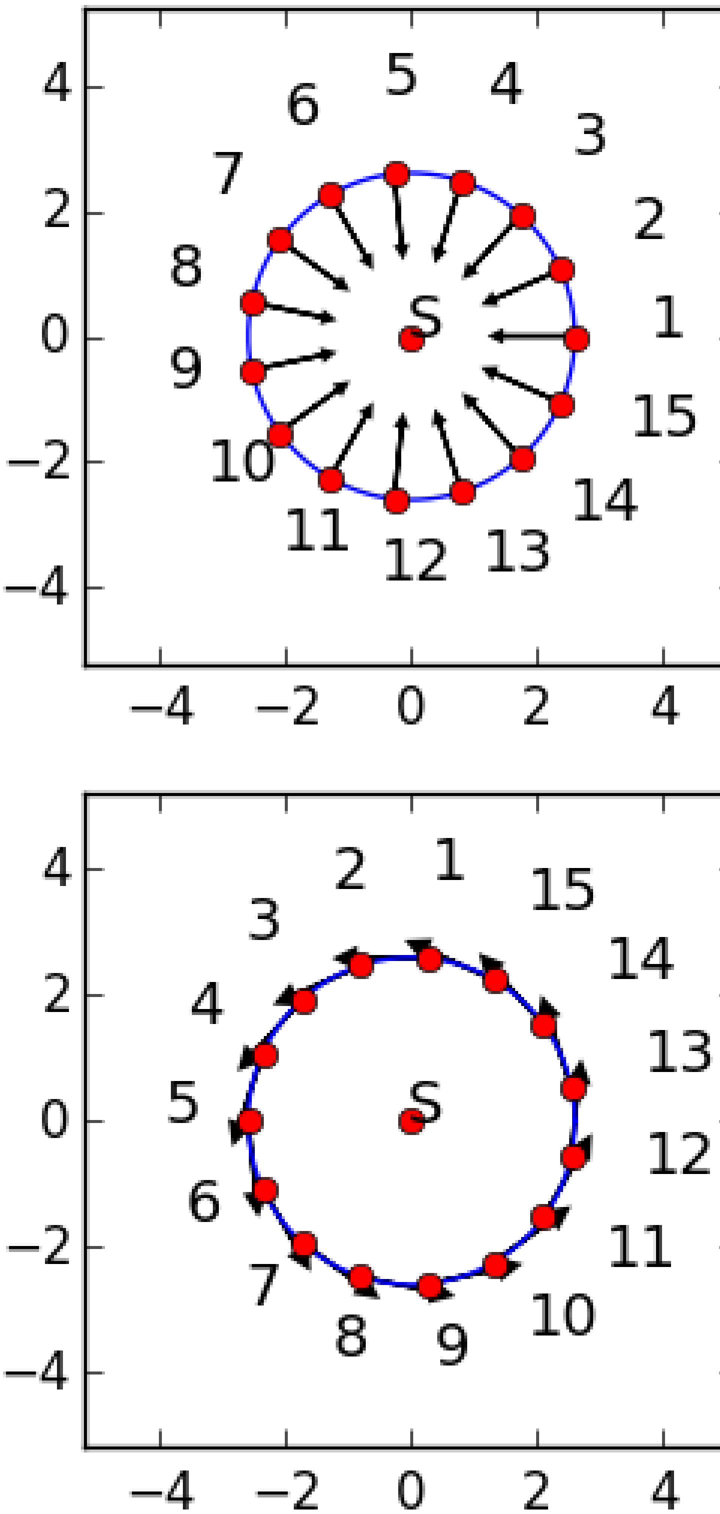


Figure 5: Top: At N equal angular intervals the changes in velocities $\Delta\vec{v}_i$ are equal in magnitude and turn by $\Delta\theta = 2\pi/N$ with respect to the previous vector. Bottom: the same vectors are plotted from a point rotated by $\pi/2 - \Delta\theta/2$. Here $N = 15$ and as $N \to \infty$ the vectors cover the circle.

