Kepler, Newton and Planetary Orbits

Andrew Smith

April 24, 2009

Abstract

In 1609 Johannes Kepler (1571-1630) published the first two of his three laws of planetary motion. After analyzing the data Tycho Brahe (1546-1601) had collected from observing the motion of Mars over a period of more than 20 years, he reached the conclusion that the orbit of a planet about the sun is an ellipse with the sun at one focus (*his first law*). His work was a scientific triumph because it created a model for the solar system that was not only more accurate than any before, but simpler as well. Though brilliant, Kepler's result amounted to fitting a curve to a set of data without discovering any fundamental principles. In 1687 Newton provided the missing principles. For example, Kepler's first law, combined with Newton's law of universal gravitation, led to Newton's law of ellipses: "objects attracted to a center by a force inversely proportional to the square of the distance travel in conic sections". Hence the behavior of the planets could be explained by the same laws which govern the path of an apple as it falls from a tree to the ground; for the first time it became clear that the so-called heavenly bodies behaved no differently than the seemingly more substantial bodies of our everyday experience. The solution of this problem is one of the greatest triumphs of the human intellect in general and of calculus in particular. Here, I present you with a more modern derivation; one that starts out in the *complex plane*, a concept that took a couple hundred years after Newton to fully develop, yet the basic concept is true to Newton and Kepler.

1 The Search for the Equation of Planetary Motion

Kepler's second law of planetary motion states that the line joining the Sun with any planet sweeps out equal areas in equal times.This tells us that as a planet approaches the Sun, the line joining the two becomes shorter, so the planet must travel faster. This law then led Kepler to his third law: "The ratio of the cube of the semi-major axis of an elliptical orbit to the square of the orbital period is the same for all planets". Newton deduced both the inverse square law for gravity, and the direction of that force, from Kepler's second and third laws. Then, using that deduction, he showed how Kepler's first law is not an independent statement, but rather a necessary consequence of these laws. Newton's derivation was completely geometrical and extremely complicated.

2 Newton's Law of Gravitation

Suppose we have two bodies, one of mass m, and a larger one of mass M. Newton's laws of motion hold in a coordinate system with the origin located at the center of mass of the two bodies. If we let M be the Sun and m a planet (or, an asteroid or comet), then the center of mass is *inside of* M. Therefore we choose *polar* coordinates in the *complex plane* so that M is at the origin, and we let $z(t)$ represent the position of m with respect to M at time t. Thus,

$$z = re^{i\theta}$$

Where r and θ are real-valued functions of t.

Using Newton's "Law of Universal Gravitation", the gravitational force between M and m is given by

$$F = \frac{GMm}{r^2}$$

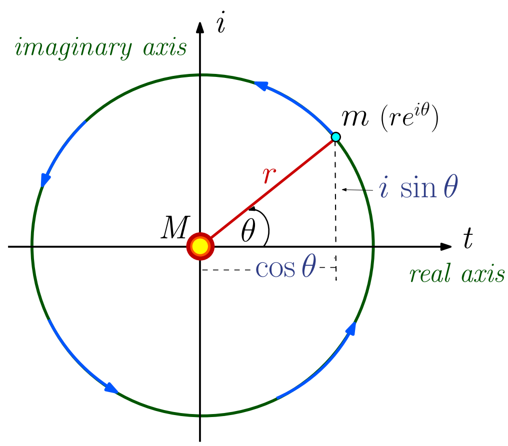


Figure 1: The Complex Plane and the orbit of m about M. The position of m at time t is given by $re^{i\theta}$

G is the *gravitational constant* $6.67 \times 10^{-11} Nm^2/kg^2$, and r is the distance between M and m at time t. (*force is in Newtons, distance in meters, and mass in kilograms*). Since gravity is an attractive force and we are assuming M to be at rest at the origin, F is directed from m toward M, giving us

$$F = -\frac{GMm}{r^2}e^{i\theta}$$

For sake of simplicity (and without loss of generality), we'll assume that this is the only force acting on the two bodies. By Newton's second law of motion, $F = ma$, we have

$$ma = -\frac{GMm}{r^2}\, e^{i\theta}$$

Substituting $k = GM$ and dividing through by m , this becomes

$$a = -\frac{k}{r^2}\, e^{i\theta} \tag{1}$$

2.1 Velocity and Acceleration at time t

Velocity is the first derivative of the position function $z(t)$. Thus,

$$v(t) = \frac{dz}{dt} = \frac{d}{dt}\, re^{i\theta} = ire^{i\theta}\frac{d\theta}{dt} + e^{i\theta}\frac{dr}{dt}$$

