AAE 5626: Orbital Mechanics for Engineers

Mrinal Kumar©

August 15, 2022

The Two Body Problem

1 Kepler's Work

- You may know that Newton "invented" calculus in his monumental masterpiece *PhilosophiæNaturalis Principia Mathematica*, which was first published in 1687. The *Principia* also contained the (Newton's) *universal law of gravitation* - the famous inverse square law, which together with calculus is the foundation for modern analysis of the **two body problem**.

  However, Johannes Kepler (1571 - 1630), who died well before Newton was ever born had already provided a macroscopic view of motion in space: in the form of his famous ***Kepler's laws of planetary motion***. Kepler induced these laws between 1609 and 1619 purely by geometrical analysis of the abundant data gathered meticulously by Tycho Brahe.

- **Kepler's laws of planetary motion**

  1. *Planets orbit the sun in elliptical orbits with the sun at one of the two foci of the ellipse.* See Fig.(1).

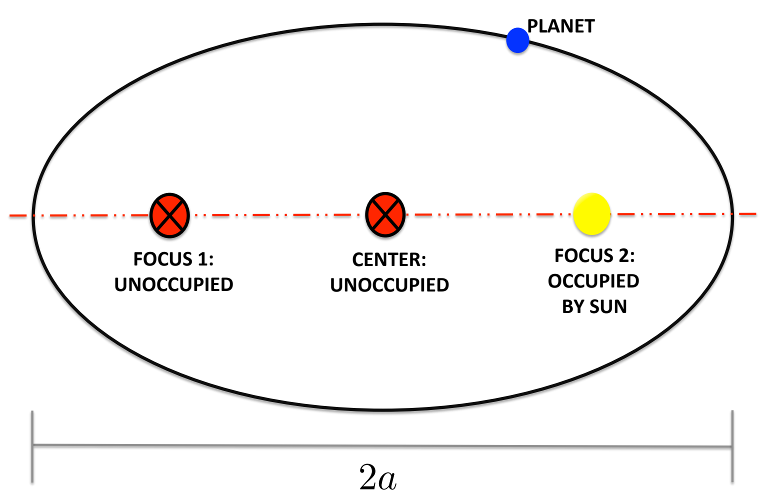


Figure 1: Kepler's First Law of Planetary Motion

The modern day generalization of Kepler's first law can be stated as follows: The orbits of the "two-body problem" are conic sections, i.e. circles, ellipses, parabolas or hyperbolas, depending on the energy-level of the orbit.

2. *The radius sweeps equal areas in equal amounts of time.* See Fig.(2), which clearly shows that in order to satisfy Kepler's second law, the planet must move *faster* when it is closer to the sun than when it is farther away.

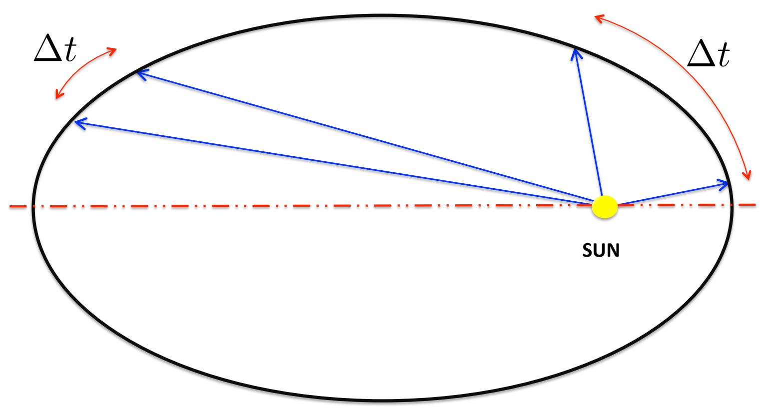


Figure 2: Kepler's Second Law of Planetary Motion

Newton's analytical approach will soon show us that Kepler's second law is equivalent to the conservation of angular momentun.

3. The square of the orbital period is directly proportional to the cube of the semi-major axis of the orbit. In other words, $P^2 \propto a^3$.

2 Newton's Approach to the Two-Body Problem

- Newton took a more analytical approach to the two body problem. Fig.(6) shows the set-up. An inertial frame $\mathcal{I}$ is used to measure the radius vectors of two masses m_1 and m_2. The relative vector, measured from m_1 to m_2 is given by $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$.

- From Newton's second law,

$$m_1\,\ddot{\mathbf{r}}_1^{\mathcal{I}} = \mathbf{F}_1 \tag{1}$$
$$m_2\,\ddot{\mathbf{r}}_2^{\mathcal{I}} = \mathbf{F}_2 \tag{2}$$

  where, $\mathbf{F}_1$ and $\mathbf{F}_2$ represent the total force acting on masses m_1 and m_2 respectively.

- The question is, what is the constitution of $\mathbf{F}_1$ and $\mathbf{F}_2$?

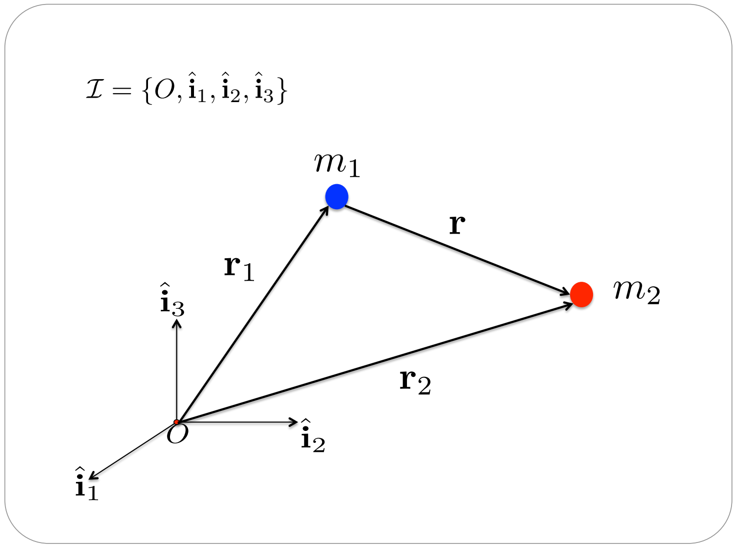


Figure 3: Geometry of the two Body Problem

- Since we are concerned with the study of motion under gravity, we answer the above question as follows:

$$\mathbf{F}_1 = \mathbf{F}_{12} + \mathbf{F}_{d1} \tag{3}$$

  where, $\mathbf{F}_{12}$ is the *gravitational force acting on mass* m_1 *due to the presence of mass* m_2, and $\mathbf{F}_{d1}$ is the sum total of *all other forces* acting on mass m_1, denoted as a "disturbance force". In other words, we will only be concerned with the inter-gravitational forces between m_1 and m_2 and collectively bundle everything else as a disturbance, eventually ignoring it.

- In the same vein, force $\mathbf{F}_2$ can be decomposed as:

$$\mathbf{F}_2 = \mathbf{F}_{21} + \mathbf{F}_{d2} \tag{4}$$

- Moving forward, the question of constitution of $\mathbf{F}_{12}$ (and $\mathbf{F}_{21}$) was given by Newton around 1687:

  Newton's Universal Law of Gravity: *Any two point masses exert on each other an attractive, radial gravitational force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.*

  - By "universal", we mean that the law applies not only to planets, but all other celestial objects, and Earthly objects, like human beings, plants, animals etc., so long as they can be replaced with an equivalent "point mass".
  - By "attractive," we mean that m_2 *pulls* m_1 towards itself (and vice-versa) by virtue of gravity.
  - By "radial", we mean that the direction of the pull is along the line joining m_1 and m_2.

- Therefore, we have:

  - Magnitude of $\mathbf{F}_{12}$: $F_{12} \propto \frac{m_1 m_2}{r^2}$; i.e.,

$$F_{12} = G\frac{m_1 m_2}{r^2} \tag{5}$$

where, G is the **universal gravitational constant**, with value $G = 6.674 \times 10^{-11} Nm^2/kg^2$. This value was determined in 1798 (well after Newton!!) by the experiments of Henry Cavendish.

- Direction of $\mathbf{F}_{12}$: Following the prescription of "attractive" and "radial", we conclude that,

$$\hat{\mathbf{F}}_{12} = \hat{\mathbf{r}} \tag{6}$$

- Combining Eqs.(5) and (6), we come to the final expression for $\mathbf{F}_{12}$:

$$\mathbf{F}_{12} = G\frac{m_1 m_2}{r^2}\hat{\mathbf{r}} = G\frac{m_1 m_2}{r^3}\mathbf{r} \tag{7}$$

- By Newton's third law of equal and opposite reactions, we get (Fig.(4))

$$\mathbf{F}_{21} = G\frac{m_1 m_2}{r^2}(-\hat{\mathbf{r}}) = -G\frac{m_1 m_2}{r^3}\mathbf{r} \tag{8}$$

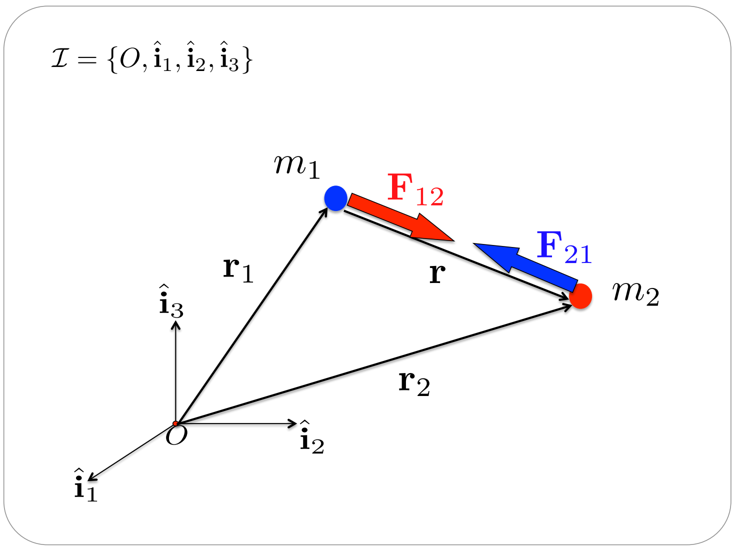


Figure 4: Gravitional Forces

- FACT: Gravity is a weak force! Example: Consider two point masses of 100 kg each (pretty heavy!), separated by a distance of 1 m. From equation (5), the magnitude of the gravitational force between them is $F = 6.674 \times 10^{-11}\frac{100^2}{1} = .667\mu N$.

- Returning to our two-body mechanics model, we have:

$$\begin{aligned} m_1\,\ddot{\mathbf{r}}_1 &= \mathbf{F}_{12} + \mathbf{F}_{d1} \\ &= G\frac{m_1 m_2}{r^3}\mathbf{r} + \mathbf{F}_{d1} \qquad (9) \end{aligned}$$

and,

$$m_2\,\ddot{\mathbf{r}}_2 = -G\frac{m_1 m_2}{r^3}\mathbf{r} + \mathbf{F}_{d2} \tag{10}$$

Note that the superscript $\mathcal{I}$ has been dropped for clarity of expression and the derivatives are understood to be in the inertial frame.

- As hinted above, we will now assume that the universe is comprised of only the two objects shown and that gravity is the only force present. This allows us to set the disturbing forces to zero, giving us the following system:

$$\ddot{\mathbf{r}}_1 = \frac{Gm_2}{r^3}\mathbf{r} \tag{11a}$$

$$\ddot{\mathbf{r}}_2 = -\frac{Gm_1}{r^3}\mathbf{r} \tag{11b}$$

All we need to do now is go ahead and solve the above system of equations!

3 Motion of the Center of Mass

- Recall that the motion of a system of point masses is equivalent to the motion of their **center of mass**. Also recall that the center of mass is a sort of a *weighted-average position* of the entire system of particles. The weights are in proportion to the masses of the particles. Mathematically speaking, the center of mass is defined as:

$$\mathbf{R}_{\text{CM}} \doteq \frac{\sum_{i=1}^{N} m_i \mathbf{r}_i}{\sum_{i=1}^{N} m_i} \tag{12}$$

where, N is the total number of particles in the system, m_i is the mass and $\mathbf{r}_i$ the radius vector of the i^{th} particle. In terms of the total mass of the system, i.e. $M = \sum_{i=1}^{N} m_i$, the center of mass can be re-written as:

$$\mathbf{R}_{\text{CM}} \doteq \sum_{i=1}^{N} \frac{m_i}{M}\mathbf{r}_i = \sum_{i=1}^{N} w_i \mathbf{r}_i \tag{13}$$

where, $w_i \doteq \frac{m_i}{M}$ is the weight associated with the i^{th} particle, clearly in proportion to its mass.

- In the current context, we have two particles, resulting in the following expression for the center of mass:

$$\mathbf{R}_{\text{CM}} = \frac{m_1}{M}\mathbf{r}_1 + \frac{m_2}{M}\mathbf{r}_2 \tag{14}$$

where, $M = m_1 + m_2$. Differentiating (inertially) twice, we get:

$$\ddot{\mathbf{R}}_{\text{CM}} = \frac{m_1}{M}\ddot{\mathbf{r}}_1 + \frac{m_2}{M}\ddot{\mathbf{r}}_2 \tag{15a}$$

$$\underset{[11]}{=} \frac{m_1}{M}\left(\frac{Gm_2}{r^3}\mathbf{r}\right) + \frac{m_2}{M}\left(\frac{Gm_1}{r^3}(-\mathbf{r})\right) \tag{15b}$$

$$= \mathbf{0} \tag{15c}$$

- Amazingly, we find that the center of mass has zero inertial acceleration. Integrating Eq.(15c) twice, we get:

$$\mathbf{v}_{\text{CM}} = \mathbf{v}_0 \quad \text{(velocity of CM)} \tag{15d}$$

$$\mathbf{R}_{\text{CM}} = \mathbf{v}_0 t + \mathbf{R}_0 \quad \text{(radius vector of CM)} \tag{15e}$$

The conclusion is that in the two-body problem (without disturbing forces), **the center of mass exhibits non-accelerating rectilinear motion**. While this may appear to be extraordinarily boring, it is actually very useful. Because we have just found a point with no inertial acceleration - and an excellent candidate for the *origin of our inertial reference frame*. See Fig.(5), in which the origin of the inertial reference frame coincides (without any approximation), with the center of mass of the two-particle system.

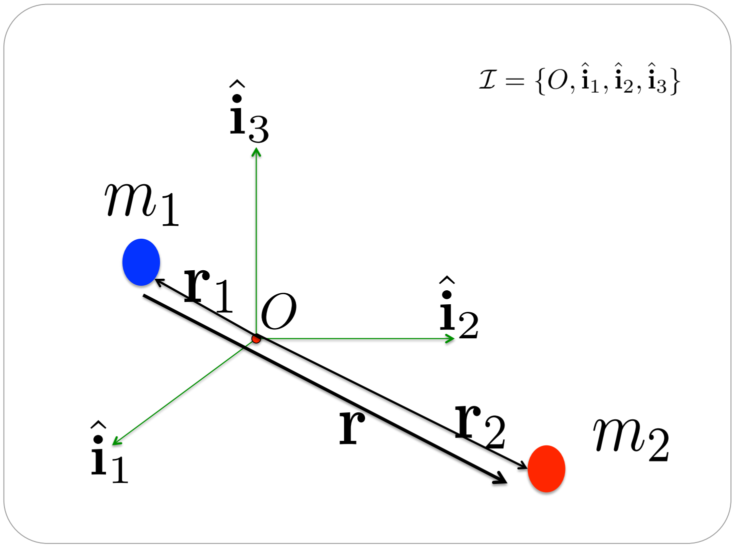


Figure 5: A Good Choice for the Origin of the Inertial Reference Frame

- Often times in the study of two-body motion, the origin of the inertial frame is *chosen* to coincide with one of the two point masses. For example, recall that the Earth-centered inertial frame (ECI) is used to study the motion of spacecraft around the Earth. From our foregoing analysis, the center of the Earth is **not** an "inertial" point. Ideally, the center of mass of the Earth-spacecraft system should be chosen as it has zero inertial acceleration.

  Let us therefore estimate the center of mass of the Earth-spacecraft system. The mass of Earth is $m_1 = 5.97 \times 10^{24}\ kg$, and consider a "heavy" spacecraft of mass $m_2 = 1000\ kg$. Let the spacecraft be in an orbit of 100 km altitude. Since the radius of the Earth is 6378.14 km, the distance between the centers of the two bodies is $6378.14 + 100 = 6478.14\ km$. Following Eq.(14), the center of mass of this system, as measured from the center of the Earth is:

$$\begin{aligned} R_{CM,\oplus/sc} &= 0 + \frac{1000}{1000 + 5.97 \times 10^{24}} 6478.14\ km \\ &= 1.085 \times 10^{-15}\ m \end{aligned}$$

  Clearly, the center of mass is not very far from the center of the Earth! (especially considering the other distances involved in this problem). The important thing is that the center of mass has zero inertial acceleration, **and it is merely 10^{-15} m away from the center of the Earth.** Therefore, we are willing to accept the error induced by shifting the origin of the inertial frame from the center of mass to the center of the Earth, a mere femto-meter away.

- The intuition we get from the above example is that if one of the two bodies is *significantly* heavier than the other one, we may safely assume that the heavier object has "zero" inertial acceleration. It is therefore also OK to use it as the origin of our inertial reference frame.

- The above is not true if the masses of the two bodies are comparable, e.g. the Earth-moon system.

4 Relative Motion

- Our analysis of the two-body system in terms of the center of mass was good, but it doesn't give any insight into the relative motion of the two objects with respect to each other (i.e., *where's the orbit??*)

  Let us therefore consider the ***relative motion***, i.e. the mechanics of the vector $\mathbf{r}$ in Fig.(6). From geometry, $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$. Differentiating twice, we have:

$$\ddot{\mathbf{r}} = \ddot{\mathbf{r}}_2 - \ddot{\mathbf{r}}_1 \tag{16}$$

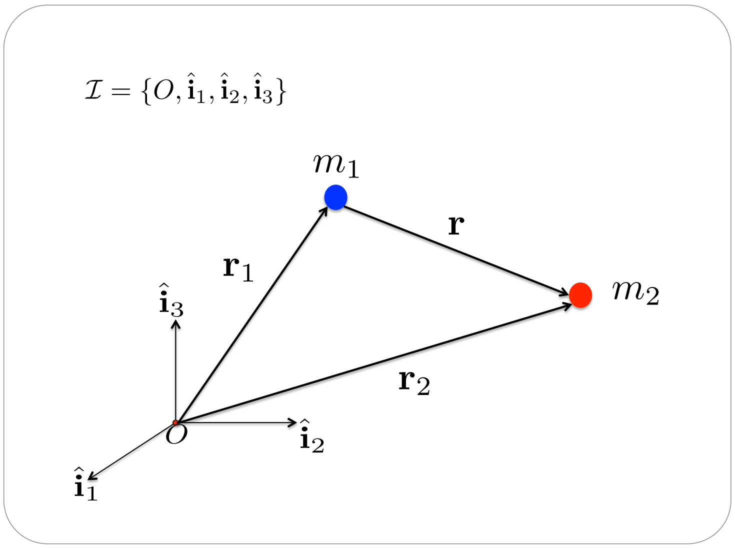


Figure 6: Geometry of the two Body Problem

From Eqs.(9) and (10),

$$\ddot{\mathbf{r}} = \left(-\frac{Gm_1}{r^3}\mathbf{r} + \frac{\mathbf{F}_{d2}}{m_2}\right) - \left(\frac{Gm_2}{r^3}\mathbf{r} + \frac{\mathbf{F}_{d1}}{m_1}\right) \tag{17a}$$

$$= -\frac{G(m_1+m_2)}{r^3}\mathbf{r} + \underbrace{\left(\frac{\mathbf{F}_{d2}}{m_2} - \frac{\mathbf{F}_{d1}}{m_1}\right)}_{=\mathbf{a}_d} \tag{17b}$$

where, $\mathbf{a}_d$ is the disturbance acceleration, i.e. the sum total of all forces *other than the mutual gravitational attraction between* m_1 and m_2. As before, in the case of pure Keplerian motion, we set $\mathbf{a}_d = \mathbf{0}$, to get:

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r} \tag{18}$$

The above is called the ***Fundamental Equation of the Two-Body Problem***, with $\mu \doteq G(m_1+m_2)$, known as the gravitational constant. Note: G (the *universal* gravitational constant) is different from μ, which is simply the gravitational constant.

- Given Eq.(18), all that is left is to solve it! Unfortunately, due to the nonlinear nature of this equation (with respect to $\mathbf{r}$), no closed-form analytical solutions exist. This is a roundabout way of saying that so far no one has succeeded in solving Eq.(18) by hand on a piece of paper. Computers must be used to numerically integrate Eq.(18) (e.g. something like "ode45") to obtain the states $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ as a function of time.

5 Integrals of Motion

- Given a dynamical system, an **integral of motion** is an entity that does not change over time. The states of the system (e.g. position and velocity) evolve with time; however, certain functions of the state remain invariant as time progresses. These functions of the state are called integrals of motion. An example is the energy of a non-dissipative system. Consider the linear spring:

$$\dot{x} = v \tag{19a}$$
$$\dot{v} = -kx \tag{19b}$$

The two equations above can be combined into a single equation of the system: $\ddot{x} = -kx$, which we easily identify as the simple linear spring, with the restorative force $-kx$.

