An Introduction to Newtonian Celestial Mechanics, and a comparison of Hohmann and Bi-elliptic transfers

March 9, 2019

Contents

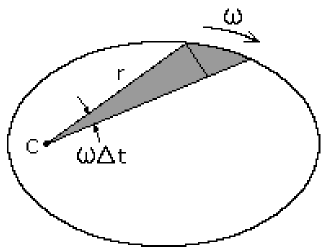


Figure 1: Kepler's Laws demonstated (test image)

1 Introduction

Throughout all of human history, our ancestors have observed the heavenly bodies. Amongst the background of fixed points of light, seven 'plantai', or "wanderers", followed fixed yet seemingly arbitrary paths through the sky. However, these bodies were widely interpreted as the Seven Heavens, which has different connotations in different religions. Some examples include the layers of paradise (Christianity), the residences of the Prophets (Islam), or the seats of the Lord (Judaism). As such, the geocentric view of the universe was the most widely accepted; the Earth lay at the centre of the universe with all of the other bodies orbiting around it along very complicated curves.
However, throughout the late medieval period, Heliocentrism became more and more widely accepted. This culminated in Newton using it as the foundation for his theory of universal gravitation which accurately described the motions of the planets. I will also show that this theory also satisfies Kepler's Laws - a set of laws that were derived from observations of the planets.
In this short essay, I will describe the foundations of Newtonian Celestial Mechanics: Newton's Laws of Motion, and a formulation of the N-Body problem. I will then show how Newton's Law of Gravity upholds all of the laws of motion, as well as maintaining the constraints assigned to a closed system - constant energy, momentum, and angular momentum.
I will then take a slight detour to describe ellipses before demonstrating that Newton's Law of Gravity upholds all of Kepler's Laws. Finally, I will use this work to describe two methods of transferring between circular orbits - the Hohmann and Bi-elliptic transfers - alongside a comparison between them.

2 Basis of Newtonian Mechanics

[1]

2.1 Laws

Newtonian mechanics is founded upon three fundamental laws:

1. Any body in motion remains moving at a constant velocity unless acted on by an exterior force
2. $\mathbf{F} = \dot{\mathbf{p}}$ where $p = m\dot{\mathbf{r}}$
3. Every action has an equal and opposite reaction

2.2 Newton's Law of Gravity

If two bodies A and B have masses m_A and m_B respecively, and the distance betwee them is r, then the force acting on both of them has a magnitude of

$$\frac{Gm_a m_b}{r^2}$$

where G is a constant, and the direction of the force on A is towards B and vice versa.

2.3 Formulation of the N-Body Problem

Let

$$M = \sum_{i=1}^{n} m_i$$

be the total mass of the system, where m_i is the mass of the i^{th} particle, and let

$$\mathbf{R} = \frac{\sum_{i=1}^{n} m_i \mathbf{r_i}}{M}$$

be the Centre of Mass of the system, where r_i is the position of the i^{th} particle. The total force on the i^{th} particle can be given as

$$\mathbf{F_i} = -G \sum_{j=1, j\neq i}^{n} \frac{m_i m_j}{|\mathbf{r_i} - \mathbf{r_j}|^2} \cdot \frac{\mathbf{r_i} - \mathbf{r_j}}{|\mathbf{r_i} - \mathbf{r_j}|}$$

$$\mathbf{F_i} = -G \sum_{j=1, j\neq i}^{n} \frac{m_i m_j}{|\mathbf{r_i} - \mathbf{r_j}|^3} \cdot (\mathbf{r_i} - \mathbf{r_j})$$

From Newton's Second Law, when we take all masses as constants,

$$\mathbf{F_i} = \frac{d}{dt} m_i \dot{\mathbf{r}} = m_i \ddot{\mathbf{r_i}}$$

Therefore

$$m_i \ddot{\mathbf{r_i}} = -G \sum_{j=1, j\neq i}^{n} \frac{m_i m_j}{|\mathbf{r_i} - \mathbf{r_j}|^3} \cdot (\mathbf{r_i} - \mathbf{r_j})$$

$$\ddot{\mathbf{r_i}} = -G \sum_{j=1, j\neq i}^{n} \frac{m_j}{|\mathbf{r_i} - \mathbf{r_j}|^3} \cdot (\mathbf{r_i} - \mathbf{r_j})$$

Let $\mathbf{r_{ij}} = \mathbf{r_i} - \mathbf{r_j}$ and $r_{ij} = |\mathbf{r_{ij}}|$. Then we get

$$\ddot{\mathbf{r}}_\mathbf{i} = -G \sum_{j=1, j\neq i}^{n} \frac{m_j \mathbf{r_{ij}}}{r_{ij}^3}$$

If $\mathbf{r_i} = x_i \mathbf{e_x} + y_i \mathbf{e_y} + z_i \mathbf{e_z}$ where $\mathbf{e_x}$, $\mathbf{e_y}$, and $\mathbf{e_z}$ make up a basis, then we can define the grad function as

$$\nabla_i = \mathbf{e_x} \frac{\partial}{\partial x_i} + \mathbf{e_y} \frac{\partial}{\partial y_i} + \mathbf{e_z} \frac{\partial}{\partial z_i}$$

Define the gravitational potential energy of a body within a system as the energy required to move said body from infinity to its current position. This can be expressed for one body as

$$V(\mathbf{r}) = -\int_{\infty}^{r} \mathbf{F} \cdot d\mathbf{r} = -\int_{\infty}^{r} -\frac{GMm}{r^2} \cdot \mathbf{r} \cdot d\mathbf{r}$$

$$V(\mathbf{r}) = -\int_{\infty}^{r} -\frac{GMm}{r^2} \cdot dr = -\frac{GMm}{r}$$