Acceleration is the first derivative of velocity. Thus,

$$\begin{aligned}
a(t) &= \frac{dv}{dt} \\
&= \frac{d}{dt}\left(ire^{i\theta}\frac{d\theta}{dt} + e^{i\theta}\frac{dr}{dt}\right) \\
&= ire^{i\theta}\,\frac{d}{dt}\left(\frac{d\theta}{dt}\right) + \frac{d\theta}{dt}\frac{d}{dt}\left(ire^{i\theta}\right) + e^{i\theta}\,\frac{d}{dt}\left(\frac{dr}{dt}\right) + \frac{dr}{dt}\frac{d}{dt}\left(e^{i\theta}\right) \\
&= ire^{i\theta}\,\frac{d^2\theta}{dt^2} + \frac{d\theta}{dt}\left(i^2re^{i\theta}\,\frac{d\theta}{dt} + ie^{i\theta}\,\frac{dr}{dt}\right) + e^{i\theta}\,\frac{d^2r}{dt^2} + \frac{dr}{dt}\, ie^{i\theta}\,\frac{d\theta}{dt} \\
&= ire^{i\theta}\,\frac{d^2\theta}{dt^2} - re^{i\theta}\left(\frac{d\theta}{dt}\right)^2 + ie^{i\theta}\,\frac{d\theta}{dt}\frac{dr}{dt} + e^{i\theta}\frac{d^2r}{dt^2} + ie^{i\theta}\,\frac{d\theta}{dt}\,\frac{dr}{dt} \\
&= -re^{i\theta}\left(\frac{d\theta}{dt}\right)^2 + e^{i\theta}\frac{d^2r}{dt^2} + i\left(re^{i\theta}\,\frac{d^2\theta}{dt^2} + 2e^{i\theta}\,\frac{d\theta}{dt}\,\frac{dr}{dt}\right) \qquad (2)
\end{aligned}$$

2.2 Deriving the Second Order O.D.E for Determining The Orbit

Putting 1 and 2 together gives

$$-\frac{k}{r^2}\, e^{i\theta} = -re^{i\theta}\left(\frac{d\theta}{dt}\right)^2 + e^{i\theta}\frac{d^2r}{dt^2} + i\left(re^{i\theta}\,\frac{d^2\theta}{dt^2} + 2e^{i\theta}\,\frac{d\theta}{dt}\,\frac{dr}{dt}\right),$$

After dividing through by $e^{i\theta}$ we have

$$-\frac{k}{r^2} = -r\left(\frac{d\theta}{dt}\right)^2 + \frac{d^2r}{dt^2} + i\left(r\frac{d^2\theta}{dt^2} + 2\,\frac{d\theta}{dt}\,\frac{dr}{dt}\right) \; . \tag{3}$$

Equating the real and imaginary parts of 3, we have

$$-\frac{k}{r^2} = -r\left(\frac{d\theta}{dt}\right)^2 + \frac{d^2r}{dt^2} \; . \tag{4}$$

And

$$0 \;=\; r\frac{d^2\theta}{dt^2} + 2\,\frac{d\theta}{dt}\,\frac{dr}{dt} \; . \tag{5}$$

Multiplying both sides of 5 by r gives us

$$0 \;=\; r^2\frac{d^2\theta}{dt^2} + 2r\,\frac{d\theta}{dt}\,\frac{dr}{dt} \; . \tag{6}$$

However,

$$r^2\frac{d^2\theta}{dt^2} + 2r\,\frac{d\theta}{dt}\,\frac{dr}{dt} \;=\; \frac{d}{dt}\left(r^2\,\frac{d\theta}{dt}\right) \; ,$$

so 6 implies that

$$\frac{d}{dt}\left(r^2\,\frac{d\theta}{dt}\right) = 0 \; .$$

Since a function with 0 for its derivative must be a constant function, it follows that

$$r^2\,\frac{d\theta}{dt} = c \; , \quad \textit{or} \qquad \frac{d\theta}{dt} \;=\; \frac{c}{r^2} \tag{7}$$