- Now, consider the quantity "$\ddot{x}\dot{x}$":

$$\ddot{x}\dot{x} = -kx\dot{x} \tag{20}$$

The LHS of the above equation is:

$$\ddot{x}\dot{x} = \frac{d}{dt}\left(\frac{1}{2}\dot{x}^2\right) \tag{21}$$

Similarly, the RHS "compactifies" into:

$$-kx\dot{x} = \frac{d}{dt}\left(-\frac{k}{2}x^2\right) \tag{22}$$

Using Eqs.(21) and (22) in Eq.(20), and noting that $\dot{x} = v$, we get:

$$\frac{d}{dt}\left(\frac{v^2}{2} + \frac{kx^2}{2}\right) = 0 \tag{23}$$

finally, upon integration we get the energy "integral of motion":

$$\frac{v^2}{2} + \frac{kx^2}{2} = c \tag{24}$$

What is the above equation telling us? Even as the states (x and v) change over time, a particular function of these states, namely energy, is constant as shown in Fig.(19).

- In general, it is not very easy to find the integrals of motion of a dynamical system. However, **if** they are found, they usually lend tremendous insight into the system under study.

5.1 Integrals of Motion of the Two Body Problem

- Amazingly enough, several integrals of motion of the two body problem are known. We will look at the following three:

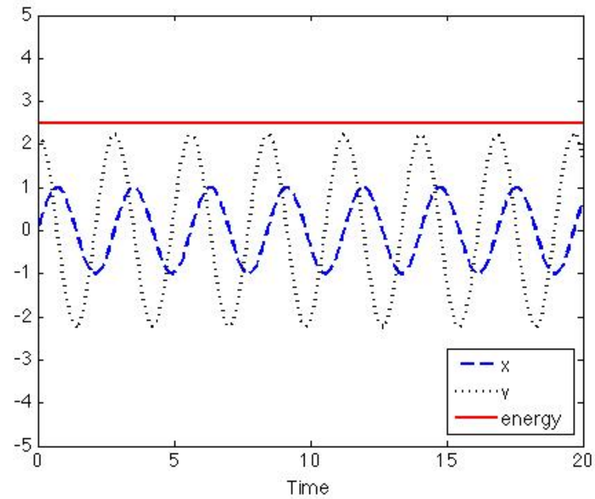


Figure 7: "Energy" is an Integral of Motion for the Linear Spring System ($k = 5$)

5.1.1 Angular Momentum

- Define the angular momentum vector (per unit mass) as follows:

$$\mathbf{h} = \mathbf{r} \times \dot{\mathbf{r}} \tag{25}$$

Suppose we take the inertial time derivative of $\mathbf{h}$ (drop the superscript $\mathcal{I}$ for clarity of expression):

$$\begin{aligned}\dot{\mathbf{h}} &= \frac{d}{dt}(\mathbf{r} \times \dot{\mathbf{r}}) && (26a)\\ &= \underbrace{\dot{\mathbf{r}} \times \dot{\mathbf{r}}}_{=\mathbf{0}} + \mathbf{r} \times \ddot{\mathbf{r}} \\ &\underset{[18]}{=} \mathbf{r} \times \left(-\frac{\mu}{r^3}\mathbf{r}\right) && (26b)\\ &= \mathbf{0} && (26c)\end{aligned}$$

The above exercise tells us that the angular momentum vector is inertially fixed. In other words, it is invariant with time, and is therefore an *integral of motion* of the two body problem.

- Since $\mathbf{h}$ is a vector, it contains three components, i.e. there are three integrals of motion contained in $\mathbf{h}$.

  Since $\mathbf{h}$ is a constant vector, we have $\mathbf{h} = \mathbf{r}(0) \times \dot{\mathbf{r}}(0) = \mathbf{r}(t) \times \dot{\mathbf{r}}(t)$, i.e. it is the same vector at the initial and all future times.

  Note also that at any time t, the vector $\mathbf{h}(t) = \mathbf{r}(t) \times \dot{\mathbf{r}}(t)$ represents the **normal** to a plane containing the vectors $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$. This is very interesting: $\mathbf{h}(t)$ is the normal to the instantaneous plane containing vectors $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ ...... However, it was shown above that $\mathbf{h}(t)$ is the same at ***all*** times, which simply implies that the plane containing the vectors $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ is the same for all times t.

We reach a very important conclusion: The vectors $\mathbf{r}(t)$ and $\dot{\mathbf{r}}(t)$ lie in the same plane at all times. In other words, the two body problem exhibits **planar motion**. The plane of motion is defined by its normal, which is nothing but the constant angular momentum vector.

- Eq.(26b) tells us that angular momentum is conserved in any system in which the acceleration acts in the radial direction, i.e. $\ddot{\mathbf{r}} = f(\mathbf{r})$. This is also called the case of a "central force-field". An example other than gravity is the elctrostatic force, which also acts along the line joining the centers.

- We can prove the Kepler's second law (equal times - equal areas) using the time invariance of $\mathbf{h}$. Consider Fig.(8), which shows the area swept by the vector $\mathbf{r}$ during an infinitesimal time dt (area of the shaded triangle):

$$d\mathbf{A} = \frac{1}{2}\mathbf{r}(t) \times d\mathbf{r}(t) \tag{27}$$

where, $d\mathbf{r}(t) = \mathbf{v}(t)dt$. Therefore,

$$d\mathbf{A} = \frac{1}{2}\mathbf{r}(t) \times \mathbf{v}(t)dt \tag{28}$$

i.e. the rate at which the area is being swept is:

$$\frac{d\mathbf{A}}{dt} = \frac{1}{2}\mathbf{r}(t) \times \mathbf{v}(t) = \frac{1}{2}\mathbf{r}(t) \times \dot{\mathbf{r}}(t) = \frac{1}{2}\mathbf{h} = \text{constant!!} \tag{29}$$

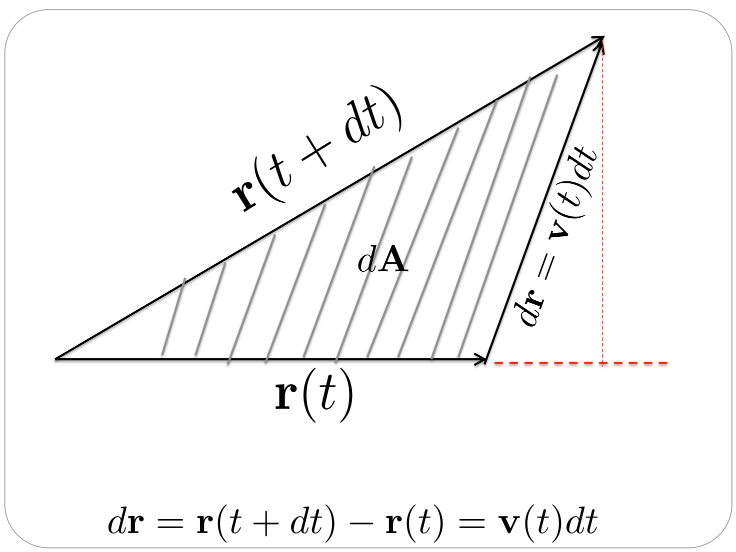


Figure 8: Rate of Area Sweep

Given that the rate of sweeping of the area is a constant, it is an obvious conclusion that equal areas will be swept in equal intervals of time - precisely the statement of Kepler's second law.

- **Flight path angle:** Recall that in the aircraft dynamics world, the flight path angle is a very important variable. It is the angle that the velocity vector makes with the horizon, defined positive *above* the horizon. Of course in the current context, the "horizon" is time-varying. We saw in Fig. (11) of the notes on particle mechanics (pg. 17) that the *local horizon plane* is defined by the radius vector $\mathbf{r}$, which is the normal to the plane. This makes complete sense because the line of sight along $\mathbf{r}$ looks to zenith, thereby the plane perpendicular to it must be the "horizontal plane". This is illustrated again in Fig. (9) below and the flight path angle (γ) is shown. Clearly, $\angle(\mathbf{r}, \mathbf{v}) = \pi/2 - \gamma$. The angular momentum (magnitude) can thus be given as:

$$h \;=\; \|\mathbf{r} \times \mathbf{v}\| \;=\; rv \sin\big(\angle(\mathbf{r}, \mathbf{v})\big) \;=\; rv \cos\gamma \tag{30}$$

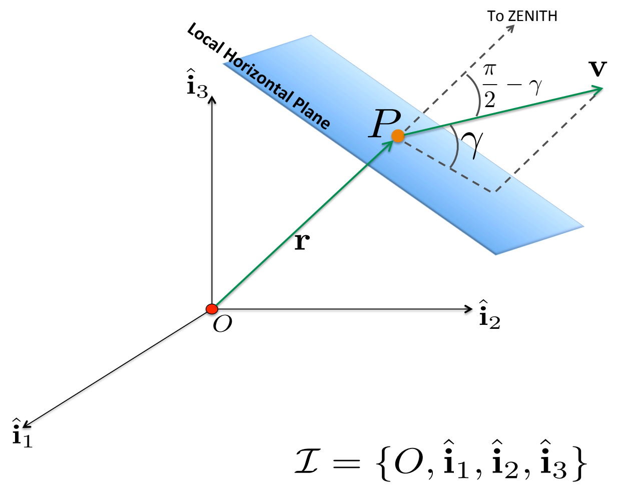


Figure 9: The Local Horizon and the Flight Path Angle

5.1.2 Eccentricity Vector

- The second integral of motion is the so-called "eccentricity vector", which will be used to prove Kepler's first law.
- For no good reason, consider the vector $\dot{\mathbf{r}} \times \mathbf{h}$. Since $\mathbf{h}$ is normal to the plane of motion, the vector $\dot{\mathbf{r}} \times \mathbf{h}$ lies in the said plane. Consider the inertial time derivative of this vector:

$$\frac{{}^{\mathcal{I}}d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) \;=\; \ddot{\mathbf{r}} \times \mathbf{h} \qquad \text{(since } \mathbf{h} \text{ is constant)} \tag{31a}$$

$$\underset{[18]}{=} \; -\frac{\mu}{r^3}\mathbf{r} \times \mathbf{h} \tag{31b}$$

$$\underset{[25]}{=} \; -\frac{\mu}{r^3}\mathbf{r} \times (\mathbf{r} \times \dot{\mathbf{r}}) \tag{31c}$$

$$= \; -\frac{\mu}{r^3}\left((\mathbf{r} \cdot \dot{\mathbf{r}})\,\mathbf{r} - (\mathbf{r} \cdot \mathbf{r})\,\dot{\mathbf{r}}\right) \tag{31d}$$

Eq.(31d) uses the vector triple cross product identity: $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \;=\; (\mathbf{a} \cdot \mathbf{c})\,\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\,\mathbf{c}$. Also, we have shown that Keplerian motion is planar. In other words, we can use the planar kinematics derived in the set of notes on particle mechanics:

$$\mathbf{r} \;=\; r\hat{\mathbf{e}}_r \tag{32a}$$

$$\dot{\mathbf{r}} \;=\; \dot{r}\hat{\mathbf{e}}_r + r\dot{\theta}\hat{\mathbf{e}}_\theta \tag{32b}$$

The dot products in Eq.(31d) are therefore given as: $\mathbf{r} \cdot \dot{\mathbf{r}} \;=\; r\dot{r}$, and $\mathbf{r} \cdot \mathbf{r} \;=\; r^2$. Thus continuing Eqs.(31), we have (after dropping superscript $\mathcal{I}$):

$$\frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) \;=\; -\frac{\mu}{r^3}\left(r\dot{r}\mathbf{r} - r^2\,\dot{\mathbf{r}}\right) \tag{33a}$$

$$= \mu\left(-\frac{\dot{r}}{r^2}\mathbf{r} + \frac{\dot{\mathbf{r}}}{r}\right) \tag{33b}$$

$$= \frac{d}{dt}\left(\frac{\mu}{r}\mathbf{r}\right) \tag{33c}$$

Thus,

$$\frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h} - \frac{\mu}{r}\mathbf{r}\right) = \mathbf{0} \tag{34}$$

And upon integration with respect to time,

$$\dot{\mathbf{r}} \times \mathbf{h} - \frac{\mu}{r}\mathbf{r} \;=\; \mathbf{c} \tag{35}$$

As shown above, the vector $\mathbf{c}$ is a constant, and we will define it as follows: $\mathbf{c} = \mu\mathbf{e} = \mu e\hat{\mathbf{i}}_e$.

- To gain insight into this newly found time invariant vector $\mathbf{c}$, consider the dot product: $\mathbf{r}\cdot\mathbf{c}$:

$$\mathbf{r}\cdot\mathbf{c} \;=\; \mathbf{r}\cdot\left(\dot{\mathbf{r}} \times \mathbf{h} - \frac{\mu}{r}\mathbf{r}\right) \tag{36a}$$

$$= \mathbf{r}\cdot(\dot{\mathbf{r}} \times \mathbf{h}) - \mu r \tag{36b}$$

$$= \mathbf{h}\cdot(\mathbf{r} \times \dot{\mathbf{r}}) - \mu r \tag{36c}$$

$$\underset{[25]}{=} h^2 - \mu r \tag{36d}$$

In going from Eq.(36b) to (36c), we have used the vector identity $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \mathbf{c}\cdot(\mathbf{a}\times\mathbf{b})$.

Now, let the angle between vectors $\mathbf{r}$ and $\mathbf{c}$ be f, i.e. $\angle(\mathbf{r},\mathbf{c}) = f$. Then, $\mathbf{r}\cdot\mathbf{c} = rc\cos f = r\mu e\cos f$. Combining this with Eq.(36d), we have:

$$r \;=\; \frac{h^2/\mu}{1 + e\cos f} \tag{37}$$

The above relationship represents the parametric equation of a conic section with "eccentricity" e. The equation is **universal**, in the sense that it could represent any one of the possible conic sections (i.e. a circle, ellipse, parabola or hyperbola), depending on the value of e:

e	**conic section**
$e = 0$	circle
$0 < e < 1$	ellipse
$e = 1$	parabola
$e > 1$	hyperbola

The quantity $h^2/\mu = p$ is called the *parameter of the conic section*, and is a constant. It is also known as the *semi-latus rectum* [Latin: *semi* - half, *latus* - side, *rectum* - straight]. There is an unfortunate mix-up of terminology here that is important to understand: even though $p = h^2/\mu$ is called the parameter of the conic section, <u>it is not due to p</u> that Eq.(37) is called a "parametric equation". Equation (37) is called parametric because it the representation of a conic section **in terms of a single <u>varying</u> parameter**, which is f. As f varies from 0 to 2π, you can sketch the complete graph of the conic section, for *fixed* values of e and p.

- We can ask the following question: Is Eq.(37) a solution of the fundamental equation of Keplerian motion (i.e. Eq.(18))? The answer is yes and no. Eq.(37) is only an implicit solution of Eq.(18), because while it gives us the shape of the orbit, it <u>does not</u> give us the explicit dependence of $\mathbf{r}$ on time, i.e. $\mathbf{r}(t)$. Instead, it gives us r as function of f, i.e. $r(f)$. Only when we obtain f as a function of time, i.e. $f(t)$ will we have a complete solution of [18], in the form of $r(f(t))$.

- The following conclusions can be drawn from our analysis of the eccentricity integral of motion:
  1. The relative paths in the two body problem are conic sections. This is a generalization of Kepler's first law which states that planets move around the sun in elliptical paths.
  2. $h^2 = \mu p$, where p is the *parameter* of the conic section and is a constant.
  3. The angle f is called the **true anomaly**. In order to get a physical feel for it, let us determine the value of f where r is minimum.

     Looking at Eq.(37), it is not hard to see that r reaches its minimum value at $f = 0$. In other words, the *point of closest approach, where* $r = r_{\text{min}}$ *is located at* $f = 0$.

     So, at $r = r_{\text{min}}$, $0 = f = \angle(\mathbf{r}_{\text{min}}, \mathbf{c})$. Therefore, $\mathbf{r}_{\text{min}} \parallel \mathbf{c}$, i.e. $\mathbf{c}$ is parallel to $\mathbf{r}_{\text{min}}$.

- We have just shown that the constant vector $\mathbf{c}$ points in the direction of the *periapsis*, which is the general term used for the point of closest approach on an orbit. If the central object is the Earth, the periapsis is called a perigee; if it is the Sun, we call it a perihelion, perilune if moon etc. Similarly, the farthest point on the orbit is called *apoapsis* in general (apogee for Earth, apohelion for Sun etc.). The line joining the apoapsis to the periapsis is called the *line of apses* or *line of apsides*.

- The above also indicates that the true anomaly is measured from the line of apses: see Fig.(10). From Eq.(37) note that $r(f = \pi/2) = h^2/\mu = p$. This is also shown in Fig.(10)

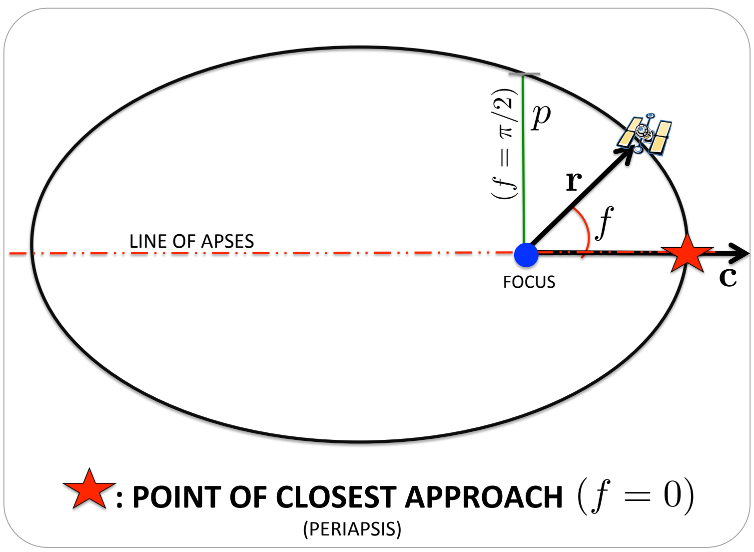


Figure 10: The Eccentricity Vector Integral of Motion

5.1.3 The Energy Integral

- There is one last integral of motion we must consider: the energy integral. As for the linear spring system above (Eqs.(19)), consider the "power" like entity $\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}}$:

$$\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} \underset{[18]}{=} -\frac{\mu}{r^3}\mathbf{r} \cdot \dot{\mathbf{r}} \tag{38a}$$

$$\underset{[32]}{=} -\frac{\mu}{r^2}\dot{r} \tag{38b}$$

$$= \frac{d}{dt}\left(\frac{\mu}{r}\right) \tag{38c}$$

To finish up, note that LHS $= \ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} = \frac{d}{dt}\left(\frac{1}{2}\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}\right)$. In conjunction with Eq.(38c), we get:

$$\frac{d}{dt}\left(\frac{v^2}{2} - \frac{\mu}{r}\right) = 0 \tag{39}$$

$$\text{or,} \quad \frac{v^2}{2} - \frac{\mu}{r} = \varepsilon \tag{40}$$

Eq.(40) is called the *vis-viva equation*, where ε is the total (constant) energy of the two-body system, comprising of kinetic energy $(v^2/2)$ and potential energy $(-\mu/r)$.

- In order to finish, we must determine the value of ε.

Note that ε is a constant. So, one can evaluate it at a point where it is most convenient to find v and r. This point is the periapsis. Before we proceed, note an important geometric relationship between the parameter p and semi-major axis a:

$$p = a(1 - e^2) \tag{41}$$

The above relationship will be used very frequently.

At the periapsis, we have $f = 0$, and let $r = r_p$ and $v = v_p$. From Eq.(37), we see that

$$r_p = \frac{p}{1+e} \underset{[41]}{=} a(1-e) \tag{42}$$

From Eq.(32b) $v_p^2 = \dot{r}_p^2 + r_p^2\dot{\theta}_p^2$. Since $r_p = r_{\min}$, $\dot{r}_p = 0$ (necessary condition for a minimum is that its first derivative is zero). This leaves us with $v_p^2 = r_p^2\dot{\theta}_p^2$.
Next, turn to the angular momentum vector. Using Eqs.(25) and (32) together, we get that $h^2 = r^4\dot{\theta}^2 = r_p^4\dot{\theta}_p^2$. Combining this with the expression for v_p^2, we get:

$$v_p^2 = r_p^2\dot{\theta}_p^2 = \frac{h^2}{r_p^2} = \frac{\mu p}{r_p^2} \tag{43}$$

Finally, using Eqs.(42) and (41):

$$v_p^2 = \frac{\mu p}{a^2(1-e)^2} = \frac{\mu(1+e)}{a(1-e)} \tag{44}$$

Therefore, we have, plugging Eqs.(42) and (44) into Eq.(40),

$$\varepsilon = \frac{v_p^2}{2} - \frac{\mu}{r_p} = \mu\left[\frac{(1+e)}{2a(1-e)} - \frac{1}{a(1-e)}\right] \tag{45a}$$

$$= -\frac{\mu}{2a} \quad \text{[Total Energy]} \tag{45b}$$

The final form of the vis-viva equation is:

$$\frac{v^2}{2} - \frac{\mu}{r} = -\frac{\mu}{2a} \tag{46}$$

- We can use the vis-viva equation to find the speed at the apoapsis (only applies to elliptic orbits!). Here, $f = \pi$ and let $r = r_a$. From Eq.(37), $r_a = p/(1-e) = a(1+e)$. Plug this is in Eq.(46) to get:

$$v_a^2 = 2\left(\frac{\mu}{a(1+e)} - \frac{\mu}{2a}\right) = \frac{\mu(1-e)}{a(1+e)} \tag{47}$$

- Note that the total energy is negative for an ellipse ($a > 0$) and positive for a hyperbola ($a < 0$). Therefore, as the size of the ellipse increases (a increases), its energy goes up.
- For **fixed angular momentum**, the semi-major axis increases with eccentricity, and so does the energy. It makes sense to refer to the semi-major axis as a measure of the size of the orbit. So we conclude that as the size of the orbit increases (for a fixed $\mathbf{h}$), so does its energy. This justifies the "modern day generalization of Kepler's first law" (pg.1). Table (1) contains the complete details for all conics.

conic	e	a (from geometry)	ε (RHS of Eq.(46))	comment
circle	$e = 0$	+ve	-ve	
ellipse	$0 < e < 1$	+ve	-ve	$\varepsilon_{\text{ell}} > \varepsilon_{\text{circ}}$
parabola	$e = 1$	∞	0	$\varepsilon_{\text{parab}} > \varepsilon_{\text{ell}} > \varepsilon_{\text{circ}}$
hyperbola	$1 < e < \infty$	-ve	+ve	$\varepsilon_{\text{hyp}} > \varepsilon_{\text{parab}} > \varepsilon_{\text{ell}} > \varepsilon_{\text{circ}}$

Table 1: Various conic sections (**fixed h**)

Fig.(11) shows the continuous variation of the semi-major axis and energy with eccentricity. In this figure, the angular momentum is held constant (with magnitude $h = 2.0 \times 10^{11}\ m^2/s$). Note that $e = 1$ (parabola) is the transition point between open and closed orbits, both in terms of the transition of a (positive to negative) and ε (negative to positive).