Extrapolating this method out for all bodies in a system, we get

$$V(\mathbf{r}) = -G \sum_{j=1}^{n} \sum_{k=j+1}^{n} \frac{m_j m_k}{r_{jk}} = -\frac{1}{2} G \sum_{j=1}^{n} \sum_{k=1, k\neq j}^{n} \frac{m_j m_k}{r_{jk}}$$

Since everything besides r_{jk} is constant,

$$\nabla_i V(\mathbf{r}) = -\frac{1}{2} G \sum_{j=1}^{n} \sum_{k=1, k\neq j}^{n} m_j m_k \cdot \nabla_i \left(\frac{1}{r_{jk}} \right)$$

The only time where $\nabla_i(\frac{1}{r_{jk}})$ is nonzero is when either j or k is equal to i. Thereforewe can cancel all non-zero terms out of our summation to get

$$\nabla_i V(\mathbf{r}) = -G \sum_{j=1, j\neq i}^{n} m_i m_j \cdot \nabla_i \left(\frac{1}{r_{ij}} \right)$$

where

$$\nabla_i (\frac{1}{r_{ij}}) = \nabla_i ((x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2)^{-\frac{1}{2}}$$

$$= -\frac{(x_i - x_j)\mathbf{e_x} + (y_i - y_j)\mathbf{e_y} + (z_i - z_j)\mathbf{e_z}}{r_{ij}^3} = -\frac{\mathbf{r_{ij}}}{r_{ij}^3}$$

This then gives

$$\nabla_i V(\mathbf{r}) = G \sum_{j=1, j\neq i}^{n} m_i m_j \cdot \frac{\mathbf{r_{ij}}}{r_{ij}^3}$$

Comparing this with

$$\ddot{\mathbf{r}}_{\mathbf{i}} = -G \sum_{j=1, j\neq i}^{n} \frac{m_j \mathbf{r}_{\mathbf{ij}}}{r_{ij}^3}$$

we note that

$$\nabla_i V(\mathbf{r}) = -m_i \ddot{\mathbf{r}}_{\mathbf{i}}$$

3 Constants of the System

[2]

There are three constants in any closed system:

1. The Centre of Mass should move at a constant velocity (This is the same as conservation of Linear Momentum)
2. The Angular Momentum should remain constant
3. The total energy should remain constant

I will now demonstrate that Newton's Law of Gravity upholds these constants.

3.1 Centre of Mass

By summing all forces,

$$\sum_{i=1}^{n} m_i \ddot{r}_i = -G \sum_{(i,j)\leq n, i\neq j} \frac{m_i m_j \mathbf{r}_{\mathbf{ij}}}{r_{ij}^3}$$

where the left hand side is the sum of the overall force on each body, and the right hand side is the sum of all gravitational forces. By Newton's Third Law, we know that the right hand side consists of pairs of forces, which cancel out. Therefore

$$\sum_{i=1}^{n} m_i \ddot{\mathbf{r}}_{\mathbf{i}} = \mathbf{0}$$

Integrating both sides twice gives

$$\sum_{i=1}^{n} m_i \mathbf{r}_{\mathbf{i}} = \mathbf{a}t + \mathbf{b}$$

where **a** and **b** are constants of integration. This results in

$$\mathbf{R} = \frac{\sum_{i=1}^{n} m_i \mathbf{r}_{\mathbf{i}}}{M} = \frac{\mathbf{a}t + \mathbf{b}}{M}$$

Therefore the Centre of Mass moves at a constant velocity.

3.2 Angular Momentum

We can express the total angular momentum as

$$\mathbf{L} = \sum_{i=1}^{n} (m_i \mathbf{r_i} \times \dot{\mathbf{r_i}})$$

Differentiating both sides gives

$$\dot{\mathbf{L}} = \frac{d}{dt}\left(\sum_{i=1}^{n} (m_i \mathbf{r_i} \times \dot{\mathbf{r_i}})\right)$$

$$\dot{\mathbf{L}} = \sum_{i=1}^{n} (m_i \dot{\mathbf{r_i}} \times \dot{\mathbf{r_i}}) + \sum_{i=1}^{n} (m_i \mathbf{r_i} \times \ddot{\mathbf{r_i}})$$

All the terms in the left hand side are zero (Since the two vectors are the same therefore parallel to each other). This then gives

$$\dot{\mathbf{L}} = \sum_{i=1}^{n} (m_i \mathbf{r_i} \times \ddot{\mathbf{r_i}})$$

Substituting the value of $\ddot{\mathbf{r_i}}$ in, we get

$$\dot{\mathbf{L}} = \sum_{i=1}^{n} m_i \mathbf{r_i} \times \left(-G \sum_{j=1, j\neq i}^{n} \frac{m_j \mathbf{r_{ij}}}{r_{ij}^3}\right) = -G \sum_{(i,j)\leq n, i\neq j} \frac{m_i m_j \cdot (\mathbf{r_i} \times \mathbf{r_{ij}})}{r_{ij}^3}$$

$$\dot{\mathbf{L}} = -G \sum_{(i,j)\leq n, i\neq j} \frac{m_i m_j \cdot (\mathbf{r_i} \times (\mathbf{r_i} - \mathbf{r_j}))}{r_{ij}^3} = G \sum_{(i,j)\leq n, i\neq j} \frac{m_i m_j \cdot (\mathbf{r_i} \times \mathbf{r_j})}{r_{ij}^3}$$

All the terms in the right hand sum come in equal and opposite pairs, so they cancel out. Therefore

$$\dot{\mathbf{L}} = \mathbf{0}$$

Therefore the angular momentum is constant.