We will now use the substitution $s = 1/r$ to get 4 into a simpler form. First,

$$\frac{dr}{dt} \;=\; \frac{d}{dt}\left(\frac{1}{s}\right) \;=\; -\frac{1}{s^2}\,\frac{ds}{dt} \;=\; -\frac{1}{s^2}\,\frac{ds}{d\theta}\,\frac{d\theta}{dt} \; . \tag{8}$$

Putting s into 7,

$$\frac{d\theta}{dt} \;=\; \frac{c}{r^2} \;=\; cs^2 \; . \tag{9}$$

Substituting 9 into 8,

$$\frac{dr}{dt} \;=\; -\frac{1}{s^2}\,\frac{ds}{d\theta}(cs^2) \;=\; -c\,\frac{ds}{d\theta} \; .$$

Differentiating again,

$$\frac{d^2r}{dt^2} \;=\; -\,c\,\frac{d}{dt}\left(\frac{ds}{d\theta}\right) = \;-\,c\,\frac{d}{d\theta}\left(\frac{ds}{d\theta}\right)\,\frac{d\theta}{dt} \;=\; -c\,\frac{d\theta}{dt}\,\frac{d^2s}{d\theta^2}\;. \tag{10}$$

Hence, using 9 on the previous page in 10

$$\frac{d^2r}{dt^2} = -c\frac{d\theta}{dt}\frac{d^2s}{d\theta^2} = -cs^2\,\frac{d^2s}{d\theta^2}\;. \tag{11}$$

Finally, substituting 11,and s into 4 on the preceding page gives us

$$-ks^2 \;=\; -\frac{1}{s}\,\left(cs^2\right)^2 - c^2s^2\,\frac{d^2s}{d\theta^2} \;\;=\;\; -c^2s^3 - c^2s^2\,\frac{d^2s}{d\theta^2}\;, \tag{12}$$

Dividing both sides of 12 by $-c^2s^2$, we have

$$\frac{d^2s}{d\theta^2} + s = \frac{k}{c^2}\,. \tag{13}$$

This is the differential equation that we've been seeking, the solution of which will be an expression for $s(\theta)$ (*ultimately* $r(\theta)$) and allow us to determine the path of motion of m.
To solve 13, we first note that if $y(\theta)$ is a solution to the equation

$$\frac{d^2y}{d\theta^2} + y = \;0\,,$$

then the function

$$x(\theta) = \;y(\theta) + \frac{k}{c^2}$$

satisfies the equation

$$\frac{d^2x}{d\theta^2} + x = \;\frac{k}{c^2}\,.$$

This is true because

$$\frac{d^2}{d\theta^2}\,\left(y + \frac{k}{c^2}\right) \,+\, \left(y + \frac{k}{c^2}\right) = \;\frac{d^2y}{d\theta^2} + y + \frac{k}{c^2} = \;0 + \frac{k}{c^2} \;=\; \frac{k}{c^2}\;.$$

Hence to solve 13, we only need to solve the equation

$$\frac{d^2s}{d\theta^2} + s = \;0\;.$$

That is, we need only find a function $s(\theta)$ such that

$$\frac{d^2s}{d\theta^2} = \;-s\;. \tag{14}$$

Now 14 says that s is a function with the property that its second derivative is the negative of itself. We already know of two such functions; $\sin\theta$ and $\cos\theta$, and, for any constants A and B, the function $A\sin\theta + B\cos\theta$ also has this property. From this it follows that our solution to (14) will be of the form

$$s = A\sin\theta + B\cos\theta + \frac{k}{c^2} \tag{15}$$

for some constants A and B.