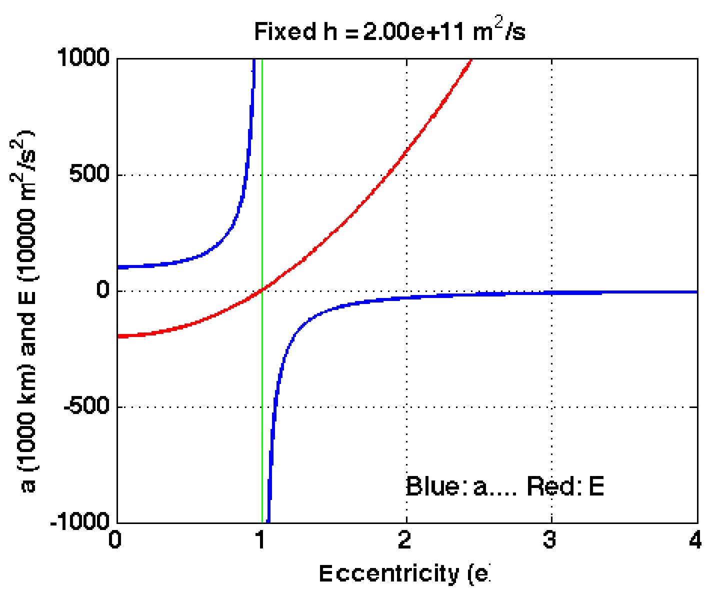


Figure 11: Variation of a and ε with e. Fixed $\mathbf{h}$ $\forall\ e$

5.2 Special Case: Speed on a Circular Orbit

- For a circular orbit, $e = 0$. Substituting this in the universal parametric equation of conic sections (Eq.(37)), we see that $r = p = a$, i.e. a confirmation of the fact that the radius in a circular orbit is constant. The vis-viva equation for a circle is:

$$\frac{v^2}{2} - \frac{\mu}{r} = -\frac{\mu}{2r} \tag{48}$$

or,

$$v_c^2 = \frac{\mu}{r} \tag{49}$$

the subscript c stands for circle.

❣ **Example** Consider Fig.(12): a spacecraft is shown at point A, through which two orbits are drawn - one is elliptical and the other circular. The distance of point A from the central body (say, Earth) is r. We would like to compute the speed of the spacecraft at point A on the elliptical and circular orbits.

Let the semimajor axis of the ellipse be $a = 12000\ km$. Also let $r = 0.8a = 9600\ km$. From Eq.(46), the speed of the spacecraft at point A in the two cases is:

$$\begin{aligned} v_e^2 = 2\left(\frac{\mu}{r} + \varepsilon_e\right) &= 2\left(\frac{\mu}{r} - \frac{\mu}{2a}\right) \\ &= 2\left(\frac{\mu}{.8a} - \frac{\mu}{2a}\right) = \frac{3\mu}{2a} \end{aligned} \tag{50a}$$

and,

$$v_c^2 \underset{[49]}{=} \frac{\mu}{.8a} = \frac{5\mu}{4a} \tag{50b}$$

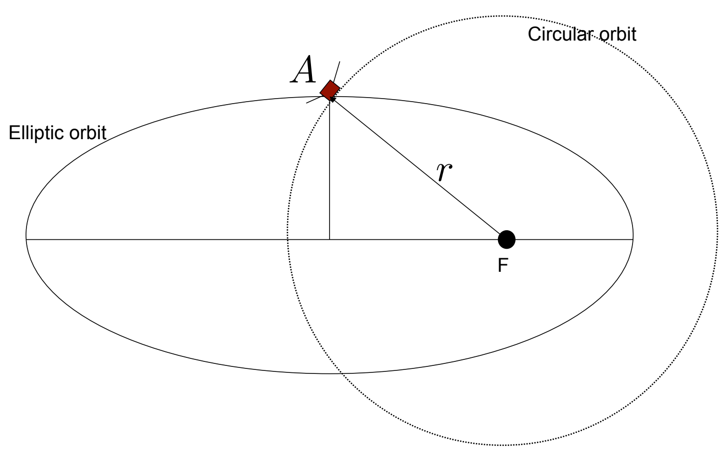


Figure 12: Velocity on a Circular Orbit

Therefore,

$$\frac{v_e^2(r)}{v_c^2(r)} = \frac{6}{5} = 1.2 \tag{51}$$

In other words, at point A the spacecraft is moving at a speed $\sqrt{1.2} = 1.095$ times faster on the shown elliptical orbit than the circular orbit. Substituting $a = 12,000\ km$, we get $v_e(r) = 7.0587\ kmps$ and $v_c(r) = 6.443\ kmps$.

Suppose the spacecraft was traveling in the elliptical orbit and it was required to *move it* from the elliptical orbit to the circular orbit at point A by making an *impulsive change* to its speed. The above analysis shows that we would need to *instantaneously slow down the spacecraft by* $\Delta v = v_e(r) - v_c(r) = 615.02\ m/s$. This is not a trivial change!

6 Escape Velocity

- In the above example, the spacecraft was slowed down to go from a high energy orbit (ellipse) to a low energy orbit (circle) sharing a common point A. We next consider the case of increasing the spacecraft energy. In fact, the objective is to increase the energy to the extent that the spacecraft moves from the closed elliptical orbit to an "open" orbit. An open orbit with the least energy is a parabola. This transfer will allow the spacecraft to "escape" the vicinity of Earth and never return.
- Note that as we transition from an ellipse to a parabola, the energy *increases* from some negative value all the way up to zero (see Table (1)). Consider a spacecraft (or any other object) at some point A on an elliptical orbit, as shown in Fig.(13). *If* the spacecraft was on a parabolic orbit at this point, its speed would be given by the following equation:

$$\frac{v_p^2}{2} - \frac{\mu}{r} = \varepsilon_p = 0 \tag{52}$$

i.e., $v_p(r) = \sqrt{2\frac{\mu}{r}} = \sqrt{2}v_c(r)$. This speed, i.e. $v_p(r)$ is called the *escape velocity*, because it is *the minimum speed needed at A to be on an open orbit around the central body.* Anything less and you would be on an elliptical orbit which would cause the spacecraft to return to its starting point, since all ellipses are closed no matter how elongated they may be.

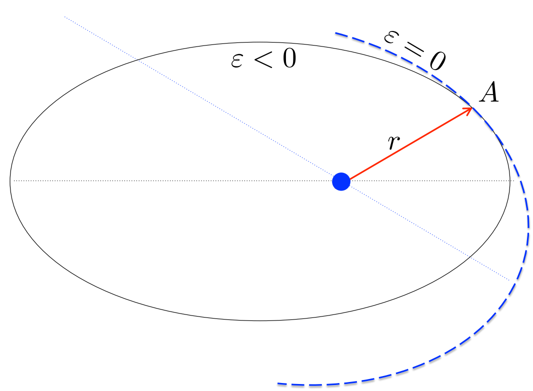


Figure 13: Escape Velocity

Therefore, the impulsive change in speed needed to move from the elliptical orbit shown to a parabolic orbit is $\Delta v = v_p(r) - v_e(r)$. Using the numbers from the previous example, we get $\Delta v = 2.054\ kmps$, which is massive!

Note that the escape velocity *from surface of the Earth is* $v_p(R_\oplus) = \sqrt{2\frac{\mu_\oplus}{R_\oplus}} = 11.18\ kmps$, where $R_\oplus$ is the radius of the Earth. In other words, if you where lying on the surface of the Earth, and you wanted to be on an "escape orbit," then you would need to have a speed of at least 11.18 $kmps$.

7 Orbit Period

- Of course, a discussion on orbital period makes sense only for closed orbits. The "period" of a parabolic or hyperbolic orbit is infinite. Kepler stated his third law ($P^2 \propto a^3$) only for elliptical orbits because he had the planets in mind. We will now derive this relationship by revisiting angular momentum.

- Recall Eq.(29), which states that the rate of sweep of area is a constant, equal to half the angular momentum vector. Fig.(14) shows the area swept by a spacecraft between times t_0 and t_f. Let the orbital period be P. If $t_f = (t_0 + P)$ then the area swept must equal the total area of the ellipse! Let us integrate Eq.(29) from t_0 to $t_f = (t_0 + P)$, such that

$$\begin{aligned} \mathbf{A}_{\text{ell}} &= \int_{t_0}^{t_0+P} \frac{\mathbf{h}}{2} dt \\ &= \frac{P\mathbf{h}}{2} \end{aligned} \tag{53}$$

where, $\mathbf{A}_{\text{ell}}$ is the total area of the ellipse. Also, we have exploited the time-invariance of the angular momentum vector in pulling it out of the integral. We can drop the "bold" and look at only the magnitudes, such that $A_{\text{ell}} = Ph/2$. Now, from geometry of ellipses, $A_{\text{ell}} = \pi ab$, where a and b are the semi- major and minor axes of the ellipse respectively. So we have, $\pi ab = Ph/2$ or,

$$\pi^2 a^2 b^2 = \frac{P^2 h^2}{4} \tag{54}$$

To finish, we use the following two relationships: (a) $h^2 = \mu p = \mu a(1 - e^2)$, and (b) $b = a\sqrt{1 - e^2}$. Thus,

$$P^2 = \frac{4\pi^2 a^3}{\mu} \tag{55}$$

The above equation is nothing but the statement of Kepler's third law and also includes the constant of proportionality. Clearly, the period of an elliptical orbit depends only on its size (a). Eq.(55) also holds for circular orbits, in which case $a = r$.

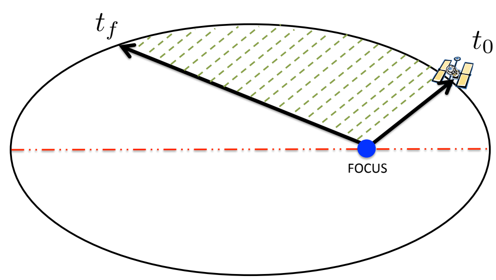
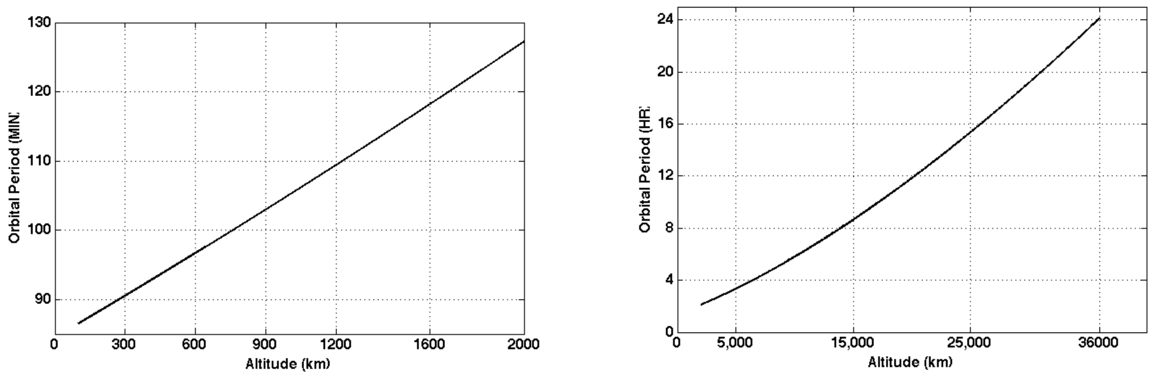


Figure 14: Kepler's Third Law: Orbital Period

- Mean motion is defined as the average angular rate around the orbit. For all orbits except circular, the instantaneous angular rate is **not** a constant. We know that by virtue of Kepler's second law, an object moves faster when it is closer to the center of force. Mean motion is simply an average angular rate, given as the total (angular) distance traveled divided by the total time over one period. I.E.,

$$n = \frac{2\pi}{P} \underset{\text{Eq.[55]}}{=} \sqrt{\frac{\mu}{a^3}} \tag{56}$$

- Figs.(15) show a plot of periods of circular orbits of various altitudes.



(a) Period of Circular Orbits (alt $< 2000\ km$)

(b) Period of Circular Orbits ($2000 km <$ alt $<$ GEO)

Figure 15: Period of Circular Orbits (upto GEO altitude)

8 Acceleration Due to Gravity

- From Newton's second law and Newton's universal gravitation law,

$$m\ddot{\mathbf{r}} = m \underbrace{\left(-\frac{GM}{r^3}\mathbf{r} \right)}_{\mathbf{g}}$$

We traditionally denote acceleration due to gravity by $\mathbf{g}$, as shown above. Taking its magnitude,

$$\|\mathbf{g}\| = g = \frac{GM}{r^2} = \frac{\mu}{r^2} \tag{57}$$

- Clearly, $g = g(r)$. It is inversely proportional to the square of the distance from the center of force. Thereby, it diminishes as r increases. For Earth, $\mu_\oplus = 3.986004 \times 10^{14}\ m^3/s^2$. On Earth's surface, $r = R_\oplus = 6378.14 \times 10^3\ m$. So the acceleration due to gravity on Earth's surface is $g(R_\oplus) = 3.986004 \times 10^{14}/(6378.14 \times 10^3)^2 = 9.8\ m/s^2$. Rigorously, this is valid only on Earth's surface.

9 Geometry of Conic Sections

We will perform a short review of conic sections here. The general definition is the following: *A conic section is a locus of points whose distance (r) from a fixed point (F) has a constant ratio (e) with the perpendicular distance (d) from a fixed line (directrix).* This is illustrated in Fig.(16).

- It turns out that for different values of the parameter e, the curves defined by the above locus represent various sections of a cone (Fig.(17)). We have already identified e as the *eccentricity* of the curve and Table (1) lists the different cases.

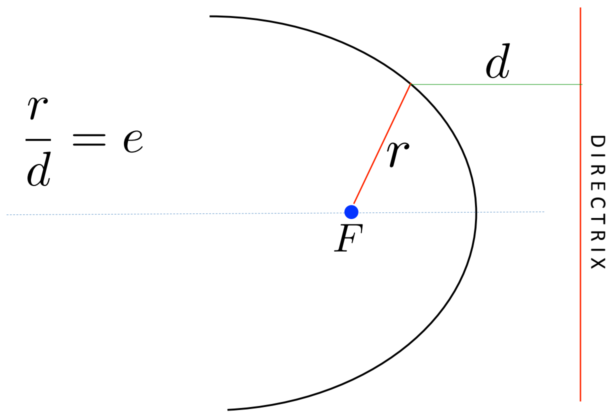


Figure 16: Locus definition of Conic Sections

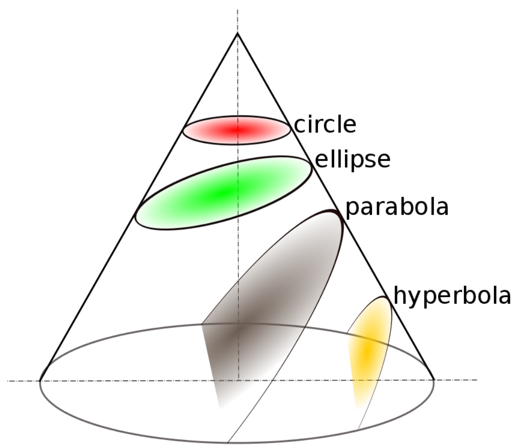


Figure 17: Sections of a Cone

- Table (2) provides useful geometrical relationships for the various conic sections.
- In the above table, "perifocal distance" is defined as the distance between the focus and the periapsis (written as r_p previously). See ellipse and hyperbola figures below for illustration.
- Fig.(18) shows the geometry of an ellipse. Similarly, Fig.(28) shows the geometry of a hyperbola. We will delve deeper into conic section geometry a little later in the course.
- The following relationships are *universal*, i.e. hold for all types of conic sections:

$$h^2 = \mu p \tag{58}$$

$$r = \frac{p}{1 + e\cos f} \tag{59}$$

$$x = r\cos f \tag{60}$$

$$y = r\sin f \tag{61}$$

$$\frac{v^2}{2} - \frac{\mu}{r} = -\frac{\mu}{2a} \tag{62}$$

Quantity	Circle	Ellipse	Parabola	Hyperbola
Defining parameters	a = semi-major axis = radius	a = semi-major axis b = semi-minor axis	p = semi-latus rectum q = perifocal distance	a = semi-transverse axis a = semi-conjugate axis By convention: $a < 0$
Equation	$x^2 + y^2 = a^2$	$x^2/a^2 + y^2/b^2 = 1$	$x^2 = 4qy$	$x^2/a^2 - y^2/b^2 = a^2$
Eccentricity, e	$e = 0$	$e = \sqrt{a^2 - b^2}/a$ $0 < e < 1$	$e = 1$	$e = \sqrt{a^2 + b^2}/(-a)$ $e > 1$
Perifocal distance, q	$q = a$	$q = a(1 - e)$	$q = p/2$	$q = a(1 - e)$
Velocity @ r: $v(r)$	$v^2 = \mu/r$	$v^2 = \mu(2/r - 1/a)$	$v^2 = 2\mu/r$	$v^2 = \mu(2/r - 1/a)$
Energy, ε	$\varepsilon = -\mu/2a$, $\varepsilon < 0$	$\varepsilon = -\mu/2a$, $\varepsilon < 0$	$\varepsilon = 0$	$\varepsilon = -\mu/2a$, $\varepsilon > 0$
Period, P	$P = 2\pi\sqrt{a^3/\mu}$	$P = 2\pi\sqrt{a^3/\mu}$	$P = \infty$	$P = \infty$
Mean angular motion, n	$n = \sqrt{\mu/a^3}$	$n = \sqrt{\mu/a^3}$	$n = \sqrt{\mu}$	$n = \sqrt{\mu/(-a)^3}$
True anomaly, f	$f = E = M$	$\tan\frac{f}{2} = \left(\frac{1+e}{1-e}\right)^{1/2}\tan\frac{E}{2}$ E = Eccentric anomaly	$\tan\frac{f}{2} = D/2q$ D = Parabolic anomaly	$\tan\frac{f}{2} = \left(\frac{1+e}{1-e}\right)^{1/2}\tanh\frac{F}{2}$ F = Hyperbolic anomaly
Distance from focus, r	$r = a$	$r = a(1 - e\cos E)$	$r = q + D^2/2$	$r = a(1 - e\cosh F)$
Parameter, p	$p = a$	$p = a(1 - e^2) = b^2/a$	$p = 2q$	$p = a(1 - e^2) = b^2/a$

Table 2: Geometrical relationships for various conic sections

where, x and y are the components of the spacecraft's position vector along and perpendicular to the line of apsides in the orbital reference frame, measured from F.

10 Orbital Elements

- Orbital elements are quantities that characterize the size, shape and orientation of the two-body orbit with respect to the inertial reference frame:
  - Size elements: a, p etc... — how big is the orbit?
  - Shape elements: e — how eccentric is the orbit?
  - Orientation elements: i, ω, Ω — what is the orientation of the orbital plane with respect to the inertial frame?
- We have already encountered the parameters a and e earlier. The quantities i, ω and Ω are actually the Euler angles of a $3-1-3$ sequence, used to ascertain the transformation from the inertial frame to the orbital reference frame. The angle Ω is called the *longitude of ascending node*, i is the *inclination* and ω is the *argument of periapsis*. We will soon get a visual feel.
- We must realize that we have already constructed an "orbital reference frame", $\mathcal{O}$ as follows

$$\mathcal{O} = \{F,\ \hat{\mathbf{i}}_e,\ \hat{\mathbf{i}}_y,\ \hat{\mathbf{i}}_h\} \tag{63}$$

  where, F is the origin of the reference frame, which is nothing but the focus of the orbit. For geocentric orbits, this point is the center of the Earth, written thus far as O. The unit vector $\hat{\mathbf{i}}_e$ points in the direction of the periapsis, i.e. is the unit eccentricity vector. The unit vector $\hat{\mathbf{i}}_h$ is the normal to the plane of the orbit and $\hat{\mathbf{i}}_y$ completes the triad, i.e. $\hat{\mathbf{i}}_y = \hat{\mathbf{i}}_h \times \hat{\mathbf{i}}_e$ (see Fig.(20)).
- **Important Fact:** The complete state information at any one point in time is enough to determine all orbital elements. Assume that at some time t_0, the inertial position and velocity vectors are known, $\mathbf{r}_\mathcal{I}(t = t_0) = \mathbf{r}_\mathcal{I}(0)$ and $\mathbf{v}_\mathcal{I}(t = t_0) = \mathbf{v}_\mathcal{I}(0)$. Next, we will use this information to determine all the orbital elements.