3.3 Energy

The total work done on each body can be expressed as

$$\sum_{i=1}^{n} m_i \ddot{\mathbf{r_i}} \cdot \dot{\mathbf{r_i}} = -\sum_{i=1}^{n} \dot{\mathbf{r_i}} \cdot \nabla_i V = -\sum_{i=1}^{n} \left(\dot{x}_i \frac{\partial V}{\partial x_i} + \dot{y}_i \frac{\partial V}{\partial y_i} + \dot{z}_i \frac{\partial V}{\partial z_i} \right)$$

V is a function dependent on all the positions of every particle, therefore by the chain rule

$$\sum_{i=1}^{n} m_i \ddot{\mathbf{r_i}} \cdot \dot{\mathbf{r_i}} = -\frac{dV}{dt}$$

Figure 2: An example of an ellipse

Integrating both sides gives

$$\int \sum_{i=1}^{n} m_i \ddot{\mathbf{r}}_\mathbf{i} \cdot \dot{\mathbf{r}}_\mathbf{i} dt = -V + E$$

$$\frac{1}{2} \sum_{i=1}^{n} m_i \dot{\mathbf{r}}_\mathbf{i} \cdot \dot{\mathbf{r}}_\mathbf{i} = -V + E$$

where E is a constant of the integration Note that the left hand side is the sum of all the kinetic energy. Therefore we can write

$$T = \frac{1}{2} \sum_{i=1}^{n} m_i |\dot{\mathbf{r}}_\mathbf{i}|^2 = -V + E$$

$$E = T + V$$

Therefore the total energy of the system is constant.

4 Ellipses

A circle can be constructed by ensuring that the distance to a point called a focus remains constant. In the same way, we can construct an ellipse by choosing two foci and ensuring that for any point on your curve, the distance to one focus plus the distance to the other focus remains constant. If we label these distances as d_1 and d_2, we are therefore keeping $d_1 + d_2$ constant.

4.1 Equation for an ellipse in Cartesian coordinates

Without loss of generality, in figure 2, we begin by placing our two foci along the horizontal axis at points $(-c, 0)$ and $(c, 0)$. We then denote the **semi-major radius** as a, and the **semi-minor radius** as b. Choose a point on the ellipse with co-ordinates (x, y). This then gives:

$$d_1 = d((-c, 0), (x, y)) = \sqrt{(x + c)^2 + (y^2)}$$

$$d_2 = d((c, 0), (x, y)) = \sqrt{(x - c)^2 + (y^2)}$$

$d_1 + d_2$ is constant, so choosing $(x, y) = (a, 0)$, we have

$$d_1 + d_2 = \sqrt{(a + c)^2 + (0^2)} + \sqrt{(a - c)^2 + (0^2)} = (a + c) + (a - c) = 2a$$

for any choice of (x, y). We can then say

$$d_1 + d_2 = \sqrt{(x + c)^2 + (y^2)} + \sqrt{(x - c)^2 + (y^2)} = 2a$$

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

Squaring both sides gives

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

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

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

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

Squaring both sides again gives

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

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

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

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

Dividing through by a^2-c^2 gives

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

Note that if we choose $(x, y) = (b, 0)$, then the two triangles formed by this point, the origin, and the two foci are congruent (as the shape is symmetrical), meaning that $d_1 = d_2$ (in this instance). Further, we know that

$$d_1 = d_2 = \sqrt{c^2+b^2}$$

Therefore we know that

$$d_1+d_2 = 2\sqrt{c^2+b^2} = 2a$$

Squaring both sides gives the following relationship between the position of the foci, the semi-major radius and the semi-minor radius:

$$a^2 = b^2+c^2$$

This allows us to rewrite the equation of our ellipse to be:

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

4.2 Equation for an ellipse in Polar coordinates centred at one focus

However, all of our observations of orbits are measured in polar co-ordinates, and are measured from one of the foci (i.e. the position of a satellite as seen from Earth). Without loss of generality, let us choose the focus at $(c, 0)$ to be the focus that we are measuring from. Therefore we have to substitute x and y out for functions of r and θ, where r is the distance of the point from our focus, and θ is the angle formed between the positive horizontal axis and the line to the point.
More specifically,

$$x - c = r\cos\theta$$

$$y = r\sin\theta$$

This gives us

$$\frac{(r\cos\theta + c)^2}{a^2} + \frac{(r\sin\theta)^2}{a^2 - c^2} = 1$$

$$(r^2\cos^2\theta + 2cr\cos\theta + c^2)(a^2 - c^2) + a^2r^2\sin^2\theta = a^2(a^2 - c^2)$$

$$a^2r^2\cos^2\theta + 2a^2cr\cos\theta + a^2c^2 - c^2r^2\cos^2\theta - 2c^3r\cos\theta - c^4 + a^2r^2\sin^2\theta = a^4 - a^2c^2$$

$$a^2r^2 + 2cr\cos\theta(a^2 - c^2) - c^2r^2\cos^2\theta = a^4 - 2a^2c^2 + c^4 = (a^2 - c^2)^2$$

This gives us a quadratic in $(a^2 - c^2)$:

$$(a^2 - c^2)^2 - 2cr\cos\theta(a^2 - c^2) + (c^2r^2\cos^2\theta - a^2r^2) = 0$$

We can solve this to get

$$(a^2 - c^2) = \frac{2cr\cos\theta \pm \sqrt{(-2cr\cos\theta)^2 - 4(1)(c^2r^2\cos^2\theta - a^2r^2)}}{2}$$

$$(a^2 - c^2) = cr\cos\theta \pm \sqrt{c^2r^2\cos^2\theta - c^2r^2\cos^2\theta + a^2r^2}$$

$$(a^2 - c^2) = cr\cos\theta \pm ar$$

Note that by construction, $c < a$, therefore $cr\cos\theta < ar$. If we were to use $cr\cos\theta - ar$, we would then get

$$(a^2 - c^2) = b^2 < 0 \implies \perp$$

Therefore we have

$$(a^2 - c^2) = cr\cos\theta + ar$$

Rearranging to get r gives

$$r = \frac{a^2 - c^2}{a + c\cos\theta} = \frac{a - \frac{c^2}{a}}{1 + \frac{c}{a}\cos\theta}$$