Now we need to find values for A and B such that s will have a local *maximum* at $\theta = 0$. Since $s = 1/r$, what this really means is that we're looking for the constant values that allow r to have a local *minimum* at $\theta = 0$. We make think of this as choosing the constants A and B in such a way that m is closest to M when the path of m crosses the positive real axis, at $t = 0$. Thus we need values for A and B such that

$$\left.\frac{ds}{d\theta}\right|_{\theta=0} = 0\,,$$

and

$$\left.\frac{d^2s}{d\theta^2}\right|_{\theta=0} \leq 0\;.$$

The first derivative of 15 is

$$\frac{ds}{d\theta} = A\cos\theta - B\sin\theta\;,$$

therefore

$$\left.\frac{ds}{d\theta}\right|_{\theta=0} = A\cos(0) - B\sin(0) \;=\; A\;, \tag{16}$$

and the second derivative of 15 is

$$\frac{d^2s}{d\theta^2} = -\,A\sin\theta - B\cos\theta\;,$$

therefore

$$\left.\frac{d^2s}{d\theta^2}\right|_{\theta=0} = -A\sin(0) - B\cos(0) \;=\; -B\;. \tag{17}$$

Hence if we set $A = 0$ and $B \geq 0$, the conditions 16 and 17 are satisfied by

$$s = B\cos\theta + \frac{k}{c^2}\,, \qquad B \geq 0\;. \tag{18}$$

In terms of r, 17 gives us

$$\frac{1}{r} = B\cos\theta + \frac{k}{c^2} \;=\; \frac{c^2\,B\cos\theta + k}{c^2},$$

So

$$r = \frac{c^2}{c^2 B \cos\theta + k} = \frac{c^2/k}{1 + (c^2 B/k)\cos\theta\ .}$$

If we let $a = c^2/k$ and $e = aB$, then our expression for $r(\theta)$ reduces to

$$r = \frac{a}{1 + e\cos\theta} \tag{19}$$

Where $a > 0$ is the *semi-major axis*, and $e \geq 0$ is the *eccentricity*, of the orbit.
(If $e = 0$ then the orbit is a circle of radius a, if $0 \leq e \leq 1$ the orbit is an ellipse with semi-major axis a, if $e = 1$ the orbit is parabolic, and if $e > 1$ the orbit is hyperbolic.)

Note that 19 does not give us values of r and θ for specified values of t, but rather gives us a value of r for any specified value of θ. In other words, our solution does not give us the position of m for a given time t, but it does tell us the location of m as a function of θ. Because we are assuming that $d\theta/dt > 0$, the motion of m will be in the counter-clockwise direction. Hence if we plot the points $z = re^{i\theta}$ for all values of θ in the interval $[-\pi, \pi]$, then the resulting curve will be the path of the orbit of m about M. If $0 < e < 1$, then $e\cos\theta$ has a minimum value of e when $\theta = 0$ and a maximum value of $-e$ when $\theta = \pi$ or $\theta = -\pi$

Thus the minimum value if r is

$$r(0) = \frac{a}{1+e}$$

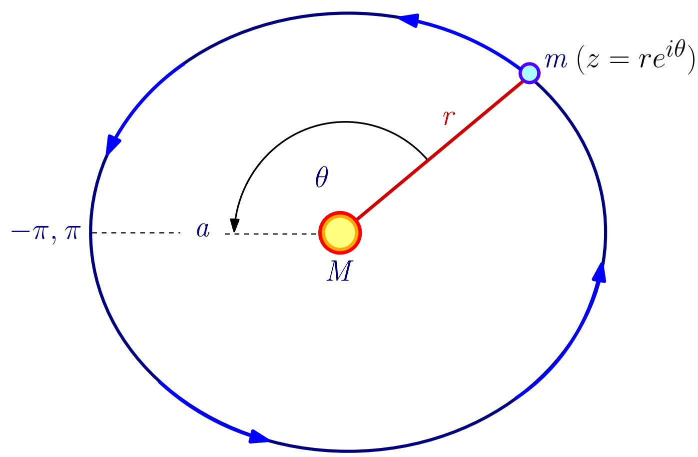
and the maximum value of r is

$$r(-\pi) = r(\pi) = \frac{a}{1-e}$$

Hence the orbit is a closed curve with

$$\frac{a}{1+e} \leq r \leq \frac{a}{1-e}$$

for all θ.