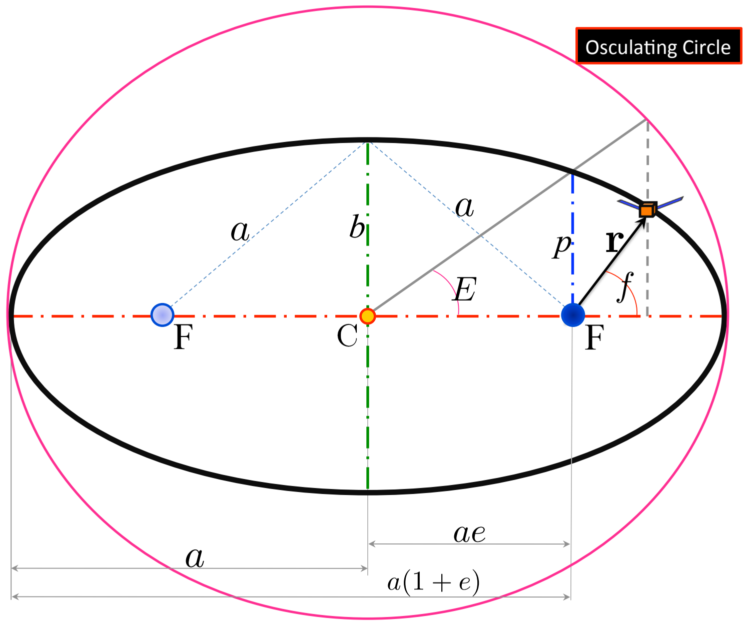


Figure 18: Geometry of an Ellipse

10.1 Evaluation of orbital elements from $(\mathbf{r}, \mathbf{v})$ data

- Semi-major axis, a: It is the easiest to compute a from the vis-viva equation. Simply use it at $t = t_0$ to get:

$$\frac{1}{a} = \frac{2}{r_0} - \frac{v_0^2}{\mu} \tag{64}$$

where $r_0 = \|\mathbf{r}_{\mathcal{I}}(0)\|$ and $v_0 = \|\mathbf{v}_{\mathcal{I}}(0)\|$.

- Eccentricity, e: The direct way to compute eccentricity is through the eccentricity vector, $\mu e \hat{\mathbf{i}}_e = \dot{\mathbf{r}} \times \mathbf{h} - \frac{\mu}{r}\mathbf{r}$ evaluated at t_0 such that

$$e = \frac{\|\mathbf{v}_{\mathcal{I}}(0) \times \mathbf{h}_{\mathcal{I}} - \frac{\mu}{r_0}\mathbf{r}_{\mathcal{I}}(0)\|}{\mu} \tag{65a}$$

$$= \frac{\|\mathbf{v}_{\mathcal{I}}(0) \times (\mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0)) - \frac{\mu}{r_0}\mathbf{r}_{\mathcal{I}}(0)\|}{\mu} \tag{65b}$$

Note that $\mathbf{h} = \mathbf{r} \times \mathbf{v} =$ constant has been used in going from Eq.(65a) to (65b). Also, the above development is 'universal' in the sense that it applies to all types of conic sections.

While the above approach is direct, it requires the evaluation of two cross products and a complicated looking numerator, which may not be desirable. If the semi-major axis has already been found, there exists a shorter route, using the (also universal) relationship between h, p and a. Recall:

$$\frac{h^2}{\mu} = p = a(1 - e^2) \tag{66}$$

Therefore,

$$e = \sqrt{\left(1 - \frac{h^2}{\mu a}\right)} \tag{67a}$$

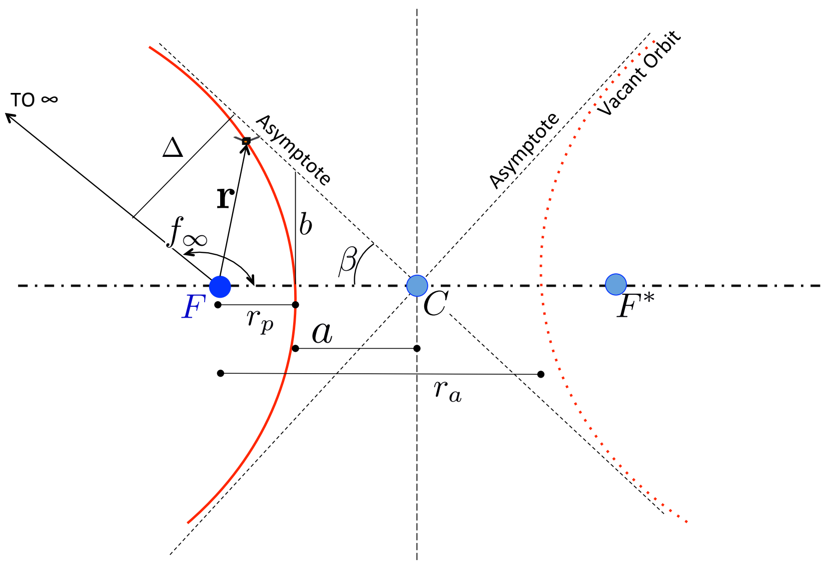


Figure 19: Geometry of a Hyperbola

$$= \sqrt{\left(1 - \frac{\|\mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0)\|^2}{\mu a}\right)} \tag{67b}$$

- Orientation angles (Ω, i and ω): The size and shape of the orbit contains information only about the characteristics on the plane of motion. All that remains is to determine its orientation with respect to the inertial reference frame.

  In other words, we must determine the transformation between the orbital reference frame of Eq.(63) and the inertial reference frame. Recall from Eq.12 (or equivalently, Eq.(16)) of the notes on particle mechanics, for two frames $\mathcal{A}$ and $\mathcal{B}$,

$$\mathbf{R}_{\mathcal{AB}} = \begin{Bmatrix} \hat{\mathbf{a}}_{1,\mathcal{B}}^{\mathrm{T}} \\ \hat{\mathbf{a}}_{2,\mathcal{B}}^{\mathrm{T}} \\ \hat{\mathbf{a}}_{3,\mathcal{B}}^{\mathrm{T}} \end{Bmatrix} \tag{68}$$

  where, the superscript "T" stands for transpose. Therefore, between frames $\mathcal{I}$ and $\mathcal{O}$, we have:

$$\mathbf{R}_{\mathcal{OI}} = \begin{Bmatrix} \hat{\mathbf{o}}_1^{\mathrm{T}} \\ \hat{\mathbf{o}}_2^{\mathrm{T}} \\ \hat{\mathbf{o}}_3^{\mathrm{T}} \end{Bmatrix} = \begin{Bmatrix} \hat{\mathbf{i}}_{e,\mathcal{I}}^{\mathrm{T}} \\ \hat{\mathbf{i}}_{y,\mathcal{I}}^{\mathrm{T}} \\ \hat{\mathbf{i}}_{h,\mathcal{I}}^{\mathrm{T}} \end{Bmatrix} \tag{69}$$

  Clearly, the above transformation is known, since

$$\hat{\mathbf{i}}_{e,\mathcal{I}} = \frac{\mathbf{v}_{\mathcal{I}}(t_0) \times \mathbf{h}_{\mathcal{I}} - \mu \mathbf{r}_{\mathcal{I}}(t_0)/r_0}{\mu e} \tag{70a}$$

$$= \frac{\mathbf{v}_{\mathcal{I}}(0) \times (\mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0)) - \mu \mathbf{r}_{\mathcal{I}}(0)/r_0}{\mu e}, \tag{70b}$$

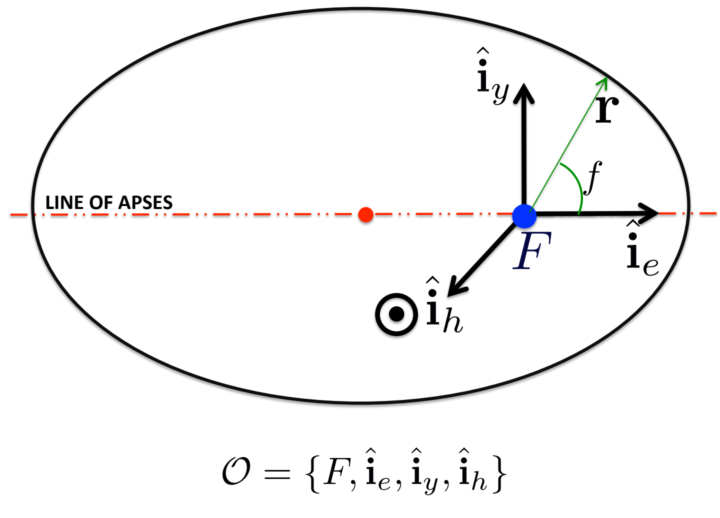


Figure 20: The Orbital Reference Frame

$$\hat{\mathbf{i}}_{h,\mathcal{I}} = \frac{\mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0)}{\|\mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0)\|} \quad \text{and,} \tag{70c}$$

$$\hat{\mathbf{i}}_{y,\mathcal{I}} = \frac{\hat{\mathbf{i}}_{h,\mathcal{I}} \times \hat{\mathbf{i}}_{e,\mathcal{I}}}{\|\hat{\mathbf{i}}_{h,\mathcal{I}} \times \hat{\mathbf{i}}_{e,\mathcal{I}}\|} \tag{70d}$$

It turns out that there exists a highly physically intuitive Euler angle sequence that accomplishes the tranformation shown in Eqs.(69) and (70). This is a $3-1-3$ sequence, given as:

$$\underset{\{F,\hat{\mathbf{i}}_j\}}{\mathcal{I}} \underset{\dot{\Omega}}{\xrightarrow{\mathbf{R}_3(\Omega)}} \underset{\{F,\hat{\mathbf{i}}'_j\}}{\mathcal{I}'} \underset{\dot{i}}{\xrightarrow{\mathbf{R}_1(i)}} \underset{\{F,\hat{\mathbf{i}}''_j\}}{\mathcal{I}''} \underset{\dot{\omega}}{\xrightarrow{\mathbf{R}_3(\omega)}} \underset{\{F,\hat{\mathbf{i}}_e,\hat{\mathbf{i}}_y,\hat{\mathbf{i}}_h\}}{\mathcal{O}} \tag{71}$$

The three rotations are illustrated in Fig.(21). The first rotation is by an angle Ω about the z-axis (Fig.(21).A). Ω is called the *longitude of ascending node.*

The second transformation is a x-rotation by the angle i, which is nothing but the inclination of the orbit. To see the orbit "inclination" with greater clarity, a different view of the $\mathbf{R}_1(i)$ transformation is shown in the top right side of Fig.(21).B, illustrating the inclination of the orbit with respect to the Earth's equatorial plane. Also from this view, it is clear that the axis $\hat{\mathbf{i}}''_3$ is normal to the plane of the orbit, and must therefore be the same as $\hat{\mathbf{i}}_h$.

Since the desired z-axis has already been achieved ($\hat{\mathbf{i}}_h$), we perform the third and final rotation by angle ω about $\hat{\mathbf{i}}''_3$, resulting in the orbital reference frame. The angle ω is called the *argument of periapsis.* Recall that $\hat{\mathbf{i}}_e$ points in the direction of the periapsis (P). Therefore, the angle ω derives its name from the fact that the rotation $\mathbf{R}_3(\omega)$ helps us locate the periapsis, P.

Also shown in Fig.((21).C) is the point A, which is one of the two points of intersection between the orbital plane and equatorial plane. For "regular" orbits (in terms of direction of motion), this is where

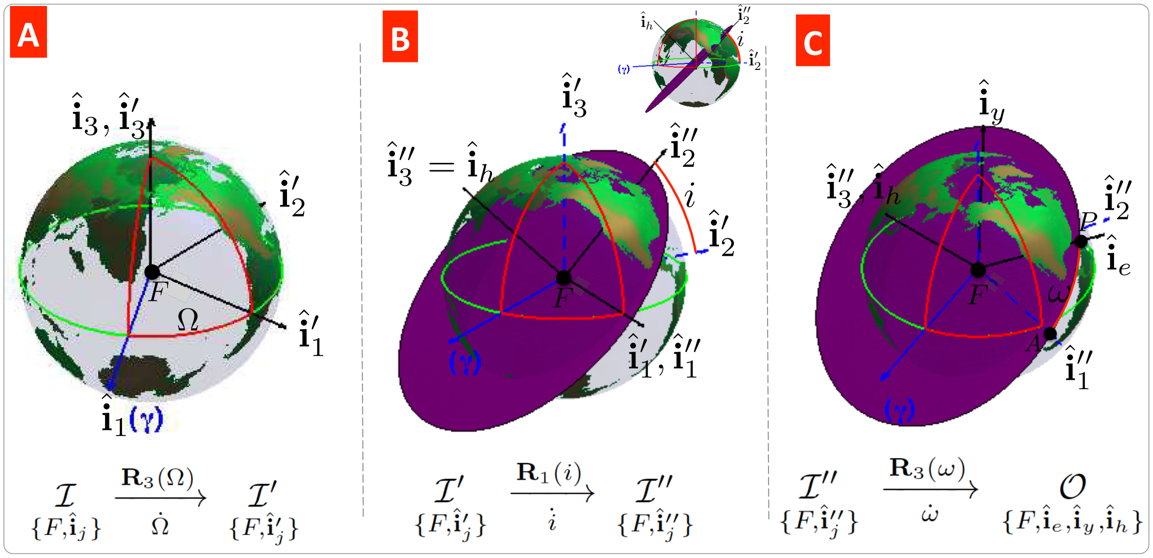


Figure 21: Going from the Inertial Reference Frame to the Orbital Reference Frame

the spacecraft "rises above the equator" and is therefore called the *ascending node*. The name of Ω (longtiude of ascending node) now makes sense because it helps locate the ascending node with respect to the inertial x-axis. The antipode of A is the descending node, where the spacecraft "sets below the equator".

The final question is, how do we determine the Euler angles Ω, i and ω? This is a simple question to answer because we already know the transformation $\mathbf{R}_{\mathcal{OI}}$ from Eq.(69). Also, from Eq.(71), we know that $\mathbf{R}_{\mathcal{OI}} = \mathbf{R}_3(\omega)\mathbf{R}_1(i)\mathbf{R}_3(\Omega)$. Therefore, we use the standard technique the extract the Euler angles, starting with i:

$$\mathbf{R}_{\mathcal{OI}} = \mathbf{R}_3(\omega)\mathbf{R}_1(i)\mathbf{R}_3(\Omega) \quad = \begin{bmatrix} (c\omega c\Omega - s\omega c i s\Omega) & (c\omega s\Omega + s\omega c i c\Omega) & s\omega s i \\ -(s\omega c\Omega + c\omega c i s\Omega) & (c\omega c i c\Omega - s\omega s\Omega) & c\omega s i \\ s i s\Omega & -s i c\Omega & c i \end{bmatrix} \tag{72a}$$

$$\underset{\text{Eq.(69)}}{=} \begin{bmatrix} i_{e_1} & i_{e_2} & i_{e_3} \\ i_{y_1} & i_{y_2} & i_{y_3} \\ i_{h_1} & i_{h_2} & i_{h_3} \end{bmatrix} \tag{72b}$$

where, we have used $\hat{\mathbf{i}}_{e,\mathcal{I}} = \{i_{e_1}, i_{e_2}, i_{e_3}\}^{\mathrm{T}}$, etc. Comparing Eqs.(72a) and (72b), we get

$$i \;=\; \cos^{-1}(i_{h_3}) \tag{73a}$$

$$\Omega \;\mapsto \begin{cases} c\Omega \;=\; -i_{h_2}/si \\ s\Omega \;=\; i_{h_1}/si \end{cases} \tag{73b}$$

$$\omega \;\mapsto \begin{cases} c\omega \;=\; i_{y_3}/si \\ s\omega \;=\; i_{e_3}/si \end{cases} \tag{73c}$$

Keep in mind that you must use *both* $\cos\Omega$ and $\sin\Omega$ expressions shown above to obtain the longitude of ascending node in the correct quadrant. Same applies to the argument of periapsis (ω).

To conclude our discussion here, we consider the following question: **How to obtain f_0**? So far, we have obtained all possible detail about the orbit - its size, shape and orientation with respect to the inertial reference frame. The only thing missing is the location of the spacecraft with respect to the periapsis at t_0, i.e. $f(t_0) = f_0$. Recall that if you choose to employ the "equation of orbit" $[r = p/(1 + e\cos f)]$, there is potential for ambiguity because $\cos f = \cos(2\pi - f)$. In other words, it is impossible to resolve which of the two cases shown in Fig.(22) is the actual location of the spacecraft at $t = t_0$.

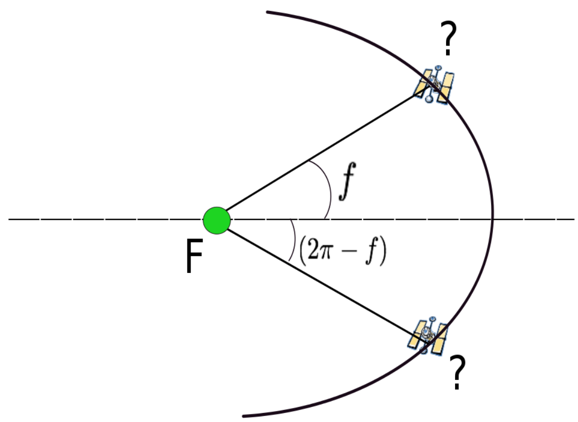


Figure 22: Ambiguity in determination of true anomaly from the equation of orbit

Given $\mathbf{r}_{\mathcal{I}}(t_0)$ and $\mathbf{v}_{\mathcal{I}}(t_0)$, one way to determine $f(t_0)$ is the following:

1. Using the rotation matrix between frames $\mathcal{I}$ and $\mathcal{O}$, determine $\mathbf{r}_{\mathcal{O}}(t_0)$:
$$\mathbf{r}_{\mathcal{O}}(0) \;=\; \mathbf{R}_{\mathcal{O}\mathcal{I}}\mathbf{r}_{\mathcal{I}}(0) \;\doteq\; [r_{0\mathcal{O}1},\; r_{0\mathcal{O}2},\; r_{0\mathcal{O}3}]'. \tag{74}$$

2. From Fig.(20), it is evident that
$$\mathbf{r}_{\mathcal{O}}(0) \;=\; r_0\cos f_0\;\hat{\mathbf{i}}_{e,\mathcal{I}} \;+\; r_0\sin f_0\;\hat{\mathbf{i}}_{y,\mathcal{I}} \;+\; 0\;\hat{\mathbf{i}}_{h,\mathcal{I}} \tag{75}$$
where $r_0 = \|\mathbf{r}_{\mathcal{I}}(t_0)\|$ is known. Sanity check: $r_{0\mathcal{O}3}$ must equal 0 in step 1 above, since there is no component of the radius vector outside the plane of the orbit, i.e. along $\hat{\mathbf{i}}_h$.

3. Compare steps (1) and (2) to see that
$$f_0 \;\mapsto \begin{cases} \cos f_0 \;=\; r_{0\mathcal{O}1}/r_0 \\ \sin f_0 \;=\; r_{0\mathcal{O}2}/r_0 \end{cases} \tag{76}$$

Alternatively, you may use the following **Method 2** for determining f_0:

1. Use the equation of orbit (Eq.(37)) to obtain $\cos f_0$. This alone cannot determine the quadrant of f_0 since $\cos f_0 = \cos(2\pi - f_0)$.

2. Given $\mathbf{v}_{\mathcal{I}}(0)$, obtain $v_{0\mathcal{O}1}$, which is the x-component of the vector $\mathbf{v}_{\mathcal{O}}(0)$. Just like $\mathbf{r}_{\mathcal{O}}(0)$, $\mathbf{v}_{\mathcal{O}}(0) = \mathbf{R}_{\mathcal{O}\mathcal{I}}\mathbf{v}_{\mathcal{I}}(0)$.

3. If $v_{0\mathcal{O}1} > 0$ (positive), $f_0 \in [\pi, 2\pi]$. If on the other hand if $v_{0\mathcal{O}1} < 0$ (negative), $f \in [0, \pi]$, i.e. the first two quadrants.

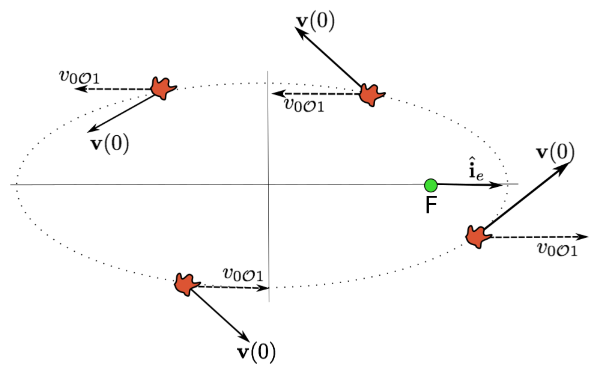


Figure 23: Determination of initial true anomaly using $v_{0\mathcal{O}1}$

The reasoning behind this method is the following: keep in mind that $v_{0\mathcal{O}1}$ is the component of $\mathbf{v}(0)$ along the $\hat{\mathbf{i}}_e$ axis. Then, look at Fig.(23). Clearly, the $\hat{\mathbf{i}}_e$-component of $\mathbf{v}(0)$ is negative if the object is in the first two quadrants and positive it it is in the third or fourth quadrants. Type of conic is irrelevant!