Let $p = a - \frac{c^2}{a}$ and define the eccentricity of our ellipse to be given as $e = \frac{c}{a}$. This gives us

$$r = \frac{p}{1 + e\cos\theta}$$

We can define the **periapsis** r_p and the **apoapsis** r_a to be the closest and furthest points on our curve. By construction, we know that

$$r_p = a - c$$

and

$$r_a = a + c$$

Therefore we have

$$r_a - r_p = 2c$$

$$r_a + r_p = 2a$$

Therefore we can write

$$e = \frac{c}{a} = \frac{2c}{2a} = \frac{r_a - r_p}{r_a + r_p}$$

This gives us a way to calculate the eccentricity of orbits via observations. The eccentricity of an ellipse can be thought of as a measure of how elongated an ellipse is. For example, the eccentricity of the Earth's orbit is 0.0167 ($r_p =$ 0.9832AU, $r_a = 1.0167$AU), whilst the eccentricity of the orbit followed by Halley's Comet is 0.9671 ($r_p = 0.586$AU, $r_a = 35.082$AU).

4.3 Area of an ellipse

Consider our Cartesian equation for our ellipse:

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

Note that the values of x are between $-a$ and a. We can also rearrange the above equation to get

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

This means that our values of y lie in the range

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

Therefore we can say that the region, R, inside our ellipse is

$$R = \left\{(x, y)\middle| x \in [-a, a], y \in \left[-b\sqrt{1 - \frac{x^2}{a^2}}, b\sqrt{1 - \frac{x^2}{a^2}}\right]\right\}$$

This then gives us the area, A, as the integral

$$A = \int_{-a}^{a} dx \int_{-b\sqrt{1-\frac{x^2}{a^2}}a}^{b\sqrt{1-\frac{x^2}{a^2}}} dy = \int_{-a}^{a} 2b\sqrt{1-\frac{x^2}{a^2}}dx$$

We can integrate this by using the substitution $x = a\sin(u)$ This gives $dx = a\cos(u)du$ and our integral is now between $\frac{-\pi}{2}$ and $\frac{\pi}{2}$. Therefore we have

$$A = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2b\sqrt{1-\frac{a^2\sin^2(u)}{a^2}}a\cos(u)du = ab\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2\sqrt{1-sin^2(u)}\cos(u)du = ab\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2\cos^2(u)du$$

Note that $2\cos^2(u) - 1 = \cos(2u)$, therefore we can write

$$A = ab\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} [\cos(2u)+1]du = ab\left[\frac{1}{2}\sin(2u)+u\right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \pi ab$$

5 Kepler's Laws

Kepler managed to generate three laws of motion under gravity based on his observations of the planets:

1. **All orbits are ellipses with the parent body as one focus.** This gives the expected equation for the motion of the planets (in polar coordinates) as

$$r = \frac{p}{1+e\cos\theta}$$

   where p is a fixed constant, and e is the eccentricity of the ellipse. Note θ is measured from the planet's closest approach.

2. **A line joining a planet and the parent body sweeps out equal areas in an equal amount of time.**

3. **The square of the orbital period is proportional to the cube of the orbit's semi-major radius**, or $T^2 \propto a^3$, where a is the semi-major radius.

We will now proceed to show that these laws hold under Newtonian Mechanics.

5.1 Foundational Work

In order to prove the laws, we must begin by defining a new set of basis vectors: $\mathbf{e_r}$, $\mathbf{e_\theta}$, and $\mathbf{e_z}$, where

$$\mathbf{e_r} = \frac{\mathbf{r}}{r}$$

$$\mathbf{e_\theta} = \mathbf{e_z} \times \mathbf{e_r}$$

Under out previous basis of $\mathbf{e_x}$, $\mathbf{e_y}$, and $\mathbf{e_z}$, the two new vectors can be expressed as

$$\mathbf{e_r} = (\cos\theta, \sin\theta, 0)$$

$$\mathbf{e_\theta} = (-\sin\theta, \cos\theta, 0)$$

Therefore we can write

$$\mathbf{r} = r\mathbf{e_r}$$

$$\mathbf{v} = \dot{\mathbf{r}} = \frac{d}{dt}(r\mathbf{e_r}) = \dot{r}\mathbf{e_r} + r\dot{\mathbf{e_r}}$$

To get $\dot{\mathbf{e_r}}$, we use our old basis:

$$\dot{\mathbf{e_r}} = \frac{d}{dt}(\cos\theta, \sin\theta, 0) = (-\dot{\theta}\sin\theta, \dot{\theta}\cos\theta, 0) = \dot{\theta}(-\sin\theta, \cos\theta, 0) = \dot{\theta}\mathbf{e_\theta}$$

Therefore

$$\mathbf{v} = \dot{r}\mathbf{e_r} + r\dot{\theta}\mathbf{e_\theta}$$

Differentiating again gives

$$\mathbf{a} = \frac{d}{dt}(\dot{r}\mathbf{e_r} + r\dot{\theta}\mathbf{e_\theta}) = \ddot{r}\mathbf{e_r} + \dot{r}\dot{\mathbf{e_r}} + \dot{r}\dot{\theta}\mathbf{e_\theta} + r\ddot{\theta}\mathbf{e_\theta} + r\dot{\theta}\dot{\mathbf{e_\theta}}$$

To get $\dot{\mathbf{e_\theta}}$, we again use our old basis:

$$\dot{\mathbf{e_\theta}} = \frac{d}{dt}(-\sin\theta, \cos\theta, 0) = (-\dot{\theta}\cos\theta, -\dot{\theta}\sin\theta, 0) = -\dot{\theta}(\cos\theta, \sin\theta, 0) = -\dot{\theta}\mathbf{e_r}$$