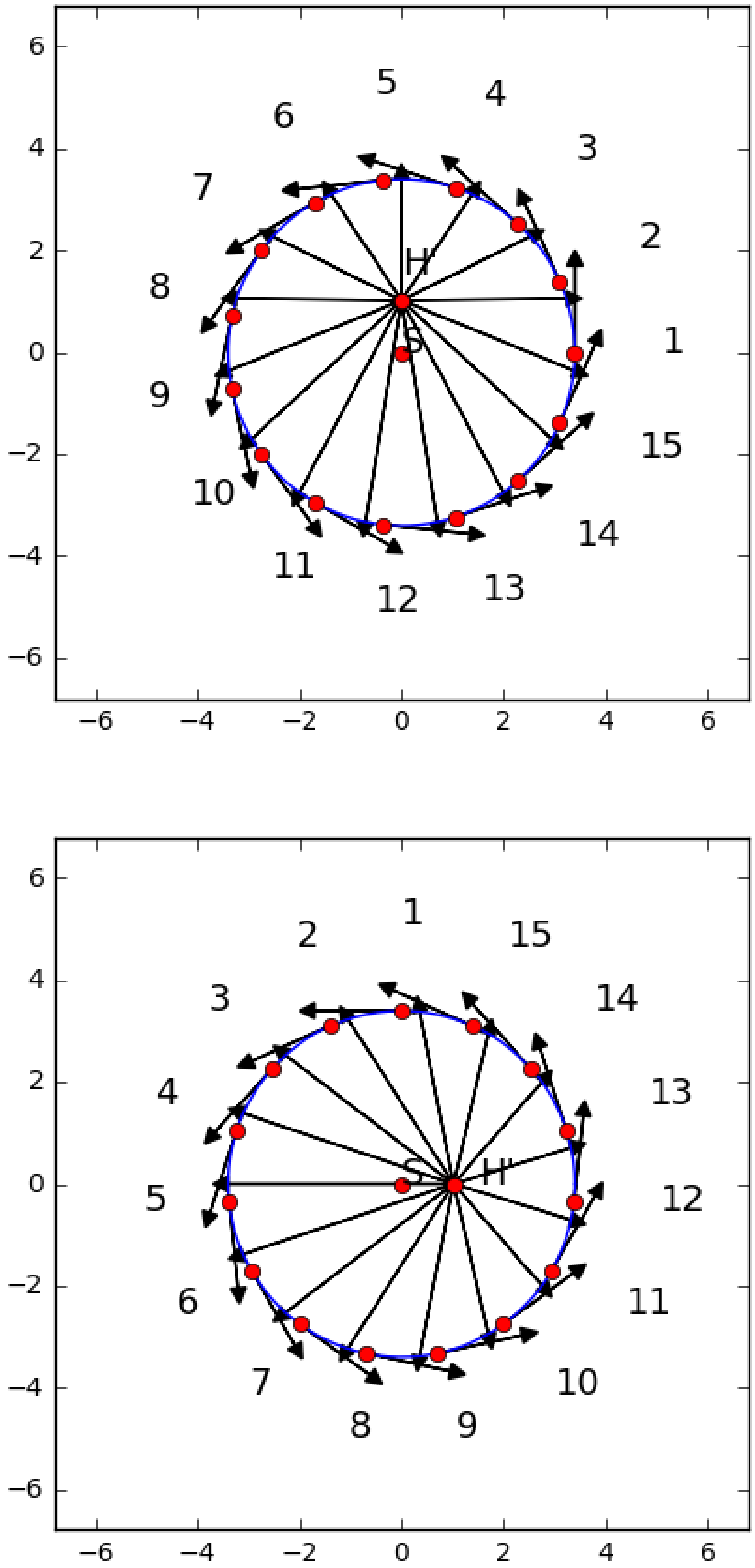


Figure 6: Drawing the velocity vectors from a common center H, such that when the planet is at $\theta = 0$ the velocity is in the y direction. Then the radiating center of the velocities is above S.

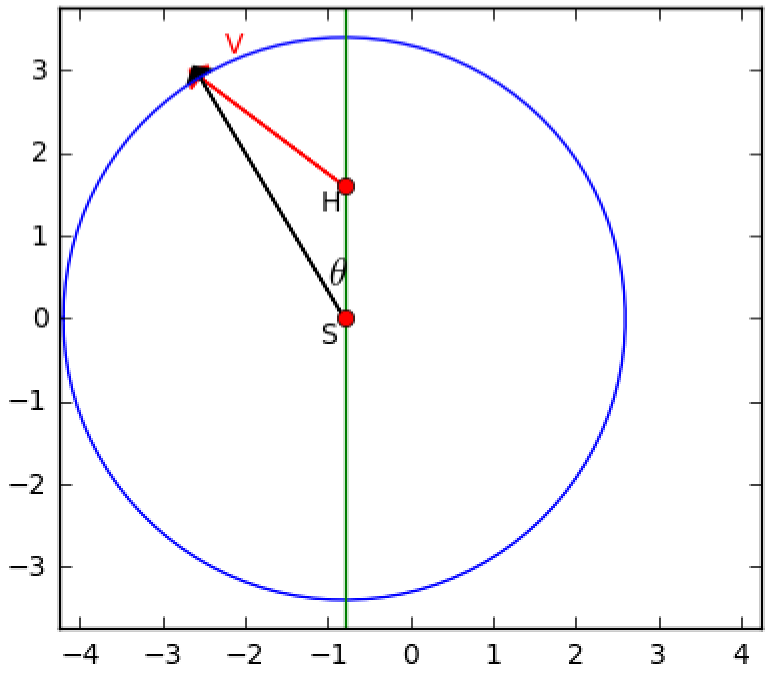


Figure 7: Pick just one point in the trajectory. Measure the angle of the planet as seen from the sun by θ.