Interestingly, you may use a **Method 3** to find f_0: Note that in the first two quadrants, the magnitude of the radius vector (i.e. r) increases, while in quadrants 3 and 4, it decreases. In terms of derivatives, $\dot{r} > 0$ if $f \in [0, \pi]$ and $\dot{r} > 0$ if $f \in [\pi, 2\pi]$. That you can use particle planar kinematics equations to get $\dot{r}$ as follows: $\mathbf{r} = r\hat{\mathbf{e}}_r$ and $\mathbf{v} = \dot{r}\hat{\mathbf{e}}_r + r\dot{f}\hat{\mathbf{e}}_f$. Therefore, $\mathbf{r} \cdot \mathbf{v} = r\dot{r}$. Since $r > 0$, therefore $\text{sign}(\dot{r}) = \text{sign}(\mathbf{r} \cdot \mathbf{v})$. So we have:

1. Given $\mathbf{r}_{\mathcal{I}}(0)$ and $\mathbf{v}_{\mathcal{I}}(0)$, determine $\mathbf{r}_{\mathcal{I}}(0) \cdot \mathbf{v}_{\mathcal{I}}(0)$.
2. If $\text{sign}[\mathbf{r}_{\mathcal{I}}(0) \cdot \mathbf{v}_{\mathcal{I}}(0)] = +1$, $f \in [0, \pi]$, else $f \in [\pi, 2\pi]$.

❣ **Example Orbital Elements.** Consider the following inertial radius and velocity vectors of a spacecraft *measured* at $t = 0$:

$$\mathbf{r}_{\mathcal{I}}(0) = \{-4777.8, 4862.6; 1760.1\}^{\mathrm{T}} \ \ km \tag{77a}$$

$$\mathbf{v}_{\mathcal{I}}^{\mathcal{I}}(0) = \{-6.7782, -4.8929, 0.9174\}^{\mathrm{T}} \ \ km/s \tag{77b}$$

Determine all the orbital elements and plot the orbit in 3 dimensions. Also obtain r_p, r_a, v_p and v_a. Given $\mu = \mu_{\oplus} = 3.986004 \times 10^{14} \ m^3/s^2$.

SOLUTION: At this point, we do not know the type of orbit the spacecraft is in. Note that the given velocity vector has been measured *and* expressed in the inertial reference frame. The magnitudes of the position and velocity vectors are : $r_0 = \|\mathbf{r}_{\mathcal{I}}\| = 7040.6 \ km$ and $v_0 = \|\mathbf{v}_{\mathcal{I}}^{\mathcal{I}}\| = 8.4099 \ km/s$. We can use these in the vis-viva equation to find the semi-major axis:

$$\varepsilon = \frac{v_0^2}{2} - \frac{\mu}{r(0)} = -\frac{\mu}{2a} \tag{78}$$

which gives us $\varepsilon = -2.1252 \times 10^7 \; m^2/s$. We can conclude that the orbit is closed ($\varepsilon < 0$). Also, we get $a = -\mu/2\varepsilon = 9378.14 \; km$. Since $r_0 \neq a$, the orbit is elliptic.

Next, compute the angular momentum vector at $t = 0$,

$$\begin{aligned} \mathbf{h}_{\mathcal{I}}(0) &= \mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0) \\ &= \{1.3073, -0.7548, 5.6337\}^{\mathrm{T}} \times 10^{10} \;\; m^2/s \end{aligned} \tag{79}$$

Recall that the angular momentum vector is an integral of motion, i.e., $\mathbf{h}(0) = \mathbf{h}(t) \; \forall t$. We will simply write it as $\mathbf{h}$. The unit angular momentum vector is,

$$\hat{\mathbf{i}}_{h,\mathcal{I}} = \frac{\mathbf{h}_{\mathcal{I}}}{h} = \{0.2241, -0.1294, 0.9659\}^{\mathrm{T}} \tag{80}$$

In the above equation, $h = \|\mathbf{h}_{\mathcal{I}}\| = 5.8324 \times 10^{10} \;\; m^2/s$. The vector $\hat{\mathbf{i}}_h$ identifies the plane of the motion since it is the normal to the orbital plane.

Next, let us look at the eccentricity vector:

$$\begin{aligned} \mathbf{c}_{\mathcal{I}} &= \mu e \hat{\mathbf{i}}_{e,\mathcal{I}} = \dot{\mathbf{r}}_{\mathcal{I}}(0) \times \mathbf{h}_{\mathcal{I}} - \frac{\mu}{r_0}\mathbf{r}_{\mathcal{I}}(0) \\ &= \{0.0176, 1.8657, 0.1547\}^{\mathrm{T}} \times 10^{14} \;\; m^3/s^2 \quad \text{(typo fixed in red)} \end{aligned} \tag{81}$$

Thus, $c = \|\mathbf{c}_{\mathcal{I}}\| = 1.1958 \times 10^{14} \; m^3/s^2$. Since $c = \mu e$, the orbit eccentricity is:

$$e = \frac{c}{\mu} = 0.3 \tag{82}$$

Again, we get confirmation that the orbit is eccentric. Also, the unit eccentricity vector is:

$$\hat{\mathbf{c}}_{\mathcal{I}} = \hat{\mathbf{i}}_{e,\mathcal{I}} = \frac{\mathbf{c}_{\mathcal{I}}}{c} = \{0.0148, 0.9915, 0.1294\}^{\mathrm{T}} \tag{83}$$

The above unit vector identifies the direction towards the point of closest approach, i.e. periapsis. To complete the triad, we get the unit vector $\hat{\mathbf{i}}_{y,\mathcal{I}}$:

$$\hat{\mathbf{i}}_{y,\mathcal{I}} = \frac{\hat{\mathbf{i}}_{h,\mathcal{I}} \times \hat{\mathbf{i}}_{e,\mathcal{I}}}{\|\hat{\mathbf{i}}_{h,\mathcal{I}} \times \hat{\mathbf{i}}_{e,\mathcal{I}}\|} = \{-0.9744, -0.0147, 0.2241\}^{\mathrm{T}} \tag{84}$$

The unit vectors $\hat{\mathbf{i}}_{e,\mathcal{I}}$, $\hat{\mathbf{i}}_{y,\mathcal{I}}$ and $\hat{\mathbf{i}}_{h,\mathcal{I}}$ can be arranged using Eq.(69) (or Eq.(72b)) to get

$$\mathbf{R}_{\mathcal{O}\mathcal{I}} = \begin{Bmatrix} \hat{\mathbf{i}}_{e,\mathcal{I}}^{\mathrm{T}} \\ \\ \hat{\mathbf{i}}_{y,\mathcal{I}}^{\mathrm{T}} \\ \\ \hat{\mathbf{i}}_{h,\mathcal{I}}^{\mathrm{T}} \end{Bmatrix} = \begin{bmatrix} 0.0148 & 0.9915 & 0.1294 \\ -0.9744 & -0.0147 & 0.2241 \\ 0.2241 & -0.1294 & 0.9659 \end{bmatrix} \tag{85}$$

Orientation angles: We next consider the orientation of the orbit with respect to the inertial reference frame via the foregoing 3-1-3 Euler sequence:

1. Inclination: Using Eq.(73a),

$$i = \cos^{-1}(i_{h_3}) \underset{\text{Eq.}[80]}{=} \cos^{-1} 0.9659 = 15 \deg \tag{86}$$

2. Longitude of ascending node: From Eq.(73b),

$$\Omega \;\mapsto \begin{cases} c\Omega \;=\; -i_{h_2}/si = 0.5000 \\ s\Omega \;=\; i_{h_1}/si = 0.8660 \end{cases} \quad ; \quad \Rightarrow \Omega \;=\; 60\,\text{deg} \tag{87}$$

3. Argument of periapsis: From Eq.(73c),

$$\omega \;\mapsto \begin{cases} c\omega \;=\; i_{y_3}/si = 0.8660 \\ s\omega \;=\; i_{e_3}/si = 0.5000 \end{cases} \quad ; \quad \Rightarrow \omega \;=\; 30\,\text{deg} \tag{88}$$

   Let us also determine the initial true anomaly. Using the rotation matrix found in Eq.(220),

$$\mathbf{r}_{\mathcal{O}}(0) \;=\; \mathbf{R}_{\mathcal{O}\mathcal{I}}\,\mathbf{r}_{\mathcal{I}}(0) \;=\; [4.9784, 4.9784, 0] \times 10^6 \; m \tag{89}$$

   Therefore, using the above in Eq.(76),

$$f_0 \;\mapsto \begin{cases} \cos f_0 \;=\; r_{0\mathcal{O}1}/r_0 = 0.7071 \\ \sin f_0 \;=\; r_{0\mathcal{O}2}/r_0 = 0.7071 \end{cases} \quad ; \quad \Rightarrow f_0 \;=\; 45\,\text{deg} \tag{90}$$

Note we used method 1 to determine f_0.

Plotting the orbit in 3D: We will now plot the orbit in 3D and show the initial location of the spacecraft in it. To do this, start with a set of true anomalies sampled between 0 and 2π, e.g. using linspace in MATLAB®. Let these be labeled $\{f_1, f_2, \dots f_N\}$, where N is the number of points chosen and $f_1 = 0$, $f_N = 2\pi$. For each of these values, determine the radius vector expressed in the *orbital frame*. This is a fairly easy two-step process:

1. Determine $r_k = r(f_k)$ = radius magnitude at $f = f_k$ using the equation of orbit: $r_k = p/(1 + e\cos f_k)$.
2. Employ Eq.(75) to get:

$$\mathbf{r}_{\mathcal{O}}(f_k) \;=\; [r_k \cos f_k,\; r_k \sin f_k,\; 0]^{\mathrm{T}} \tag{91}$$

Finally, use the $\mathbf{R}_{\mathcal{O}\mathcal{I}}$ matrix to obtain $\mathbf{r}_{\mathcal{I}}(f_k) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\;\mathbf{r}_{\mathcal{O}}(f_k)$. For each $f(k)$, plot $\mathbf{r}_{\mathcal{I}}(f_k)$ using "plot3" in MATLAB®(see code in Carmen) to obtain the 3D orbit. Fig.(24) shows the orbit corresponding to the current example problem. (all lengths are scaled by the semi-major axis, i.e. $r \to r/a$).

Finally, we can determine the other desired quantities:

- radius of periapsis: $r_p \;=\; a(1-e) \;=\; 6564.7\; km$. Using vis-viva, $v_p \;=\; 8.8845\; kmps$.
- radius of apoapsis: $r_a \;=\; a(1+e) \;=\; 12,191.7\; km$. Using vis-viva, $v_a \;=\; 4.7839\; kmps$.
- Period: $P \;=\; 2\pi\sqrt{a^3/\mu} \;=\; 2.511\; hrs$,

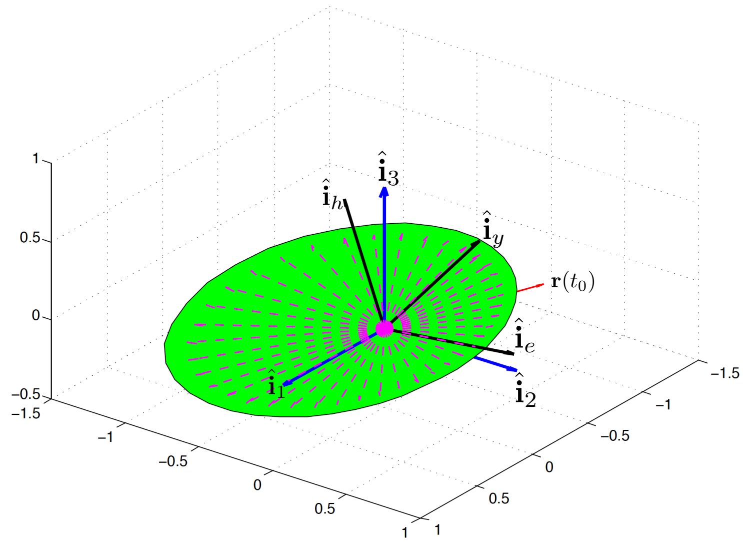


Figure 24:

11 Solution of the Two Body Problem

In the previous section, we studied the orbital elements of the two-body problem. They provided us tremendous (analytical) detail into the mechanics of the problem, but we fell short of actually solving the system. In this section, we will look at different ways of solving the fundamental equation. To be clear, let us begin with the problem statement.

11.1 Problem Statement - Two Body Problem

Given the dynamics:

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r} \quad \text{(dynamics)} \tag{92}$$

and initial conditions:

$$\mathbf{r}_{\mathcal{I}}(t_0) = \mathbf{r}_{\mathcal{I}}(0) \tag{93a}$$
$$\mathbf{v}_{\mathcal{I}}(t_0) = \mathbf{v}_{\mathcal{I}}(0) \quad \text{(initial conditions)} \tag{93b}$$

determine the states $\{\mathbf{r}_{\mathcal{I}}(t),\ \mathbf{v}_{\mathcal{I}}(t)\}$ for all times $t > t_0$.

In other words, we would like the time histories of all the states of the two-body system. Recall that the above system involves six scalar first order ordinary differential equations.

11.2 Popular Methods of Solution

The following is a list of popular methods of solving the fundamental equation. Please note that the list is not exhaustive. This is a very old problem ($\sim$ 450 years) and many types of methods have been developed.

I. Explicit numerical integration: numerical integrators (e.g. `ode45` of MATLAB®) are used to explicitly solve the two body problem in "table form" (details below).

II. Implicit solution via Kepler's equation: This approach utilizes the orbital elements and is an elegant way of solving the problem.

III. F&G solution: This method also uses orbital elements and in particular, the planarity of motion.

IV. Universal solution: Developed by Sam Herrick (John Junkins' advisor @ UCLA), this is an elegant generalization of the implicit solution.

V. Power series solution: sort of a "brute force method": we will not consider this approach.

12 Explicit Numerical Integration of TBP

- The difficulties in solving Eq.(92) arise because of its nonlinearity. The "obvious" way of solving any nonlinear system is to numerically integrate it using an "adequate" integrator. The adequacy of a numerical algorithm is determined by the level of accuracy required by the system. E.g., the two body problem requires highly accurate numerical integrators because the time duration of integration can be very long - on the order of hours (e.g. low Earth orbits) or days (e.g geostationary orbits), sometimes years (e.g. mission to Mars). As a result, errors can *pile up* over time leading to incorrect orbits.

- Eq.(92) can be split into two first order vector ODEs:

$$\dot{\mathbf{r}} = \mathbf{v} \tag{94a}$$

$$\dot{\mathbf{v}} = -\frac{\mu}{r^3}\mathbf{r} \tag{94b}$$

where, ($\mathbf{r}$, $\mathbf{v}$) together form the "state" of the TBP system.

- Note that Eqs.(94) are abstract, i.e. can be expressed in any desired reference frame. For further analysis, we need to consider the individual components of the vectors $\mathbf{r}$ and $\mathbf{v}$. In other words, we need to select a reference frame. We will express all vectors in the inertial reference frame, such that:

$$\dot{\mathbf{r}}_{\mathcal{I}} = \mathbf{v}_{\mathcal{I}} \tag{95a}$$

$$\dot{\mathbf{v}}_{\mathcal{I}} = -\frac{\mu}{r^3}\mathbf{r}_{\mathcal{I}} \tag{95b}$$

Of course, recall that all derivatives appearing above are measured in the inertial reference frame. The superscript has been dropped for notational convenience.

- Let the components of vectors $\mathbf{r}_{\mathcal{I}}$ and $\mathbf{v}_{\mathcal{I}}$ be given as:

$$\mathbf{r}_{\mathcal{I}} = \begin{Bmatrix} r_x \\ r_y \\ r_z \end{Bmatrix}, \quad \mathbf{v}_{\mathcal{I}} = \begin{Bmatrix} v_x \\ v_y \\ v_z \end{Bmatrix} \tag{96}$$

- Thus Eq.(95a) becomes:

$$\dot{\mathbf{r}}_{\mathcal{I}} \underset{\text{Eq.(96)}}{=} \begin{Bmatrix} \dot{r}_x \\ \dot{r}_y \\ \dot{r}_z \end{Bmatrix} = \mathbf{v}_{\mathcal{I}} \underset{\text{Eq.(96)}}{=} \begin{Bmatrix} v_x \\ v_y \\ v_z \end{Bmatrix} \tag{97}$$

Also, $r = \|\mathbf{r}\| = \|\mathbf{r}_{\mathcal{I}}\| = \sqrt{r_x^2 + r_y^2 + r_z^2}$. Therefore, Eq.(95b) becomes

$$\dot{\mathbf{v}}_{\mathcal{I}} \underset{\text{Eq.(96)}}{=} \begin{Bmatrix} \dot{v}_x \\ \dot{v}_y \\ \dot{v}_z \end{Bmatrix} = -\frac{\mu}{r^3}\mathbf{r}_{\mathcal{I}} \underset{\text{Eq.(96)}}{=} -\frac{\mu}{(r_x^2 + r_y^2 + r_z^2)^{3/2}} \begin{Bmatrix} r_x \\ r_y \\ r_z \end{Bmatrix} \tag{98}$$

Collect the six scalar first order ODEs from Eqs.(97) and (98) to get the complete two-body system in "state-space form":

$$\dot{r}_x = v_x \tag{99a}$$

$$\dot{r}_y = v_y \tag{99b}$$
$$\dot{r}_z = v_z \tag{99c}$$
$$\dot{v}_x = -\frac{\mu r_x}{(r_x^2 + r_y^2 + r_z^2)^{3/2}} \tag{99d}$$
$$\dot{v}_y = -\frac{\mu r_y}{(r_x^2 + r_y^2 + r_z^2)^{3/2}} \tag{99e}$$
$$\dot{v}_z = -\frac{\mu r_z}{(r_x^2 + r_y^2 + r_z^2)^{3/2}} \tag{99f}$$

Equations (99a)-(99f) must be integrated simultaneously to obtain $\mathbf{r}_{\mathcal{I}}(t) = \{r_x(t), r_y(t), r_z(t)\}$ and $\mathbf{v}_{\mathcal{I}}(t) = \{v_x(t), v_y(t), v_z(t)\}$, given the initial conditions $\mathbf{r}_{\mathcal{I}}(t_0) = \mathbf{r}_{\mathcal{I}}(0) = \{r_x(0), r_y(0), r_z(0)\}$ and $\mathbf{v}_{\mathcal{I}}(t_0) = \mathbf{v}_{\mathcal{I}}(0) = \{v_x(0), v_y(0), v_z(0)\}$.

- Due to the nonlinearities in Eqs.(99) ($1/r^3$ term), the above equations cannot be integrated "by hand" and a numerical integrator is needed. MATLAB®'s `ode45` is convenient for this purpose. See attached code for examples.

- Accuracy is a big issue in explicit integration of the two body problem. Recall that in numerical integration, error accumulates over time. Given that typical time durations of integration in the current context can be very long (e.g. several hours), severe error buildup can occur. As a result, active *error control* is needed to ensure accurate results. You must explore the use of MATLAB®command "`odeset`" and its "`RelTol`" and "`AbsTol`" options.

  Important note: as mentioned above, error buildup is governed by the time-duration of integration. However, the length of time must be measured against the "degree of nonlinearity" of the underlying system. E.g. a system with "mild nonlinearity" can be integrated longer than a system with "strong nonlinearity" with the same integrator before error buildup becomes a concern. Unfortunately, there is no definitive way of measuring the degree of a system's nonlinearity.

- Battin's chapter 12 provides a useful collection of *higher order* (read *more accurate*) numerical integrators especially suited to the challenges of integrating Eqs.(99a). This system suffers the dual curse of significant nonlinearity and long time durations of proagation.

- To finish, note that the solution obtained in this *explicit* approach is in the form of a table, something like shown in Table.(3) Clearly, the first row is specified by the initial conditions. The final time

time (t)	$r_x(t)$	$r_y(t)$	$r_z(t)$	$v_x(t)$	$v_y(t)$	$v_z(t)$
	$\longleftarrow$	$\mathbf{r}_{\mathcal{I}}(t)$	$\longrightarrow$	$\longleftarrow$	$\mathbf{v}_{\mathcal{I}}(t)$	$\longrightarrow$
t_0	$r_x(0)$	$r_y(0)$	$r_z(0)$	$v_x(0)$	$v_y(0)$	$v_z(0)$
t_1	$r_x(1)$	$r_y(1)$	$r_z(1)$	$v_x(1)$	$v_y(1)$	$v_z(1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
t_{N-1}	$r_x(N-1)$	$r_y(N-1)$	$r_z(N-1)$	$v_x(N-1)$	$v_y(N-1)$	$v_z(N-1)$
$t_N = t_f$	$r_x(N)$	$r_y(N)$	$r_z(N)$	$v_x(N)$	$v_y(N)$	$v_z(N)$

Table 3: Solution in Tabular Form: Explicit Integration of TBP

$(t_f = t_N)$ is typically also set by the user. The number of points selected between t_0 and t_f is determined by the accuracy of the numerical integrator used. This is sometimes referred to as the resolution of time differentiation. Note that depending on the integrator, the *step-size* may not be uniform, i.e. $(t_k - t_{k-1})$ may be different for different values of $k = 2, 3, \ldots N$. E.g. `ode45` is a *variable step integrator*. It is accurate up to fourth-order Runge Kutta (RK) scheme of integration and imposes a threshold on the order of the fifth-order RK term to control the step-size.