Therefore we have

$$\dot{\mathbf{e_r}} = \dot{\theta}\mathbf{e_\theta}$$

and

$$\dot{\mathbf{e_\theta}} = -\dot{\theta}\mathbf{e_r}$$

Substituting these into our equation of acceleration gives

$$\mathbf{a} = \ddot{r}\mathbf{e_r} + \dot{r}\dot{\theta}\mathbf{e_\theta} + \dot{r}\dot{\theta}\mathbf{e_\theta} + r\ddot{\theta}\mathbf{e_\theta} - r\dot{\theta}^2\mathbf{e_r}$$

Therefore

$$\mathbf{a} = \mathbf{e_r}(\ddot{r} - r\dot{\theta}^2) + \mathbf{e_\theta}(2\dot{r}\dot{\theta} + r\ddot{\theta})$$

In our context, the only force present is gravity, which is a radial force. Therefore we can say that

$$\mathbf{a} = \frac{-GM}{r^2}\mathbf{e_r}$$

Which gives us

$$\frac{-GM}{r^2}\mathbf{e_r} = \mathbf{e_r}(\ddot{r} - r\dot{\theta}^2) + \mathbf{e_\theta}(2\dot{r}\dot{\theta} + r\ddot{\theta})$$

$\mathbf{e_r}$ and $\mathbf{e_\theta}$ are orthogonal by construction, therefore we can solve this equation component-wise, giving us

$$\frac{-GM}{r^2} = \ddot{r} - r\dot{\theta}^2$$

and

$$0 = 2\dot{r}\dot{\theta} + r\ddot{\theta}$$

Multiplying this second equation by r gives

$$0 = 2r\dot{r}\dot{\theta} + r^2\ddot{\theta} = \frac{d}{dt}(r^2\dot{\theta})$$

Therefore $r^2\dot{\theta}$ is constant. Note that angular momentum is given by $\mathbf{L} = \mathbf{r} \times \mathbf{v}$.

$$\mathbf{L} = \mathbf{r} \times \mathbf{v} = r\mathbf{e_r} \times (\dot{r}\mathbf{e_r} + r\dot{\theta}\mathbf{e}_\theta) = (r\mathbf{e_r} \times \dot{r}\mathbf{e_r}) + (r\mathbf{e_r} \times r\dot{\theta}\mathbf{e}_\theta) = r^2\dot{\theta}\mathbf{e_z}$$

This gives $|\mathbf{L}| = r^2\dot{\theta}$, which we know is a constant. Therefore the angular momentum of this system is constant.

5.2 Kepler's First Law

Consider the first equation given at the end of our foundation work:

$$\frac{-GM}{r^2} = \ddot{r} - r\dot{\theta}^2$$

along with

$$h := |L| = r^2\dot{\theta}$$

Therefore we have that

$$\dot{\theta} = \frac{h}{r^2}$$

Substituting this into our first equation gives

$$\frac{-GM}{r^2} = \ddot{r} - \frac{h^2}{r^3}$$

Let $u = \frac{1}{r}$ so that $u = u(\theta)$ and $\theta = \theta(t)$. Note that real orbits do not pass through the body they are orbiting around, so u is always well defined. Using this, we can now write

$$\dot{r} = \frac{dr}{dt} = \frac{d(\frac{1}{u})}{dt} = \frac{d(\frac{1}{u})}{du}\frac{du}{dt} = -\frac{1}{u^2}\frac{du}{dt} = -r^2\frac{du}{dt} = -r^2\frac{du}{d\theta}\frac{d\theta}{dt} = -h\frac{du}{d\theta}$$

h is a constant, so we can now write

$$\ddot{r} = -h\frac{d}{dt}\left(\frac{du}{d\theta}\right) = -h\frac{d}{d\theta}\left(\frac{du}{d\theta}\right) \cdot \frac{d\theta}{dt} = -h\dot{\theta}\frac{d^2u}{d\theta^2}$$

We already know that

$$\dot{\theta} = \frac{h}{r^2} = hu^2$$

Therefore we can write

$$\ddot{r} = -h^2u^2\frac{d^2u}{d\theta^2}$$

Substituting this into the first equation gives

$$-GMu^2 = -h^2u^2\frac{d^2u}{d\theta^2} - h^2u^3$$

$$GM = h^2\frac{d^2u}{d\theta^2} + h^2u$$

$$\frac{GM}{h^2} = \frac{d^2u}{d\theta^2} + u$$

This is solvable using our work from "Differential Equations", giving us the solution

$$u(\theta) = \frac{GM}{h^2}[1 + e\cos(\theta - \theta_0)]$$

where both e and θ_0 are arbitrary constants. Since θ_0 represents a rotation of our system, without loss of generality we can let $\theta_0 = 0$. Let $p = \frac{h^2}{GM}$, so that we can write

$$u(\theta) = \frac{1}{r(\theta)} = \frac{1}{p}[1 + e\cos(\theta)]$$

giving

$$r(\theta) = \frac{p}{1 + e\cos(\theta)}$$

which gives shapes dependent on the value of e (note $e > 0$ in reality):
$e < 1 \implies$ Ellipse
$e = 1 \implies$ Parabola
$e > 1 \implies$ Hyperbola

Note that the latter two of these do not give orbits, instead they give the path followed by an object on an escape trajectory, so we shall only consider $e < 1$. Therefore all orbits are elliptical.

5.3 Kepler's Second Law

Suppose that between times t and $t + \delta t$, an angle of $\delta\theta$ is swept out. By approximating the area of this segment as the area of a triangle, we get that

$$\delta A \approx \frac{1}{2} \cdot r \cdot r\delta\theta = \frac{1}{2}r^2\delta\theta$$

Dividing through both sides by δt, we get

$$\frac{\delta A}{\delta t} \approx \frac{1}{2}r^2\frac{\delta\theta}{\delta t}$$

Taking the limit as $\delta t \to 0$ gives the equation

$$\frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt} = \frac{1}{2}r^2\dot{\theta} = \frac{1}{2}|\mathbf{L}|$$

which we know is constant. This gives that equal areas are swept out in equal amounts of time, proving Kepler's Second Law. Furthermore, it shows that this law is a direct result of the conservation of angular momentum.