The orbit of m about M.
(*the eccentricity has been highly exagerated*)

3 Statistics for the Solar System

Mercury	$e=0.2056$	$a=$ 0.390	$m=3.303\,\mathrm{e}\,23\ kg$	$P \approx 0.2408$ yrs
Venus	$e=0.0068$	$a=$ 0.720	$m=4.870\,\mathrm{e}\,24\ kg$	$P \approx 0.6152$ yrs
Earth	$e=0.0167$	$a=$ 1.000	$m=5.976\,\mathrm{e}\,23\ kg$	$P = 1.0000$ yrs
Mars	$e=0.0934$	$a=$ 1.520	$m=6.421\,\mathrm{e}\,23\ kg$	$P \approx 1.8810$ yrs
Jupiter	$e=0.0483$	$a=$ 5.200	$m=1.900\,\mathrm{e}\,27\ kg$	$P \approx 11.862$ yrs
Saturn	$e=0.0556$	$a=$ 9.540	$m=5.688\,\mathrm{e}\,26\ kg$	$P \approx 29.957$ yrs
Uranus	$e=0.0461$	$a=$ 19.19	$m=8.684\,\mathrm{e}\,25\ kg$	$P \approx 84.210$ yrs
Neptune	$e=0.0097$	$a=$ 30.06	$m=1.024\,\mathrm{e}\,26\ kg$	$P \approx 64.732$ yrs
Pluto	$e=0.2480$	$a=$ 39.53	$m=1.290\,\mathrm{e}\,22\ kg$	$P \approx 247.68$ yrs

e = *eccentricity*, a = *semi-major axis*, m = *mass* (i.e. $3.3\,\mathrm{e}\,23\,\mathrm{kg} = 3.3 \times 10^{23}\ kg$)
P = *Sidereal Orbital Period*: The time it takes to complete one orbit, with respect to the distant stars and galaxies.

Most of the solar system's asteroids have $0 < e \leq 0.35$. *Short-period comets have* $0 < e < 1$, *most long-period comets have* $e \backsimeq 1$, *while others have* $e > 1$. *This latter group of comets' paths are hyperbolic, hence they never complete a full orbit.*

.

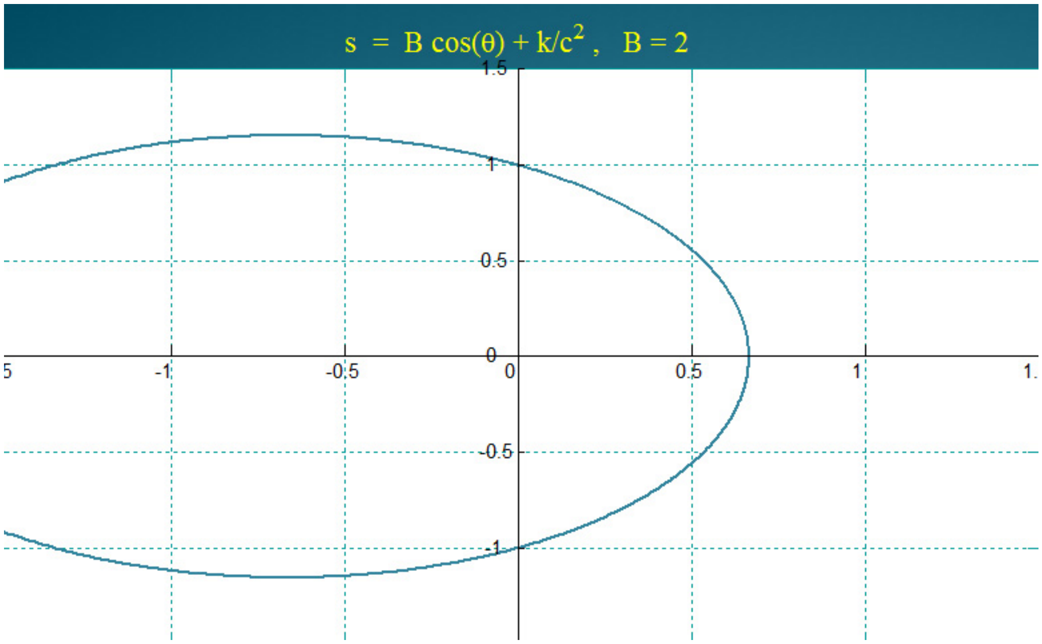


Figure 2: One possible orbit of m about M given by the solution to the second order differential equation $d^2s/d\theta^2 + s = k/c^2$, *or* $d^2(1/r)/d\theta^2 + (1/r) = k/c^2$

4 References

1. *Difference Equations to Differential Equations*, section 7.4.pdf
   Copyright by Dan Sloughter 2000

2. Solar system statistics from: "The Solar System", by Giovanni Caprera, 2003. Published by: Firefly Books Ltd. 2003