- Suppose you need the answer at a time t^* which is not included in the first column of the table above. You must interpolate using the *available data*, i.e. Table (3) to obtain the answer $\{\mathbf{r}_{\mathcal{I}}(t^*), \mathbf{v}_{\mathcal{I}}(t^*)\}$. (See help on `interp1` in MATLAB®)
- **PROS** of Direct Numerical Integration:
  i. The method is direct and hence *simple*.
  ii. It is universally applicable to all types of initial conditions, i.e. you need not pre-determine whether the underlying trajectory is a circle, ellipse, parabola or hyperbola.
- **CONS** of Direct Numerical Integration:
  i. There can be significant error buildup and therefore close monitoring of the performance of integration is required.
  ii. It is not possible to obtain any real insight into the physical properties of the orbit. E.g. how big is the orbit? What is its eccentricity, inclination, etc. None of this information is available in Table (3).
  iii. The integration is good only for the particular initial conditions given in the first row. If the initial conditions change, the entire integration process must be repeated from start to finish. This can be quite time-consuming if there is uncertainty in initial conditions.
  iv. The process shown above is called *time marching.* Suppose that given $\{\mathbf{r}_{\mathcal{I}}(t_0),\ \mathbf{v}_{\mathcal{I}}(t_0)\}$, we are only interested in determining $\{\mathbf{r}_{\mathcal{I}}(t_0 + P/4),\ \mathbf{v}_{\mathcal{I}}(t_0 + P/4)\}$ where P is the orbital period. Using numerical integration, it is not possible to directly *map the initial state* $\{\mathbf{r}_{\mathcal{I}}(t_0),\ \mathbf{v}_{\mathcal{I}}(t_0)\}$ *into the desired final state* $\{\mathbf{r}_{\mathcal{I}}(t_0 + P/4),\ \mathbf{v}_{\mathcal{I}}(t_0 + P/4)\}$. In other words, we must build the table shown above, perhaps setting $t_f = t_0 + P/4$ and march through *all intermediate times, spaced closely enough to ensure that error buildup is under control.* This is a tremendous cost to pay to simply obtain the answer at a single point in time in the future.

13 Implicit Solution via Kepler's Equation

Clearly, direct numerical integration is good in its simplicity but has its limitations. Recall that we have already built a wealth of information about the two-body problem using the integrals of motion in Sec.(5). Given the initial conditions of Eqs.(93), we were able to perform the following computations:

$$\{\mathbf{r}_{\mathcal{I}}(0), \mathbf{v}_{\mathcal{I}}(0)\} \longrightarrow \{\mathbf{h}_{\mathcal{I}},\, \mathbf{c}_{\mathcal{I}}, \varepsilon\} \longrightarrow \mathbf{R}_{\mathcal{OI}} \;=\; \mathbf{R}_{\mathcal{OI}}(\hat{\mathbf{i}}_{e,\mathcal{I}},\ \hat{\mathbf{i}}_{y,\mathcal{I}},\ \hat{\mathbf{i}}_{h,\mathcal{I}}) \tag{100}$$

where, $\mathbf{h}_{\mathcal{I}}$ and $\mathbf{c}_{\mathcal{I}}$ are the angular momentum and eccentricity vectors respectively, both expression in the inertial frame, and $\hat{\mathbf{i}}_{e,\mathcal{I}}$ and $\hat{\mathbf{i}}_{h,\mathcal{I}}$ are the unit eccentricity and angular momentum vectors in the inertial frame and $\hat{\mathbf{i}}_{y,\mathcal{I}}$ completes the triad (pointing in the direction of $(f = \pi/2)$). Using one of the three methods described in Sec.(10.1), it is also possible to find $f(t_0) = f(0)$.

- Conversely, given $f(t)$, it is possible to find $\mathbf{r}_{\mathcal{I}}(f(t))$ as follows:

$$\mathbf{r}_{\mathcal{O}}(f(t)) \;=\; \begin{Bmatrix} r(t)\cos f(t) \\ r(t)\sin f(t) \\ 0 \end{Bmatrix}, \quad \text{where, } r(t) \;=\; \frac{p}{1 + e\cos f(t)} \tag{101a}$$

$$\text{Therefore,} \quad \mathbf{r}_{\mathcal{I}}(f(t)) \underset{\text{Eq.(100)}}{=} \mathbf{R}_{\mathcal{OI}}^{\mathrm{T}}\, \mathbf{r}_{\mathcal{O}}(f(t)) \tag{101b}$$

While it is not shown here, it is also possible to obtain $\mathbf{v}_{\mathcal{I}}(f(t))$. In essence, we can build a table similar to Table (3), except that instead of the first column being time, we must use true anomaly. Over a full period, f lies in the range $[0, 2\pi]$ and therefore, by selecting *enough number of points* f_k (you decide how many points is "smooth enough" for you), the entire orbit can be drawn in three dimensions: see Table (4).

True Anomaly (f)	$r_x(f)$	$r_y(f)$	$r_z(f)$
	$\longleftarrow$	$\mathbf{r}_{\mathcal{I}}(f)$	$\longrightarrow$
$f_0 = 0$	$r_x(f_0)$	$r_y(f_0)$	$r_z(f_0)$
f_1	$r_x(f_1)$	$r_y(f_1)$	$v_z(f_1)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_{N-1}	$r_x(f_{N-1})$	$r_y(f_{N-1})$	$r_z(f_{N-1})$
$f_N = 2\pi$	$r_x(f_N)$	$r_y(f_N)$	$r_z(f_N)$

Table 4: Implicit Solution of TBP in Tabular Form

- This is great, because no numerical approximations are needed - a fully analytical and closed-form system of equations have been derived to obtain $\mathbf{r}_{\mathcal{I}}$ as a function of the true anomaly.
- The problem is, that we need not $\mathbf{r}_{\mathcal{I}}(f)$, but $\mathbf{r}_{\mathcal{I}}(t)$. This is why the current method is called an "implicit approach". In order to complete the solution procedure, we must find the true anomaly as a function of time, i.e. $f(t)$ so that to each f in Table (4) we can assign a *time stamp* and obtain the desired states $\{\mathbf{r}_{\mathcal{I}}(t), \mathbf{v}_{\mathcal{I}}(t)\}$. This "missing link" between f and t is the Kepler's equation that we will soon encounter.
- If we can somehow determine an expression for $\dot{f}$, we can hopefully integrate it to obtain $f(t)$. To this end, let us revisit the equations of planar kinematics:

$$\mathbf{r}_{\mathcal{E}} = r\hat{\mathbf{e}}_r \tag{102a}$$
$$\mathbf{v}_{\mathcal{E}} = \dot{r}\hat{\mathbf{e}}_r + r\dot{f}\hat{\mathbf{e}}_f \tag{102b}$$

where the local frame $\mathcal{E} = \{F, \hat{\mathbf{e}}_r, \hat{\mathbf{e}}_f, \hat{\mathbf{i}}_h\}$ has been used to coordinateize the vectors $\mathbf{r}$ and $\mathbf{v}$ above. Consider the angular momentum vector:

$$\mathbf{h}_{\mathcal{E}} = \mathbf{r}_{\mathcal{E}} \times \mathbf{v}_{\mathcal{E}} \underset{\text{Eq.(102)}}{=} r^2\dot{f}\,\hat{\mathbf{i}}_h \tag{103}$$

Considering only the magnitudes in the above equation, $h = r^2\dot{f}$. Recall the following universal (i.e. valid for all conic types) relationships: $h^2 = \mu p$ and $r = p/(1+e\cos f)$ (equation of orbit). Combining these into Eq.(103),

$$\sqrt{\mu p} = \frac{p^2}{(1+e\cos f)^2}\frac{df}{dt} \tag{104}$$

Or,

$$\boxed{\frac{\sqrt{\mu}}{p^{3/2}}dt = \frac{df}{(1+e\cos f)^2}} \tag{105}$$

The above equation captures the elusive time-dependence of true-anomaly that we have been missing. Note that Eq.(105) is <u>universal</u>. Integrating, we get

$$\frac{\sqrt{\mu}}{p^{3/2}}(t-t_0) = \underbrace{\int_{f_0}^{f(t)} \frac{df}{(1+e\cos f)^2}}_{\textbf{not fun!!!}} \tag{106}$$

The integral appearing on the right hand side of Eq.(106) is brutal! There is no good way to evaluate it analytically in its present form. So, we have ultimately hit the "bottleneck" caused by nonlinearity of the governing dynamics (Eq.(92)): we almost got away with it and Eq.(106) was basically the final step!

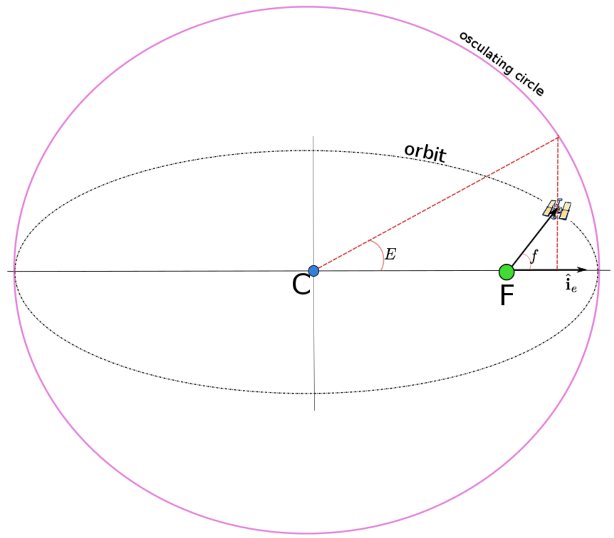


Figure 25: The Eccentric Anomaly

- When faced with difficult integrals, a strategy commonly adopted is a change of variables. We will do the same here to try and extricate the RHS of Eq.(106). Instead of using true anomaly as our measure of the angular position in orbit, we will try and use the *eccentric anomaly*. The eccentric anomaly first appeared in Table (2) in Sec.(9). As a reminder, it is illustrated again in Fig.(25) below. The eccentric anomaly is obtained by projecting the object's location in orbit up on to the osculating circle, then joining the point of projection *with the center of the ellipse* (not focus): see Fig.(25). There is a one-one correspondence between the true and eccentric anomalies and the relationship between them is (from geometrical considerations):

$$\tan\frac{E}{2} \;=\; \sqrt{\frac{1-e}{1+e}}\tan\frac{f}{2} \tag{107}$$

  The above equation is the transformation between our old variable (f) and the new one (E). Eq.(107) is very useful because you need not worry about the correctness of quadrant of the eccentric anomaly: it always gives an answer in the same quadrant as the true anomaly, which is where E lies (see Fig.(25)).

- There exists **another method** for finding the eccentric anomaly at a point in the orbit: given $\{\mathbf{r},\ \mathbf{v}\}$, define

$$\sigma \;\triangleq\; \frac{\mathbf{r}\cdot\mathbf{v}}{\sqrt{\mu}} \tag{108}$$

  Then,

$$\tan E \;=\; \frac{\sigma}{\sqrt{a}(1-r/a)} \tag{109}$$

- Before we proceed any further, it is important to realize that in exchanging f for E, we have made a serious compromise! Recall that the eccentric anomaly is defined only for ellipses (and of course circles, where $f = E$). In other words, our change of variable comes at a steep cost: all developments from this point onward are valid only for elliptical (and circular) orbits. You must therefore first make sure before that the underlying orbit is an ellipse (i.e. determine the eccentricity).

- We now repeat the process above with the new angular variable, E. Instead of the $\mathcal{E}$ frame (Eqs.(102)), we will now work in the $\mathcal{O}$ frame. Let

$$\mathbf{r}_{\mathcal{O}} \;=\; x\,\hat{\mathbf{i}}_e \;+\; y\,\hat{\mathbf{i}}_y \tag{110}$$

Of course, the z-component of $\mathbf{r}_\mathcal{O}$ is zero because $\hat{\mathbf{i}}_e - \hat{\mathbf{i}}_y$ is the plane of the orbit. It is known from geometry of ellipses that:

$$x = a(\cos E - e) \tag{111}$$
$$y = a\sqrt{1-e^2}\sin E \tag{112}$$

While Eq.(110) is universal, we emphasize again that Eqs.(246a)-(246b) hold only for ellipses. Differentiating (inertially), we get:

$$\mathbf{v}_\mathcal{O}^\mathcal{I} = \dot{x}\,\hat{\mathbf{i}}_e + \dot{y}\,\hat{\mathbf{i}}_y \tag{113}$$

Note that the chain-rule was applied in going from Eq.(110) to Eq.(113), but $\dot{\hat{\mathbf{i}}}_e = \dot{\hat{\mathbf{i}}}_y = 0$. This is so because the vectors $\hat{\mathbf{i}}_e$ and $\hat{\mathbf{i}}_y$ are non-rotating (after-all the orbital reference frame, $\mathcal{O} = \{F, \hat{\mathbf{i}}_e, \hat{\mathbf{i}}_y, \hat{\mathbf{i}}_h\}$ is an inertial reference frame). Please be sure you are convinced of this.

Now, differentiating Eqs.(246a) and (246b), we get:

$$\dot{x} = -a\sin E\,\frac{dE}{dt} \tag{114}$$
$$\dot{y} = a\sqrt{1-e^2}\cos E\,\frac{dE}{dt} \tag{115}$$

Therefore, the angular momentum vector becomes:

$$\mathbf{h}_\mathcal{O} = \mathbf{r}_\mathcal{O}\times\mathbf{v}_\mathcal{O} = (x\dot{y} - y\dot{y})\,\hat{\mathbf{i}}_h \tag{116}$$

Taking only the magnitude,

$$\begin{aligned} h &= (x\dot{y} - y\dot{x}) \\ &\underset{\text{Eqs.(246a),(246b),(246c),(246d)}}{=} a^2\sqrt{1-e^2}(\cos^2 E + \sin^2 E - e\cos E)\dot{E} \\ &= a^2\sqrt{1-e^2}(1-e\cos E)\dot{E} \end{aligned} \tag{117}$$

Now, we use $h = \sqrt{\mu p} = \sqrt{\mu a(1-e^2)}$ to get

$$\sqrt{\mu a(1-e^2)} = a^2\sqrt{1-e^2}(1-e\cos E)\dot{E} \tag{118}$$

or,

$$\sqrt{\frac{\mu}{a^3}}dt = (1-e\cos E)dE \tag{119}$$

The above equation is clearly much easier to integrate (compared to Eq.(106)):

$$\sqrt{\frac{\mu}{a^3}}(t-t_0) = \int_{E_0}^{E(t)}(1-e\cos E)dE \tag{120}$$

Or,

$$\sqrt{\frac{\mu}{a^3}}(t-t_0) = (E - e\sin E)\Big|_{E_0}^{E(t)} \tag{121}$$

Eq.(121) opens up to give:

$$\boxed{(E(t) - e\sin E(t)) - (E_0 - e\sin E_0) = \sqrt{\frac{\mu}{a^3}}(t-t_0)} \tag{122}$$

The above is a version of the well known **Kepler's equation**. We reiterate that it holds only for elliptical orbits. E is the eccentric anomaly at the current time (t), and E_0 is the initial eccentric anomaly (at t_0). For brevity, we will write $E(t) \equiv E$.

- Recall the "mean motion", which is the average angular speed of the object in orbit:

$$n = \frac{2\pi}{P} = \sqrt{\frac{\mu}{a^3}} \tag{123}$$

Also, define the so-called *mean anomaly*:

$$M \triangleq E - e\sin E \tag{124}$$

such that $E_0 - e\sin E_0 = M_0$. Eq.(122) then becomes

$$\boxed{E - e\sin E = M = M_0 + n(t - t_0)} \tag{125}$$

The above version of Kepler's equation is cleaner than Eq.(122) in the sense that everything known (t, E_0) is pushed to the RHS and the LHS contains only the unknown, namely, $E = E(t)$.

- The sequence of steps to go from $\{\mathbf{r}_\mathcal{I}(t_0),\, \mathbf{v}_\mathcal{I}(t_0)\}$ to M_0 is the following:

$$\{\mathbf{r}_\mathcal{I}(t_0),\, \mathbf{v}_\mathcal{I}(t_0)\} \;\underset{\longrightarrow}{\text{Sec.(10)}}\; f_0 \;\underset{\longrightarrow}{\text{Eq.(107)}}\; E_0 \;\underset{\longrightarrow}{\text{Eq.(124)}}\; M_0$$

- **Kepler's equation: special case.** The most popular form of the Kepler's equation arises when t_0 is the moment of periapsis passage, i.e. $t_0 = t_P$. In this case, the time elapsed, namely $(t-t_0) = (t-t_P)$ is defined as the *time since periapsis passage* and is denoted by τ_P:

$$\tau_P \triangleq (t - t_P) \tag{126}$$

In this case, $f_0 = f_P = 0$. From Eq.(107), $E_0 = E_P = 0$ and from Eq.(124), $M_0 = M_P = 0$. Eq.(125) therefore specializes to:

$$\boxed{E - e\sin E = n\,\tau_P} \tag{127}$$

As mentioned above, the above is the most popular version of Kepler's equation, sometimes called the *Classical Keplers's equation.* It is only a special case of the more general form given in Eq.(125).

- Clearly, the unknown in Eqs.(125) or (127) is the current eccentric anomaly, E and it is locked inside a nonlinear, in fact transcendental equation. All that is left to solve the two-body problem (for ellipses only) is to solve Eq.(125) for E.

13.1 Solution of Kepler's Equation via Newton Raphson Iterations

We have walked down a truly (mathematically) elegant path involving the integrals of motion of the two body problem, leading to the orbital elements and finaling culminating in Kepler's equation. All developments were analytical and we derived closed form expressions for all quantities involved. The nonlinear in the governing dynamics has however manifested in our final step and there is no way to avoid it. Eq.(125) (or Eq.(127)) is a transcendental equation that researchers have tried to solve for the past about 450 years. No one has succeeded but the result of these efforts has been the development of an incredibly large number of elegant numerical root-finding algorithms.

We must realize that Eq.(125) is only an algebraic equation. In other words, we have found a *mapping* from the states at the initial time $\{\mathbf{r}_\mathcal{I}(t_0),\, \mathbf{v}_\mathcal{I}(t_0)\}$, to the states at the final time, $\{\mathbf{r}_\mathcal{I}(t),\, \mathbf{v}_\mathcal{I}(t)\}$, precluding the need for marching through time from t_0 to t in small increments. By itself, this is a significant improvement over the explicit numerical integration method involving `ode45`.

As mentioned above, many numerical techniques have been designed to approximate the eccentric anomaly that solves Eq.(125). We will only consider one of the simplest method, namely the Newton-Raphson (NR) itertive approach.

- The essence of NR iterations is very simple: begin with a guess and then find a correction that improves the guess. Keep updating (i.e. finding corrections) until the most recent approximation is close enough to the truth (in terms of equation error). Below, we lay out the method.
- Remember our unknown is E.
- Begin with an *initial guess*: $E^{(0)}$. We will call this the **zeroth approximation**.
  Note the difference between the notation ($E(0) = E_0$) and $E^{(0)}$.
- *Solution correction:* Find an update that improves the guess:

$$E^{(1)} \ = \ E^{(0)} \ + \ \Delta E^{(1)} \tag{128}$$

where, $E^{(1)}$ is the **first approximation** and $\Delta E^{(0)}$ is the **first correction**.

The obvious question is, what should be the criterion for selection of the correction term?

The answer is just as obvious: find a correction such that the updated approximation ($E^{(1)}$) exactly solves the Kepler's equation! In other words,

$$E^{(1)} \ - \ e \sin E^{(1)} \ = \ M \tag{129}$$

or, using Eq.(128),

$$\left(E^{(0)} \ + \ \Delta E^{(1)}\right) \ - \ e \sin\left(E^{(0)} \ + \ \Delta E^{(1)}\right) \ = \ M \tag{130}$$

... solve for $\Delta E^{(1)}$. $E^{(0)}$ is known.

Did we really achieve anything by introducing the above prediction-correction type of methodology? We started with a nonlinear equation, and our current unknown ($\Delta E^{(1)}$) is still trapped inside a transcendental nonlinear equation (Eq.(130)).

This is what we achieved: we re-formulated the problem to replace the unknown from E (a potentially large value) to ΔE (hopefully a small value)... thus giving us the license to ***linearly approximate*** the nonlinear term appearing in Eq.(130). The linear approximation can made via the Taylor's series expansion of nonlinear functions. We will expand the term $\sin\left(E^{(0)} + \Delta E^{(1)}\right)$ "about" the reference point $E^{(0)}$, "in terms of" the correction term $\Delta E^{(1)}$:

$$\begin{aligned}\sin\left(E^{(0)} + \Delta E^{(1)}\right) \ = \ \sin E^{(0)} \ + \ \Delta E^{(1)} \, \frac{d \sin E}{dE}\Big|_{E^{(0)}} \ + \ \frac{\left(\Delta E^{(1)}\right)^2}{2} \, \frac{d^2 \sin E}{d^2 E}\Big|_{E^{(0)}} \\ \dots \ + \ \frac{\left(\Delta E^{(1)}\right)^k}{k!} \, \frac{d^k \sin E}{d^k E}\Big|_{E^{(0)}} + \dots\end{aligned} \tag{131}$$

Some observations:

- Taylor's expansion is an incredible tool - it converts a general form of nonlinearity (so long as its derivatives exist at the point of expansion) into polynomial nonlinearity. Note that the RHS in the above equation contains only polynomial terms in the unknown, $\Delta E^{(1)}$. Note: this is related to the so-called *Weierstrass approximation theorem* which states that any continuous function can be approximated arbitrarily well using only polynomial functions.
- The equality holds in Eq.(131) in general only when an infinite number of terms are included in the expansion in the RHS.

In order to simplify the above expansion, we will make the ***assumption*** that the correction term, $\Delta E^{(1)}$ is small. In other words, $\Delta E^{(1)} \gg \left(\Delta E^{(1)}\right)^2$. In general,

$$\Delta E^{(1)} \gg \left(\Delta E^{(1)}\right)^k, \quad k \geq 2 \tag{132}$$

Note that this assumption may be false! If the initial guess is already close to the true answer (which there is no definitive way of knowing), the assumption may be acceptable (since the required correction is small). But if not, i.e. we start with a bad guess, then the assumption may be completely false. The Newton-Raphson method makes this assumption nevertheless and hopes for the best!