5.4 Kepler's Third Law

From our proof of Kepler's First law, we have that

$$r(\theta) = \frac{p}{1 + e\cos(\theta)}$$

We know that the area is given by $A = \pi ab$. Also, since $\frac{dA}{dt}$ which is constant, we know that

$$\frac{A}{T} = \frac{dA}{dt}$$

where T is the orbital period. Therefore we have

$$T = \frac{A}{\frac{dA}{dt}} = \frac{\pi ab}{\frac{1}{2}h} = \frac{2\pi ab}{h}$$

Note that $a^2 - c^2 = b^2$, therefore

$$b = \sqrt{a^2 - c^2} = a\sqrt{1 - \frac{c^2}{a^2}} = a\sqrt{1 - e^2}$$

Therefore

$$T = \frac{2\pi a^2\sqrt{1 - e^2}}{h}$$

Also recall that

$$p = a - \frac{c^2}{a} = a(1 - \frac{c^2}{a^2}) = a(1 - e^2)$$

Therefore

$$\sqrt{1 - e^2} = p^{\frac{1}{2}}a^{-\frac{1}{2}}$$

Substituting this into our equation for T gives

$$T = \frac{2\pi a^{\frac{3}{2}}p^{\frac{1}{2}}}{h}$$

Squaring both sides gives

$$T^2 = \frac{4\pi^2 a^3 p}{h^2}$$

Resulting in

$$T^2 \propto a^3$$

as required.

6 Useful Orbital Equations

For the next section, I will be using a variety of orbital formula. As such, this section will be focussed on generating these formula.

6.1 Equation linking orbital velocity, position, and semi-major radius

We will now use the fact that energy and angular momentum are conserved to get a very useful result. Consider the energy at an orbit's periapsis and apoapsis (in this section, anything with a in the subscript is a value measured at the apoapsis, and similarly with p for the periapsis):

$$E = \frac{1}{2}mv_a^2 - \frac{GMm}{r_a} = \frac{1}{2}mv_p^2 - \frac{GMm}{r_p}$$

Some simple algebra of the last two equations gives

$$\frac{1}{2}\left(v_a^2 - v_p^2\right) = GM\left(\frac{1}{r_a} - \frac{1}{r_p}\right)$$

Also consider the angular momentum at an orbit's periapsis and apoapsis:

$$h = mr_av_a = mr_pv_p$$

We then have

$$r_av_a = r_pv_p$$

giving

$$v_p = \frac{r_av_a}{r_p}$$

Substituting this into our first equation gives

$$\frac{1}{2}\left(v_a^2 - \left(\frac{r_av_a}{r_p}\right)^2\right) = GM\left(\frac{1}{r_a} - \frac{1}{r_p}\right)$$

$$\frac{v_a^2}{2}\left(1 - \left(\frac{r_a}{r_p}\right)^2\right) = GM\left(\frac{r_p - r_a}{r_ar_p}\right)$$

$$\frac{v_a^2}{2}\left(\frac{r_p^2 - r_a^2}{r_p^2}\right) = GM\left(\frac{r_p - r_a}{r_ar_p}\right)$$

$$\frac{v_a^2}{2} = GM\left(\frac{r_p - r_a}{r_ar_p}\right)\left(\frac{r_p^2}{r_p^2 - r_a^2}\right) = GM\left(\frac{r_p}{r_a(r_p + r_a)}\right)$$

Note that $2a = r_p + r_a$, therefore

$$\frac{v_a^2}{2} = GM\left(\frac{r_p}{r_a(2a)}\right) = GM\left(\frac{r_a + r_p - r_a}{r_a(2a)}\right) = GM\left(\frac{2a - r_a}{r_a(2a)}\right) = GM\left(\frac{1}{r_a} - \frac{1}{2a}\right)$$

Finally, consider the orbital energy at an arbitrary point, r, where our object has a velocity of v:

$$E = \frac{1}{2}mv^2 - \frac{GMm}{r} = \frac{1}{2}mv_a^2 - \frac{GMm}{r_a}$$

Rearranging the last two equations gives

$$\frac{v^2}{2} - \frac{GM}{r} = \frac{v_a^2}{2} - \frac{GM}{r_a}$$

$$\frac{v^2}{2} - \frac{GM}{r} = GM\left(\frac{1}{r_a} - \frac{1}{2a}\right) - \frac{GM}{r_a} = -\frac{GM}{2a}$$

Therefore we have

$$\frac{v^2}{2} = GM\left(\frac{1}{r} - \frac{1}{2a}\right)$$

$$v^2 = GM\left(\frac{2}{r} - \frac{1}{a}\right)$$

7 Transfer Orbits

When companies are positioning satellites, they typically begin by launching the satellite into a Low Earth Orbit (LEO), which is any orbit with an apoapsis lower than 2,000km (1,200mi). They then position the satellite into the desired higher orbit via a series of 'burns', or instantaneous changes of your velocity. In this section, I will outline two methods of transferring a satellite from a low circular orbit to a higher one, before providing a model detailing which to use in order to minimise the amount of energy the satellite would expend.
The energy expended by a spacecraft primarily comes in the form of changes to kinetic energy. In order to achieve higher orbits, burns are performed in the direction of travel (or **prograde**), and to achieve lower orbits you burn in the opposite direction (or **retrograde**). Therefore we have that changes to the energy of the spacecraft caused by burns are just changes in its kinetic energy, or

$$\Delta E \approx \Delta E_k$$

Further, we know that changes to the kinetic energy are proportional to changes to the square of the velocity, so we then have

$$\Delta E \propto (v + \Delta v)^2 - v^2 = 2v\Delta v + \Delta v^2$$

where v is our starting velocity, and Δv is the instantaneous change of velocity. Notice that our change in velocity Δv is multiplied by our starting velocity, but the energy expended by the spacecraft are only proportional to Δv^2. This means that changes in velocity that occur at higher speeds are more efficient. Therefore performing burns when more of our energy is stored as kinetic energy makes the spacecraft more efficient - this is known as the **Oberth effect**, and this effect is what allows the Bi-elliptic transfer to be more energy efficient (in some cases).