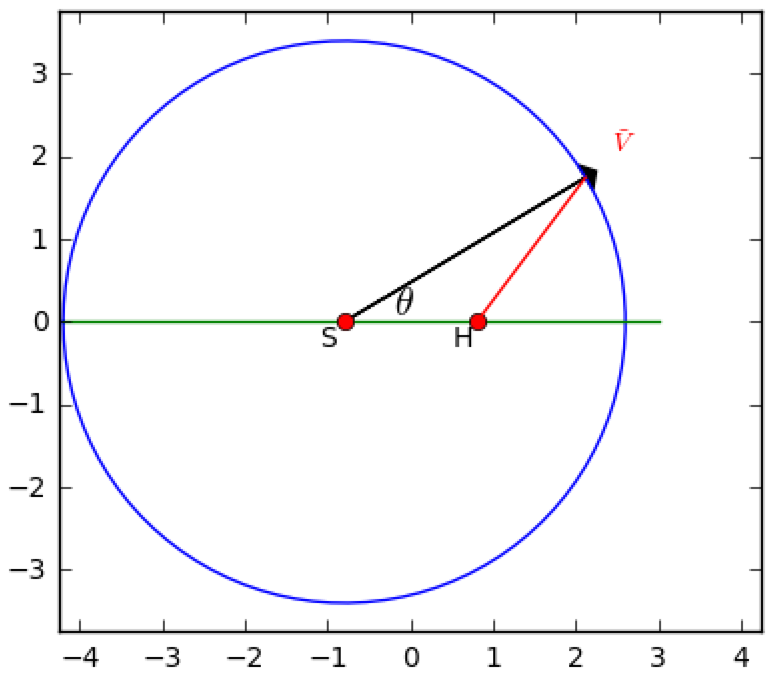


Figure 8: Rotate the previous picture by $\pi/2$. Since we measure the angle of the planet as seen from the sun by θ, now the line $S\tilde{V}$ points in the direction of the planet.

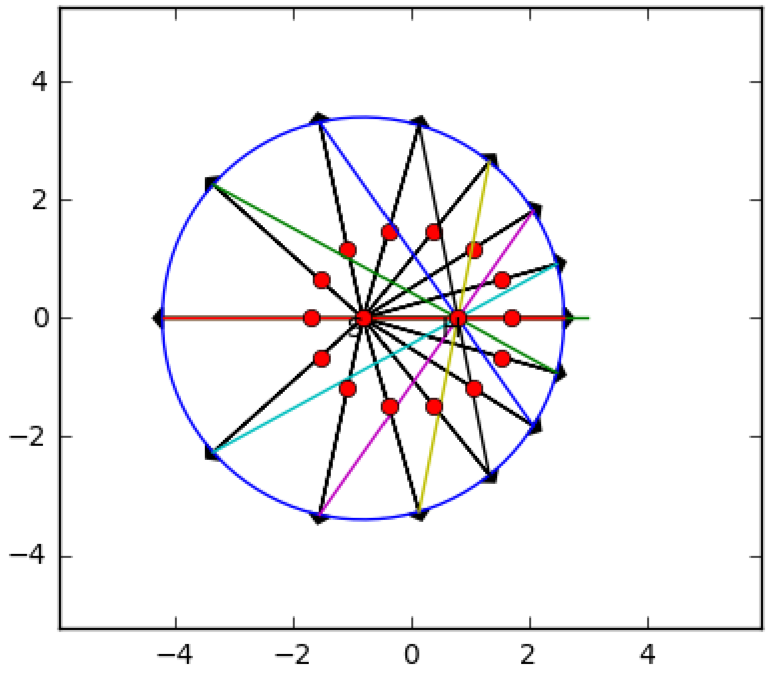


Figure 9: The result of choosing the mid point, for a set of points where the eccentric anomaly varies by a constant angle. Note that this is, first, easier to program and second irrelevant, since the construction should be valid for any θ