The benefit of making this assumption is that the Taylor series now reduces to a linear expansion:

$$\begin{aligned} \sin\left(E^{(0)} + \Delta E^{(1)}\right) &\approx \sin E^{(0)} + \Delta E^{(1)} \frac{d\sin E}{dE}\Big|_{E^{(0)}} \\ &= \sin E^{(0)} + \Delta E^{(1)} \cos E^{(0)} \end{aligned} \tag{133}$$

Using Eq.(133) in Eq.(130), we get:

$$\left(E^{(0)} + \Delta E^{(1)}\right) - e\left(\sin E^{(0)} + \Delta E^{(1)} \cos E^{(0)}\right) \approx M \tag{134}$$

The equality is "wavy" ($\approx$) because of the approximation made in assuming linearity. We can now solve for the correction!

$$\Delta E^{(1)} = \frac{M - \left(E^{(0)} - e\sin E^{(0)}\right)}{\left(1 - e\cos E^{(0)}\right)} \tag{135}$$

Thus Eq.(128) becomes:

$$\underbrace{E^{(1)}}_{\text{1st approximation}} = \underbrace{E^{(0)}}_{\text{0th approximation}} + \underbrace{\frac{M - \left(E^{(0)} - e\sin E^{(0)}\right)}{\left(1 - e\cos E^{(0)}\right)}}_{\text{1st correction}} \tag{136}$$

- Following the same set of steps, we can now write a **second approximation**, as an improvement over the first:

$$E^{(2)} = E^{(1)} + \Delta E^{(2)} \tag{137}$$

Once again, the objective is to have the new approximation ($E^{(2)}$) exactly satisfy the Kepler's equation. We failed to meet this objective last time around because of the linearity approximation (Eq.(133)). Employing the Taylor's expansion again,

$$\Delta E^{(2)} = \frac{M - \left(E^{(1)} - e\sin E^{(1)}\right)}{\left(1 - e\cos E^{(1)}\right)} \tag{138}$$

This keeps going, such that the $k-th$ correction and approximation can be given as:

$$\Delta E^{(k)} = \frac{M - \left(E^{(k-1)} - e\sin E^{(k-1)}\right)}{\left(1 - e\cos E^{(k-1)}\right)} \qquad \text{(correction)} \tag{139}$$

$$E^{(k)} = E^{(k-1)} + \Delta E^{(k)} \qquad \text{(update)} \tag{140}$$

- It is expected (but not guaranteed) that

$$\left|\Delta E^{(k+1)}\right| < \left|\Delta E^{(k)}\right| \tag{141}$$

I.E. smaller and smaller corrections are needed. Correspondingly, the approximation becomes progressively better. There is no guarantee though that this trend will be displayed. Also, the following limiting behavior is to be expected:

$$\lim_{k\to\infty} \Delta E^{(k)} = 0 \tag{142}$$

- **Stopping criteria.** In actual implementation, the sequence of corrections and updates cannot go on forever! A threshold must be set such that if the magnitude of the latest correction is less than the threshold, iterations are stopped:

$$\left|\Delta E^{(k)}\right| \le \epsilon \tag{143}$$

where, ϵ is the specified threshold on the correction magnitude and the final approximation returned is $E^{(k)}$. Another breaking criterion can be "$k \geq N$", where N is an upper limit on the number of iterations to be performed (sort of an upper limit on computational capability). This type of a breaking condition makes your code robust - in the case when there is no convergence in sight and the iterations are simply meandering, setting an upper limit on the number of iterations will stop further computations. The answer will most likely not be very good, but at least the method will not enter an "infinite loop".

Discussion on initial guesses (e.g. what does $E^{(0} = 0$ correspond to and what does $E^{(0} = M(t)$ correspond to (circle) and uniqueness of solution.

ALGORITHM: Newton Raphson for Kepler's Equation

- **Step 1.** Initial guess: $E^{(0)}$. Looking at Eq.(125), $E^{(0)} = M$ is a good guess!
- **Step 2.** Set $k = 1$.
  - **Step 2a.** $k-th$ correction: Get $\Delta E^{(k)}$ from Eq.(140).
  - **Step 2b.** $k-th$ approximation: $E^{(k)} = E^{(k-1)} + \Delta E^{(k)}$.
- **Step 3.** Check for stopping criteria:

IF $\left|\Delta E^{(k)}\right| \le \epsilon$ (correction magnitude less than set threshold) OR $k \geq N$ (too many iterations!);

Return $E^{(k)}$ as the final approximation.

ELSE $k = k+1$ and return to step 2a.

Some notes on the above algorithm:

- The threshold ϵ is a user defined entity and its value depends on the application at hand.
- For moderately eccentric orbits, the convergence is typically very fast. You will hardly ever need more than five to six iterations to reach a satisfactory answer.
- The convergence becomes slower as the eccentricity increases. This fact can be tied to the discussion following immediately after Eq.(132): If e is large, our initial guess ($E^{(0)} = M$) will likely be off by a significant amount. This is because contribution of the nonlinear term ($e \sin E$) is large and the linearity assumption will likely be a poor one. As a result, many more iterations are needed (sometimes even more than a hundred, e.g. for $e > 0.95$) to get a satisfactory answer.

Newton Raphson for solving "$f(x) = a$"

The following steps can be used to solve the following general single variable nonlinear equation $f(x) = a$. The only requirement is that $f(x)$ be differentiable:

- **Step 1.** Initial guess: $x^{(0)}$.
- **Step 2.** Set $k = 1$.
  - **Step 2a.** $k - th$ correction:

$$\Delta x^{(k)} = \frac{a - f(x^{(k-1)})}{\left.\frac{df}{dx}\right|_{x^{(k-1)}}} \tag{144}$$

  - **Step 2b.** $k - th$ approximation:

$$x^{(k)} = x^{(k-1)} + \Delta x^{(k)} \tag{145}$$

- **Step 3.** Stopping criteria:

IF $\left|\Delta x^{(k)}\right| \leq \epsilon$ (threshold on correction magnitude) **OR** $k \geq N$ (limit on number of iterations for robustness),

Return $x^{(k)}$ as final approximation.

ELSE $k = k + 1$; go to Step 2a.

❣ **Example Kepler's Equation.** This example is based on the same data as that used in Sec.(10.1). The following initial conditions are known:

$$\mathbf{r}_{\mathcal{I}}(t_0) = \{-4777.8, 4862.6, 1760.1\}^{\mathrm{T}}\ km \tag{146a}$$
$$\mathbf{v}_{\mathcal{I}}(t_0) = \{-6.7782, -4.8929, 0.9174\}^{\mathrm{T}}\ kmps \tag{146b}$$

Question: Determine $\mathbf{r}_{\mathcal{I}}(t)$ and $\mathbf{v}_{\mathcal{I}}(t)$ at $t = (t_0 + P/4)$, where P is the orbital period.

In the example in Sec.(10.1), we had already worked through the integrals of motion for the initial conditions given above and obtained the following numbers:

i. semi-major axis: $a = 9378.14\ km$.

ii. eccentricity: $e = 0.3$.

iii. period: $P = 2\pi\sqrt{a^3/\mu} = 9038.4\ s = 2.51$ hours.

iv. mean motion: $n = \sqrt{\mu/a^3} = 6.9517 \times 10^{-4}\ rad/s$.

v. rotation matrix between inertial and orbital frames:

$$\mathbf{R}_{\mathcal{O}\mathcal{I}} = \begin{Bmatrix} \hat{\mathbf{i}}^{\mathrm{T}}_{e,\mathcal{I}} \\ \hat{\mathbf{i}}^{\mathrm{T}}_{y,\mathcal{I}} \\ \hat{\mathbf{i}}^{\mathrm{T}}_{h,\mathcal{I}} \end{Bmatrix} = \begin{bmatrix} 0.0148 & 0.9915 & 0.1294 \\ -0.9744 & -0.0147 & 0.2241 \\ 0.2241 & -0.1294 & 0.9659 \end{bmatrix} \tag{147}$$

vi. initial true anomaly: $f(t_0) = 45\deg = \pi/4\ rad$.

The initial eccentric anomaly can be found using Eq.(107):

$$E(t_0) \;=\; E_0 \;=\; 2\tan^{-1}\left(\sqrt{\frac{1-e}{1+e}}\tan\frac{f(0)}{2}\right) \;=\; 33.81\,\text{deg} \;=\; 0.5902\;\;rad \tag{148}$$

And the initial mean anomaly:

$$M_0 \;=\; E_0 \;-\; e\sin E_0 \;=\; 0.4232\;\;rad \tag{149}$$

Be careful and <u>use the E_0 value in radians in the above equation.</u> We can now set up the Kepler's equation for this problem:

$$E \;-\; e\,\sin E \;=\; M \;=\; M_0 + n\,\underbrace{(t-t_0)}_{P/4} \;=\; 1.9940\;\;rad \tag{150}$$

Let us solve the above Kepler's equation with a threshold of $\epsilon = 10^{-4}$ rad. Table (5) shows the evolution of the various iterations. Note that the upper limit on number of iterations is not really needed here since convergence is very fast. In fact, after only three corrections, the desired threshold is met, with $E^{(3)} = 2.2310$ rad as the final answer.

k (iteration #)	$\Delta E^{(k)}$ (rad) (correction: Eq.(139))	$E^{(k)}$ (rad) (approximation: Eq.(140))	$\lvert\Delta E^{(k)}\rvert \le \epsilon$? (stopping)	comment
0	-	1.9940	-	initial guess, $E^{(0)} = M$.
1	0.2435	2.2375	No	assumption of Eq.(132) barely valid!
2	-0.0066	2.2310	No	sudden drop in correction!
3	-4.29×10^{-6}	2.2310	YES	reached end.

Table 5: Newton Raphson Iterations for Example Problem

We can now use Eq.(107) to find $f(t)$:

$$f(t) \;=\; 2\tan^{-1}\left(\sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2}\right) \;=\; 2.4518\;\;rad \;=\; 140.48\,\text{deg} \tag{151}$$

and the equation of orbit gives:

$$r(t) \;=\; \frac{p}{1+e\cos f} \;=\; 11{,}103.71\;\;km \tag{152}$$

Recall that our objective is to find $\mathbf{r}_{\mathcal{I}}(t)$ and $\mathbf{v}_{\mathcal{I}}(t)$. We will first find $\mathbf{r}_{\mathcal{I}}(t)$ and leave $\mathbf{v}_{\mathcal{I}}(t)$ for a little later (next section). In order to get $\mathbf{r}_{\mathcal{I}}(t)$, first find $\mathbf{r}_{\mathcal{O}}(t)$ and use the rotation matrix $\mathbf{R}_{\mathcal{I}\mathcal{O}} = \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}$, i.e. $\mathbf{r}_{\mathcal{I}}(t) = \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\,\mathbf{r}_{\mathcal{O}}(t)$. There are two ways to determine $\mathbf{r}_{\mathcal{O}}(t)$ - either use the true anomaly approach (Eq.(101a)) or the eccentric anomaly approach (Eqs.(110), (246a), (246b)):

$$\mathbf{r}_{\mathcal{O}}(t) \;=\; r(t)\begin{Bmatrix}\cos f(t)\\ \sin f(t)\\ 0\end{Bmatrix} \;=\; \begin{Bmatrix}a(\cos E - e)\\ a\,\sqrt{1-e^2}\sin E\\ 0\end{Bmatrix} \tag{153}$$

Having solved the Kepler's equation, both f and E are available. Either way, we get $\mathbf{r}_{\mathcal{O}}(t) = \{-8565.16,\; 7066.15,\; 0\}^{\mathrm{T}}$ km. Using the rotation matrix from Eq.(147),

$$\mathbf{r}_{\mathcal{I}}(t) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\,\mathbf{r}_{\mathcal{O}}(t) \underset{\text{Eqs.(147),(152),(153)}}{=} \begin{Bmatrix}-7012.0\\ -8596.4\\ 475.5\end{Bmatrix}\;\; km \tag{154}$$

Determination of $\mathbf{v}_{\mathcal{I}}(t)$

- We are yet to obtain the expression for the inertial velocity vector at the current time. Return to Eq.(113) along with Eqs.(246c) and (246d) to get:

$$\mathbf{v}_{\mathcal{O}}(t) \;=\; \begin{Bmatrix} -a\sin E \\ a\sqrt{1-e^2}\cos E \\ 0 \end{Bmatrix} \dot{E} \tag{155}$$

- All we need is an expression for $\dot{E}$. But this is already available: From Eq.(119),

$$\dot{E} \;=\; \frac{n}{(1-e\cos E)} \tag{156}$$

- The above can also be obtained by differentiating the Kepler's equation (obviously since Kepler's equation was obtained by integrating the above equation!). Eq.(155) along with Eq.(156) provides the desired answer: $\mathbf{v}_{\mathcal{I}}(t) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\mathbf{v}_{\mathcal{O}}(t)$.

- It is possible to obtain more compact expression for the terms appearing above. For example, consider

$$r^2 \underset{\text{Eq.(101a)}}{=} x^2 \;+\; y^2 \tag{157a}$$

$$\underset{\text{Eqs.(246a),(246b)}}{=} a^2(\cos E - e)^2 \;+\; a^2(1-e^2)\sin^2 E \tag{157b}$$

$$= \quad a^2(1-e\cos E)^2 \tag{157c}$$

Or,

$$r \;=\; a(1-e\cos E) \tag{158}$$

Substitute for $(1-e\cos E)$ in Eq.(156),

$$\dot{E} \;=\; \frac{na}{r} \tag{159}$$

- Therefore, the expression for $\mathbf{v}_{\mathcal{O}}(t)$ becomes:

$$\mathbf{v}_{\mathcal{O}}(t) \;=\; -\frac{na^2}{r}\sin E\;\hat{\mathbf{i}}_e \;+\; \frac{na^2\sqrt{1-e^2}}{r}\cos E\;\hat{\mathbf{i}}_y \tag{160}$$

- More simplifications are possible:

$$\frac{na^2}{r} \;=\; \underbrace{\sqrt{\frac{\mu}{a^3}}}_{=n}\frac{a^2}{r} \;=\; \frac{\sqrt{\mu a}}{r} \tag{161}$$

and,

$$\frac{na^2\sqrt{1-e^2}}{r} \;=\; \sqrt{\frac{\mu}{a^3}a^4(1-e^2)}\frac{1}{r} \;=\; \sqrt{\mu a(1-e^2)}\frac{1}{r} \;=\; \sqrt{\mu p}\frac{1}{r} \;=\; \frac{h}{r} \tag{162}$$

- Putting together Eqs.(161) and (162) we get:

$$\mathbf{v}_{\mathcal{O}}(t) \;=\; -\frac{\sqrt{\mu a}}{r}\sin E\;\hat{\mathbf{i}}_e \;+\; \frac{h}{r}\cos E\;\hat{\mathbf{i}}_y \tag{163}$$

And of course, $\mathbf{v}_{\mathcal{I}}(t) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\mathbf{v}_{\mathcal{O}}(t)$.

- For the example studied above, Eq.(163) gives us: $\mathbf{v}_{\mathcal{O}}(t_0+P/4) \;=\; \{-4.3493, -3.2214, 0\}^{\mathrm{T}}$ *kmps*. And, $\mathbf{v}_{\mathcal{I}}(t_0+P/4) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\mathbf{v}_{\mathcal{O}}(t_0+P/4) \;=\; \{3.0749, -4.2647, -1.2848\}^{\mathrm{T}}$ *kmps*.

13.2 Summary: Implicit Solution of the Two Body Problem via Kepler's Equation

- We have walked through a long sequence of concepts and equations to finally solve the fundamental equation of the two body problem without the need for explicit time integration. Along the way, we specialized our solution methodology to only ellipses (and circles) through the introduction of the eccentric anomaly.
- Figure (26) summarizes the implicit method.

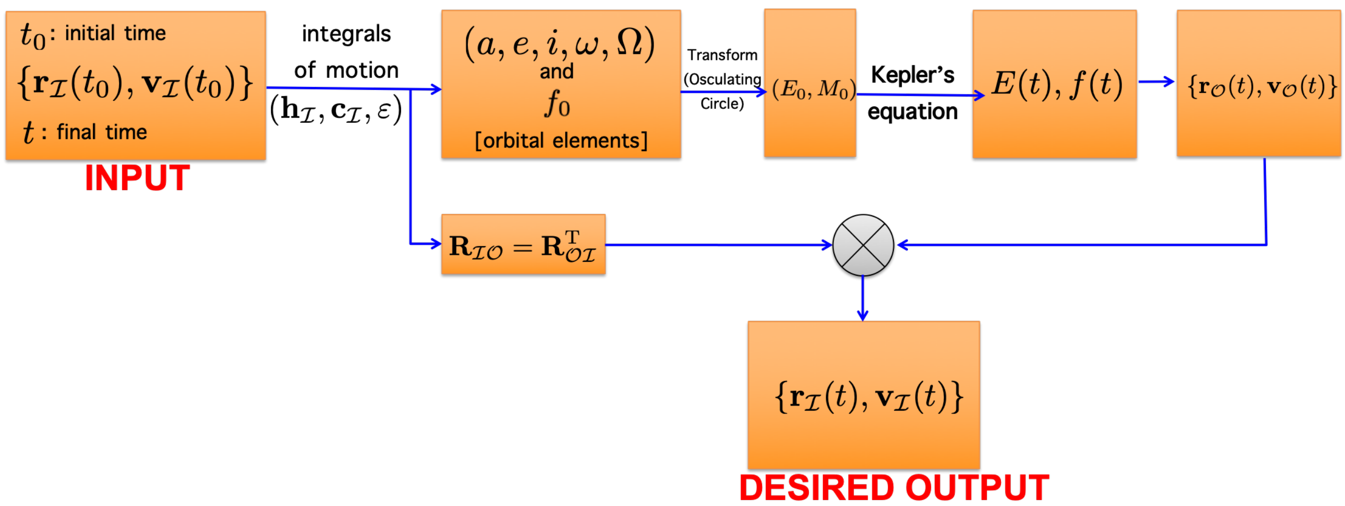


Figure 26: A Schematic of the Implicit Method

- **PROS** of the implicit method: There are many good things about the implicit method:
  i. Its elegance! Remember that the original fundamental equation is quite complex due to its non-linearity ($1/r^2$ term). However, through truly ingenious use of the *integrals of motion* and *orbital elements*, we have been able to reduce the fundamental equation (a nonlinear six dimensional ODE system) into a single nonlinear **algebraic** equation (the Kepler's equation). This is nothing short of outright beautiful.
  ii. There is **no error accumulation** in this method. The Kepler's equation represents a *direct mapping* of the initial states into the states at the final time. The mapping is nonlinear (Kepler's equation) but more importantly, does not require time-marching an in `ode45` and thereby preventing accumulation of error.
  iii. **High accuracy:** It is possible to obtain as accurate an answer as you desire: simply set a very low threshold in Eq.(143). The process may need a few extra iterations, but will give a highly accurate mapping.
- **CONS** of the implicit method: there is probably only one drawback: the method is **not universal**. As it was developed above, it applies only to circles and ellipses. We must first determine the orbit eccentricity before proceeding.
- In the next section, we consider the implicit solution of the two-body problem for open orbits (parabolas and hyperbolas).

13.3 Implicit Solution: Open Orbits

The implicit solution method described in Sec.(13) concerns only closed orbits (ellipses and circles). In this section, we will extend our technique to open orbits: i.e. parabolas and hyperbolas. Fig.(27) shows the schematic for the implicit scheme again. The shaded portion of this picture is *universal*, i.e. follows the same steps for all types of conic sections. The only non-universal step is the one that involves the time integration equation (the "bad integral") that must be solved to obtain $f(t)$. For ellipses, we defined the so called eccentric anomaly (E) to simplify the evaluation of the time integral. We will now revisit that equation and explore what can be done for parabolas and hyperbolas.

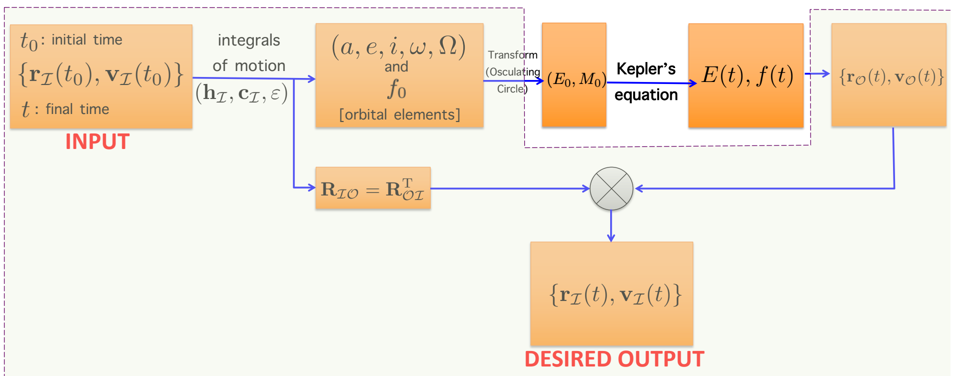


Figure 27: A Schematic of the Implicit Method: Closed Orbits Only!

13.4 Implicit Solution: Parabolas ($e = 1$)

As mentioned above, the integral equation involving time and the true anomaly was the reason that led us to specialize the implicit solution method to ellipses via the introduction of the eccentric anomaly. We revisit that integral below,

$$\int_{t_0}^{t} \frac{\sqrt{\mu}}{p^{3/2}}\, dt \;=\; \int_{f_0}^{f(t)} \frac{df}{(1 + e\cos f)^2} \tag{164}$$

This equation is the "unshaded link" in the schematic shown in Fig.(27).