Figure 3: An example of a Hohmann transfer

Figure 4: An example of a Bi-Elliptic transfer

7.1 Hohmann Transfer

The Hohmann Transfer is arguably the simplest transfer method; first make a burn to increase the apoapsis of your orbit to be at the same height as your desired new orbit. Then once you reach your apoapsis, make a secondary burn to circularise / bring your periapsis up to the same value as your apoapsis.

7.2 Bi-Elliptic Transfer

Similar to the Hohmann Transfer, the Bi-elliptic Transfer begins by raising your apoapsis to some height above the desired orbital radius. Then you make a second burn at apoapsis to raise your periapsis to the orbital radius desired, before finally circularising once your reach periapsis. Note that whilst this appears to use more energy (since you are reaching a higher altitude), this can still lead to a lower energy expenditure. Note that this maneuver requires 3 burns instead of the 2 used for Hohmann transfers. Some types of engines can only be ignited a limited number of times, so this extra burn requirement needs to be considered. Also note that this method takes significantly longer to execute, so whilst in some situations it is more energy viable, you might choose a Hohmann transfer in order to save time.

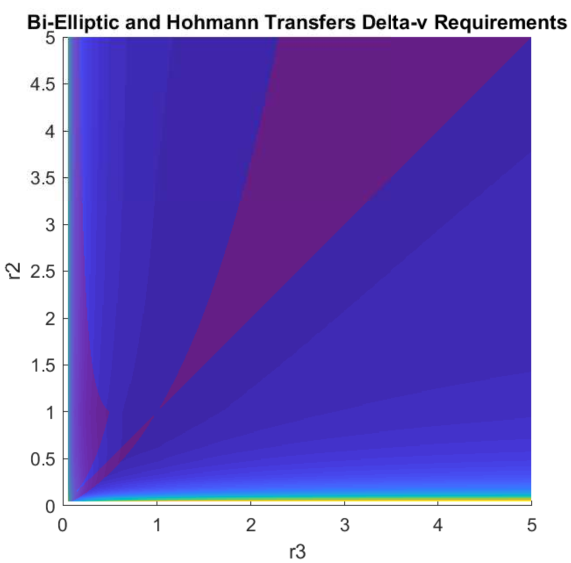
8 A Qualitative Comparison via Matlab

In order to make this comparison, I have created some Matlab code to graph the Bi-elliptic and Hohmann transfer energy requirements. Note that in this code, I have assumed that the mass remains unchanged throughout, and all changes in velocity occur instantaneously. Considering the mass change caused by fuel consumption requires knowledge of the "wet" and "dry" mass of the spacecraft (the mass with and without the fuel respectively), and the engine efficiency. However this would then make the graph constructed less general, and in practice does not change the plot drastically. The assumption of instantaneous burns also does not alter our calculations a lot; usually burn times are far shorter than any orbital periods involved. In this graph, note that the areas shown in red are the areas where the bi-elliptic transfer is more efficient, with $r2$ and $r3$ chosen according to the position on the graph. The other areas are where the Hohmann transfer is more efficient.

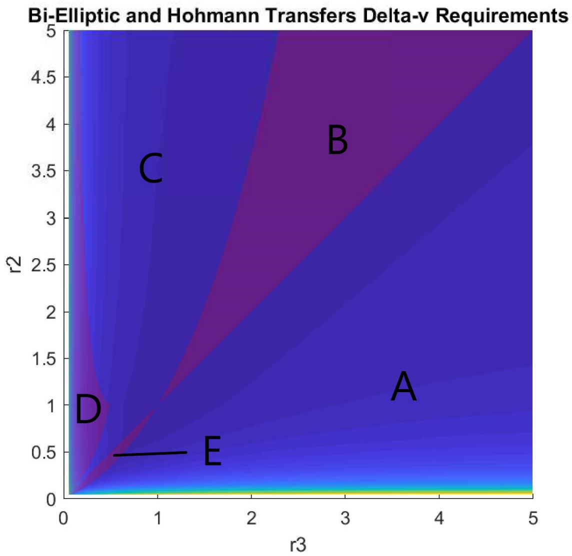
```
%Drawing a graph for both the Hohmann and Bi-elliptic transfer

%r2 - height of transfer orbit's apoapse
```

```
%r3 - desired height
%Note the Hohmann transfer is the case where r2 = r3, and is shown in red

%Begin by defining our mesh for points to sample
step = 0.1;
start = 0.3;
endp = 10;
r2mesh = [start:step:endp];
r3mesh = [start:step:endp];
[r2, r3] = meshgrid(r2mesh,r3mesh);

%Draw the graph for the Bi-elliptic transfer
hold off
BE_E = arrayfun(@(a,b) energy(a,b),r2,r3); %Gives imaginary values when r2<r3
BE_E(imag(BE_E) ~= 0) = NaN;

%Graph labelling
surf(r2,r3,BE_E,'EdgeColor','none');
title('Bi-Elliptic Transfers Delta-v Requirements')
xlabel('r3')
ylabel('r2')
zlabel('Energy Required')
%Having the z axis logarithmically scaled punches out a hole at (1,1), but
%makes the rest of the data clearer
%set(gca,'zscale','log')
pbaspect([1 1 1])
grid off
hold on

%Draw the graph for the Hohmann transfer
H_E = arrayfun(@(a,b) energy(a,b),r2,r2); %Having the transfer radius equal the final rad
H_E(isnan(BE_E)) = NaN;  %Removes extra values not included in first plot

surf(r2,r3,H_E,'FaceColor','r', 'FaceAlpha',0.2, 'EdgeColor','none');
hold off
view(0, 90);
%Defining our functions

%p - current position, Ra - Apoapsis, Rp - Periapsis. Note that if the
%labelling for the apoapsis and periapsis changes, no overall changes occur
%to the program

%The formula for velocity at a given position, as derived
function v = vel(p,Ra,Rp);
a = (Ra + Rp)./2;
v = sqrt(2./p - 1./a);
```

```
end

%Deriving the energy required for a Bi-elliptic transfer
function E = energy(R2,R3);
v1 = vel(1,1,1);
dv1 = vel(1,R2,1) - v1;

v2 = vel(R2,R2,1);
dv2 = vel(R2,R2,R3) - v2;

v3 = vel(R3,R2,R3);
dv3 = vel(R3,R3,R3) - v2;

E = abs(2.*v1.*dv1 + dv1.^2) + abs(2.*v2.*dv2 + dv2.^2) + abs(2.*v3.*dv3 + dv3.^2);

end
```



In the above graphic, the multicoloured area denotes the varies energy requirements of performing Bi-elliptic transfers, where $r2$ is the "middle" orbit apoapse (aka transfer orbit), and $r3$ is the desired orbital radius. The area shaded in red is the region where performing the Hohmann transfer is less efficient than the Bi-elliptic. For the sake of clarity, I will split the regions of the graph into 5 areas: A, B, C, D, and E. A is the lower right-hand blue section, and B is the larger red section. C is the top left blue section, and then D and E are the 2 red sections above and below the 45°line respectively. Note that this entire section is going to be qualitative, and most bounds will simply be referred to "upper" or "lower", since whilst all these curves are distinct, they are unnamed since as far as I can tell, I am the first to generate this graph.



8.1 Region A

This area is blue since it is (almost) always more efficient to go to a higher orbit via either a direct route (Hohmann transfer), or a Bi-elliptic transfer that takes your apoapsis above your desired height. This also means it is always more efficient to perform a single Hohmann transfer instead of two separate ones, each raising your orbital radius. Note that the 45°line bounding A above (almost everywhere) is caused by the equivalence of the Hohmann and Bi-Elliptic transfers along it.

8.2 Region B

In this region, the Bi-elliptic transfer wins out. However, note that the curve bounding the region from above implies that the Bi-elliptic transfer is more efficient provided that your transfer orbit apoapsis is not raised too high. The curve bounding this region above is caused by simply going into such a high orbit that the gains made back from the Oberth effect are not as much as the losses due to completing the initial radius raising burn.

8.3 Region C

Raising the transfer apoapsis too high results in the maneuver becoming inefficient. Note that this region is only joined to region A due to the resolution of the plot - in reality the two regions are only joined at (1,1), where the Hohmann and Bi-elliptic transfers are equivalent. Furthermore, the section of this region that is below $r2 = 1$ also implies that it is not always efficient to perform a burn as a bi-elliptic transfer when decending. If this were the case, it would actually imply a continuous solution - if splitting the burn up was always more efficient, we would be encouraged to split the burn into infinitely many infinitesimal burns.

8.4 Region D

When your desired orbital radius is less than half your starting radius, it is actually more efficient to have a transfer orbital radius slightly higher or lower than your original (Note that along the line $r2 = 1$, the Bi-elliptic and Hohmann transfers are again equal, but are shaded as to fit to their surroundings). Having a transfer radius slightly higher than your initial position allows for a greater percentage of the total δv to be expended lower down, and having a transfer radius lower than your start allows for a percentage of your burn normally conducted at a radius of 1 to be completed lower. Both of these strategies take advantage of the Oberth Effect, where the former is maximising the amount burnt lower down and the latter is minimising the amount burnt higher up. The curve bounding this region above is caused by the same effect as what causes the upper bound of B, whereas the lower bound is caused by the gains made from the Oberth effect being less than the energy expended to lower your orbit by that much.

8.5 Region E

This region is implying that whilst setting your transfer radius slightly higher than your target is inefficient, setting it slightly lower can yield a more efficient solution. Again, this is caused by the Oberth effect. This also means that there are more efficient solutions where going below your target radius is more efficient, but only for maneuvers that aim to decrease your orbital radius.

9 Summary of this comparison

Whilst our simplifications and assumptions make this graph unsuited for finding solutions directly, the various divisions of the graph imply that there do exist scenarios where the bi-elliptic transfer is the more efficient method. Furthermore, this graph suggests where these solutions might exist, and the 'smoothness' of the various curves involved suggests that it is likely possible to obtain these various regions as regions bounded by known curves. If this can be done, our simplifications regarding the mass of the parent body and the value of G would just result in a linear transformation of these equations. However, since satellites have different dry and wet masses, different engine efficiencies, and different thrust outputs, dropping our assumptions on the constant mass of the spacecraft and the instantaneous changes in velocity would result in a unique graph being generated for every spacecraft. Between this and the added effects from factors such as other gravitationally attractive bodies, air resistance (which is certainly

not negligible in low orbits), and radiation pressure, it rapidly becomes appropriate to use iterative methods and simulations. This has the affect of being far more accurate for specified satellites, but losing the general solutions provided by the graph.
In conclusion, whilst this method does have its flaws, this graph demonstrates the general regions where bi-elliptic and Hohmann transfers should be used. As such, it is a useful resource when trying to get a quick idea as to which method to use.

References

[1] Mauri Valtonen and Hannu Kattunen. The three-body problem. *Cambridge University Press*, pages 20–23, 2005.

[2] Mauri Valtonen and Hannu Kattunen. The three-body problem. *Cambridge University Press*, pages 25–27, 2005.