- In the current context, we are looking at a very special case of the above integral, since $e = 1$. Making this substitution, we get

$$\int_{t_0}^{t} \frac{\sqrt{\mu}}{p^{3/2}} dt \;\underset{[e=1]}{=}\; \int_{f_0}^{f(t)} \frac{df}{(1 + \cos f)^2} \tag{165}$$

- Incredibly enough, the above integral (note: **only for** $e = 1$) can be evaluated analytically in closed form as follows

$$\frac{\sqrt{\mu}}{p^{3/2}}(t - t_0) \;=\; \frac{1}{2}\tan\frac{f}{2}\bigg|_{f_0}^{f(t)} + \frac{1}{6}\tan^3\frac{f}{2}\bigg|_{f_0}^{f(t)} \tag{166}$$

- To simplify the notation above, define the so-called *parabolic mean anomaly*, M_p as follows:

$$M_p \;\triangleq\; \frac{1}{2}\tan\frac{f}{2} + \frac{1}{6}\tan^3\frac{f}{2} \tag{167}$$

- So that Eq.(166) becomes

$$\boxed{M_p(t) \;=\; \frac{1}{2}\tan\frac{f}{2} \;+\; \frac{1}{6}\tan^3\frac{f}{2} \;=\; M_p(0) \;+\; \frac{\sqrt{\mu}}{p^{3/2}}(t - t_0)} \tag{168}$$

where the unknown is the true anomaly, $f(t)$. The above is the parabolic version of Kepler's equation and is widely known as the ***Barker's equation.*** The problem is the following: given t, t_0 and f_0, solve for $f(t)$.

- Just as for ellipses, we have a special case of the Barker's equation if we set $t_0 = t_p$, i.e. if the object is at the periapsis at t_0. In this case, $f(0) = 0$, thus $M_p(0) = 0$ and Eq.(168) becomes:

$$\frac{1}{2}\tan\frac{f}{2} + \frac{1}{6}\tan^3\frac{f}{2} = \frac{\sqrt{\mu}}{p^{3/2}}(t - t_p) = \frac{\sqrt{\mu}}{p^{3/2}}\tau_p \tag{169}$$

where τ_p is known as the *time since periapsis.*

- Something very interesting has happened above: note that we did not need to make a "change of variables" as for ellipses. For the special case of $e = 1$, the integral in Eq.(164) was evaluated *as is*, in terms of the true anomaly.

- We are not out of jail yet, because the algebraic equation involving the current true anomaly i.e. the Barker's equation (Eq.(168)) is nonlinear. However, the nonlinearity is not as bad as that for the elliptical case (for starters, Eq.(168) is not transcendental) and the following simple change of variable reduces Eq.(168) to a polynomial root-finding equation: $z \triangleq \tan f/2$ as follows

$$\frac{1}{2}z + \frac{1}{6}z^3 = M_p(t) = M_p(0) + \frac{\sqrt{\mu}}{p^{3/2}}(t - t_0) \tag{170}$$

solve for z.

- Clearly, the above is a cubic polynomial equation in the unknown z. The obvious problem is that three roots are anticipated and one would need to decide which of the three is the actual answer. *Luckily*, it turns out that only *one* of the three roots of Eq.(170) is real (the other two appear as an imaginary pair and are obviously inadmissible). Moreover, the unique real root can be determined analytically as

$$z = \tan\frac{f}{2} = \left[3M_p(t) + \sqrt{(3M_p(t))^2 + 1}\right]^{1/3} - \left[3M_p(t) + \sqrt{(3M_p(t))^2 + 1}\right]^{-1/3} \tag{171}$$

Amazingly, we have just found an analytically closed form solution of the two-body problem! The caveat is that it holds only for the case of the parabola. No approximations are needed and we can use the true anomaly found above to next determine our desired quantities $\mathbf{r}_{\mathcal{I}}(t)$ and $\mathbf{v}_{\mathcal{I}}(t)$.

- **Determination of $\mathbf{r}_{\mathcal{I}}(t)$ and $\mathbf{v}_{\mathcal{I}}(t)$.** To begin this process first determine the magnitude of the radius vector

$$r(t) = \frac{p}{1 + e\cos f(t)} \tag{172}$$

You must set $e = 1$ in the above equation for the present case (parabola).

- We can now obtain $\mathbf{r}_{\mathcal{O}}(t)$ as

$$\mathbf{r}_{\mathcal{O}}(t) = \begin{Bmatrix} x \\ y \\ 0 \end{Bmatrix} = r(t)\begin{Bmatrix} \cos f(t) \\ \sin f(t) \\ 0 \end{Bmatrix} \tag{173}$$

or in other words,

$$\mathbf{r}_{\mathcal{O}}(t) = r(t)\cos f(t)\,\hat{\mathbf{i}}_e + r(t)\sin f(t)\,\hat{\mathbf{i}}_y \tag{174}$$

- Now, as far as $\mathbf{v}_{\mathcal{O}}(t)$ is concerned, there is a slight problem.. we cannot determine it like we did for ellipses because there exists no such thing as an eccentric anomaly (recall that $\mathbf{v}_{\mathcal{O}}(t)$ was derived *in terms of E*). We must develop all expressions purely in terms of the true anomaly. To this end, differentiate (inertially) Eq.(174)

$$\mathbf{v}_{\mathcal{O}}(t) = \dot{\mathbf{r}}^{\mathcal{I}}_{\mathcal{O}}(t) = \left(\dot{r}\cos f - r\dot{f}\sin f\right)\hat{\mathbf{i}}_e + \left(\dot{r}\sin f + r\dot{f}\cos f\right)\hat{\mathbf{i}}_y \tag{175}$$

Keep in mind that in going from Eq.(174) to Eq.(175) we have used the fact that both vectors $\hat{\mathbf{i}}_e$ and $\hat{\mathbf{i}}_y$ are inertially non-rotating so that $\dot{\hat{\mathbf{i}}}_e = \dot{\hat{\mathbf{i}}}_y = 0$. So that now the problem reduces to finding two terms: $\dot{r}(t)$ and $\dot{f}(t)$ for use in the above expression. For this purpose we revert to the expression of kinematics in the local $\hat{\mathbf{e}}_r - \hat{\mathbf{e}}_\theta$ coordinate system ($\mathcal{E} = \{F, \hat{\mathbf{e}}_r, \hat{\mathbf{e}}_\theta, \hat{\mathbf{i}}_3\}$):

$$\mathbf{r}_\mathcal{E} = r\,\hat{\mathbf{e}}_r \tag{176a}$$
$$\mathbf{v}_\mathcal{E} = \dot{r}\,\hat{\mathbf{e}}_r + r\dot{f}\,\hat{\mathbf{e}}_\theta \tag{176b}$$

Eqs.(176) can be used to obtain the magnitude of angular momentum as $h = \|\mathbf{r}_\mathcal{E} \times \mathbf{v}_\mathcal{E}\| = r^2\dot{f}$. Thus we have

$$\dot{f}(t) = \frac{h}{r^2(t)} \tag{177}$$

Next, Eq.(176b) gives us the magnitude of the velocity vector as $v^2 = \dot{r}^2 + r^2\dot{f}^2$. This can be used in the vis-viva equation to give

$$\frac{v^2}{2} - \frac{\mu}{r} = \frac{\dot{r}^2 + r^2\dot{f}^2}{2} - \frac{\mu}{r} = \varepsilon = 0 \quad \text{(parabola)} \tag{178}$$

or,

$$\dot{r}(t) = \pm\sqrt{\frac{2\mu}{r(t)} - r^2(t)\dot{f}^2(t)} \tag{179}$$

To resolve ambiguity in the sign above, recall that for $f \in [0, \pi]$, $\dot{r} > 0$ and for $f \in [\pi, 2\pi]$, $\dot{r} < 0$. Since we have already found $f(t)$, we can determine the sign of $\dot{r}(t)$ based on its quadrant.

Eqs.(177) and (179) can be used in Eq.(175) to obtain $\mathbf{v}_\mathcal{O}(t)$.

- The final step is to use the $\mathbf{R}_{\mathcal{O}\mathcal{I}}$ matrix (obtained from the integrals of motion) to determine

$$\mathbf{r}_\mathcal{I}(t) = \mathbf{R}_{\mathcal{O}\mathcal{I}}^{\mathrm{T}}\,\mathbf{r}_\mathcal{O}(t) \tag{180a}$$
$$\mathbf{v}_\mathcal{I}(t) = \mathbf{R}_{\mathcal{O}\mathcal{I}}^{\mathrm{T}}\,\mathbf{v}_\mathcal{O}(t) \tag{180b}$$

Eqs.(180) complete the implicit solution for parabolas.

Dicuss the other expression for $\dot{r}$ involving $\sin f$ (the one that does not require disambiguation)

13.5 Implicit Solution: Hyperbolas ($e > 1$)

Before we look into the details of implicit solution of Keplerian motion for hyperbolic orbits, let us revisit the geometry of hyperbolas: refer to Fig.(28).

- The (occupied) orbit has a mirror image about the minor axis, which is vacant (unoccupied orbit) shown using a dotted line. The corresponding unoccupied focus ($F*$) is also shown.
- From the equation of orbit, $r = p/(1 + e\cos f)$. At infinity, i.e. $r = \infty$, let $f = f_\infty$. Then we must have

$$1 + e\cos f_\infty = 0$$
$$\Rightarrow f_\infty = \cos^{-1}(-1/e) \tag{181}$$

- Clearly, the radius vector at infinity must be parallel to the asymptote and from Fig.(28) it is evident that

$$\beta = \pi - f_\infty = \cos^{-1}(1/e) \tag{182}$$

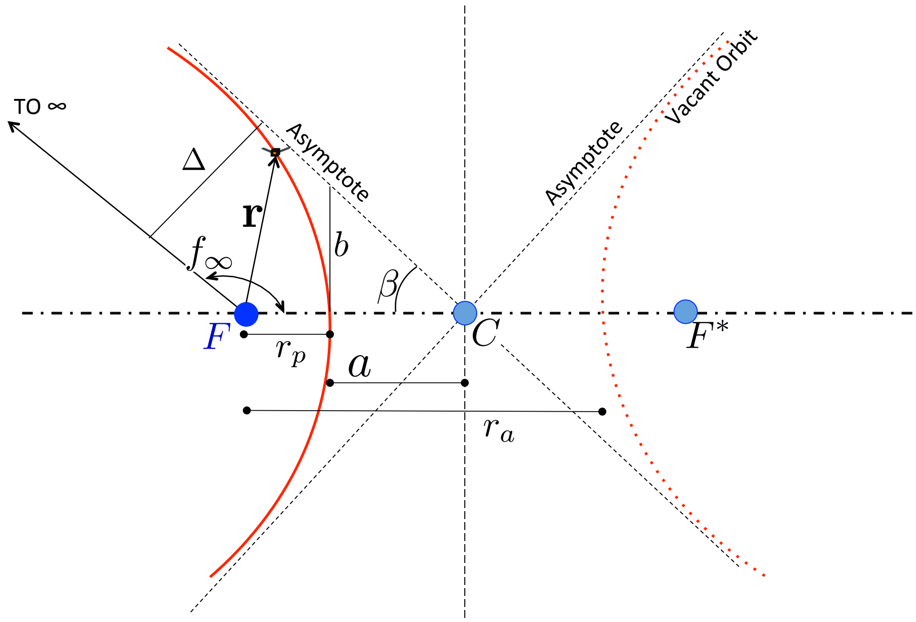


Figure 28: Geometry of a Hyperbolic Orbit

- From Fig.(28),

$$\begin{aligned} b &= a\tan\beta \\ &= a\frac{\sin\beta}{\cos\beta} \underset{\text{Eq.(182)}}{=} ae\sin(\pi - f_\infty) = ae\sin f_\infty \\ &= \frac{ae\sqrt{e^2-1}}{e} = a\sqrt{e^2-1} \end{aligned} \tag{183}$$

- Also shown in Fig.(28) is the so-called *aiming radius*, which is a useful measure in the design of interplanetary approach trajectories (it is used to maintain safe distance from the target planet).

$$\Delta = (r_p + a)\sin\beta \tag{184}$$

- Energy on a hyperbolic orbit: $\varepsilon = -\mu/2a > 0$ (since $a < 0$). Let us evaluate the vis-viva equation at $r = \infty$, i.e. $f = f_\infty$. Let the "speed at infinity" be denoted as v_∞. Then,

$$\frac{v_\infty^2}{2} - \underbrace{\frac{\mu}{\infty}}_{=0} = -\frac{\mu}{2a} \tag{185}$$

or,

$$v_\infty = \sqrt{-\frac{\mu}{a}} \tag{186}$$

The above is a meaningful number because for hyperbolas $a < 0$. The speed v_∞ is called the *hyperbolic excess speed*, and it is the residual speed an object would have upon reaching infinity.

On a parabolic orbit, the "excess speed at infinity" is zero (this is very easy to prove: simply set $\varepsilon = 0$ and $r = \infty$ to get $v(r = \infty) = 0$). In comparison, a hyperbolic orbit has a non-zero finite speed at infinity - thus the term "excess". Another way to think about this is the following - an object on a parabolic orbit *barely* reaches infinity... it does get there, but its speed reduces to zero in the process. On the other hand, the same object, if on a hyperbolic orbit would reach infinity with a finite non-zero speed, an *excess*, called the hyperbolic excess.

In terms of energy, an object on a parabolic orbit expends all its kinetic energy by the time it reaches infinity. However on a hyperbolic orbit, it still has some kinetic energy left even at infinity, represented by the hyperbolic excess.

Note that the above discussion is purely a mathematical one. In the physical world, an object would never really reach "infinity".

Eq.(185) is interesting because it presents an alternate way of writing the total energy on a hyperbolic orbit:

$$\varepsilon = -\mu/2a = v_\infty^2/2 \tag{187}$$

Therefore the vis-viva equation becomes

$$\frac{v^2(r)}{2} - \frac{\mu}{r} = \frac{v_\infty^2}{2} \tag{188}$$

Now recall that at any point r, the speed on a parabolic orbit is given by

$$\frac{v_p^2(r)}{2} = \frac{\mu}{r} \qquad (\text{since } \varepsilon_p = 0) \tag{189}$$

Therefore Eq.(188) becomes:

$$\frac{v^2(r)}{2} = \frac{v_p^2(r)}{2} + \frac{v_\infty^2}{2} \tag{190}$$

The above equation carries significant insight. We can write

$$(\text{K.E. @ } r \text{ on hyperbola}) = (\text{K.E. @ } r \text{ on parabola}) + (\text{Excess K.E.}) \tag{191}$$

Thus we have a direct measure of how much excess kinetic energy is needed (over a parabola) at any point r to be on a hyperbolic orbit of interest. This excess kinetic energy is called $C3 = v_\infty^2/2$ and is of tremendous interest in interplanetary mission design. For a spacecraft to be "injected" into a hyperbolic escape orbit from Earth, the excess kinetic energy must ultimately be provided by the launching rocket. A shorter mission (if desired) needs greater speed at the point of injection, thus the spacecraft must be injected into a hyperbolic orbit of greater energy, or equivalently greater excess KE (Eq.(187)), perhaps requiring a bigger rocket!

- We are now ready to lay out the implicit solution of Keplerian motion for hyperbolic orbits. Once again we return to the starting point: the "bad time integral" involving the true anomaly and time:

$$\frac{\sqrt{\mu}}{p^{3/2}} \int_{t_0}^{t} dt = \int_{f_0}^{f(t)} \frac{df}{(1 + e\cos f)^2} \tag{192}$$

Since the above integral was hard to evaluate, we adopted a change of variables ($f \to E$) that was valid only for ellipses. For $e > 1$ (i.e. for hyperbolas), the above integral assumes the following form,

$$\frac{\sqrt{\mu}}{p^{3/2}} (t - t_0)(e^2 - 1)^{3/2} = \left[\frac{e\sqrt{e^2 - 1}}{1 + e\cos f} \sin f - \ln\left(\frac{\sqrt{e+1} + \sqrt{e-1}\tan\frac{f}{2}}{\sqrt{e+1} - \sqrt{e-1}\tan\frac{f}{2}} \right) \right] \Big|_{f_0}^{f(t)} \tag{193}$$

The above equation is horrendously nonlinear in the unknown ($f(t)$). In order to begin the simplification process, a change of variables is proposed (like in the case of the ellipse), involving the so-called *hyperbolic eccentric anomaly*, F defined as follows:

$$\sinh F = \frac{y}{b} \tag{194}$$

where sinh(·) is the *hyperbolic sine* function defined as

$$\sinh x \;=\; \frac{e^x - e^{-x}}{2} \tag{195}$$

Also in Eq.(194), y is the $\hat{\mathbf{i}}_{y,\mathcal{I}}$ coordinate of the position vector, shown in Fig.(29).

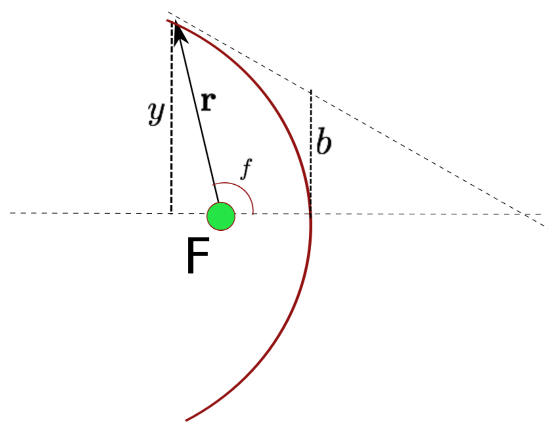


Figure 29: Hyperbolic Eccentric Anomaly

Now, $y = r\sin f$, $r = a(1-e^2)/(1+e\cos f)$ and $b = a\sqrt{e^2-1}$. Therefore

$$\begin{aligned}\sinh F &= \frac{a(1-e^2)}{(1+e\cos f)}\frac{\sin f}{a\sqrt{e^2-1}} \\ &= \frac{\sqrt{e^2-1}\sin f}{(1+e\cos f)} \end{aligned}\tag{196}$$

It is also possible to derive (derivation can be found in book) the following two astonishing relationships between f and F:

$$F \;=\; \ln\left(\frac{\sqrt{e+1}+\sqrt{e-1}\tan\frac{f}{2}}{\sqrt{e+1}-\sqrt{e-1}\tan\frac{f}{2}}\right) \tag{197}$$

$$\tanh\frac{F}{2} \;=\; \sqrt{\frac{e-1}{e+1}}\tan\frac{f}{2} \tag{198}$$

But of course, there is a problem of inversion: note what happens when the true anomaly is (even moderately) close to π. I think the sinh relationship above must be used for inversion. **Actually not... the true anomaly is not allowed to get close to π for hyperbolas! It must be less than f_∞, where the limit is reached. f_∞ is always less than π since there is no such thing as a apoapsis.**

Using Eqs.(197) and (196) in Eq.(193), we obtain

$$\frac{\sqrt{\mu}}{p^{3/2}}(t-t_0)(e^2-1)^{3/2} \;=\; (e\sinh F \;-\; F) \;-\; (e\sinh F_0 \;-\; F_0) \tag{199}$$

And now, introduce the so called *mean hyperbolic anomaly*, M_h as

$$M_h(t) \;=\; e\sinh F(t) \;-\; F(t) \tag{200}$$

So that Eq.(199) can be written as

$$\boxed{M_h(t) \;=\; e\sinh F(t) \;-\; F(t) \;=\; M_h(0) \;+\; \frac{\sqrt{\mu}}{p^{3/2}}\,(t-t_0)\,(e^2-1)^{3/2}} \tag{201}$$

which is known as the ***hyperbolic Kepler's equation*** with unknown $F(t)$ and given t, t_0 and F_0. Just like for ellipses, the above hyperbolic version of Kepler's equation must be solved numerically to obtain $F(t)$, which in turn can be used in Eq.(198) to determine the current true anomaly, $f(t)$. The algorithm below can be developed exactly like for the elliptical case.

ALGORITHM: Newton Raphson for Hyperbolic Kepler's Equation

- **Step 1.** Initial guess: $F^{(0)}$. Looking at Eq.(201), $F^{(0)} \;=\; M_h(t)$ is a good guess!
- **Step 2.** Set $k = 1$.
  - **Step 2a.** $k-th$ correction: Get $\Delta F^{(k)}$ as follows

$$\Delta F^{(k)} \;=\; \frac{M_h(t) - (e\sinh F^{(k-1)} - F^{(k-1)})}{e\cosh F^{(k-1)} - 1} \tag{202}$$

  - **Step 2b.** $k-th$ approximation: $F^{(k)} \;=\; F^{(k-1)} \;+\; \Delta F^{(k)}$.
- **Step 3.** Check for stopping criteria:

IF $\left|\Delta F^{(k)}\right| \leq \epsilon$ (correction magnitude less than set threshold) OR $k \geq N$ (too many iterations!);

Return $F^{(k)}$ as the final approximation.

ELSE $k \;=\; k+1$ and return to step 2a.

algorithm end.

Discussion on initial guesses (e.g. what does $F^{(0)} \;=\; 0$ correspond to and what does $F^{(0} = M_h(t)$ correspond to (circle) and uniqueness of solution. Use i.) Barkers solution as initial guess and ii.) F_∞ as initial guess.

Can we converge to the Barkers' equation by setting $e \downarrow 1$ to the hyperbola solution?

- Using the obtained approximation for the hyperbolic eccentric anomaly, we next obtain the true anomaly at the current time, i.e. $f(t)$ using Eq.(198). Once $f(t)$ is available, follow the same steps as for the parabola to obtain $\mathbf{r}_\mathcal{I}(t)$ and $\mathbf{v}_\mathcal{I}(t)$ (see Eqs.(172) - (180) above). The example problem below outlines in detail the implicit solution procedure.

❣ **Example Hyperbolic Keplerian Motion.** It is given that time t_0, an object has the following geocentric position and velocity vectors expressed in the inertial reference frame:

$$\mathbf{r}_\mathcal{I}(0) \;=\; \{-6.9786,\, 5.7203,\, 4.7745\}^{\mathrm{T}} \times 10^6\; m \tag{203a}$$

$$\mathbf{v}_\mathcal{I}(0) \;=\; \{-7.4157,\, -6.5515,\, 0.3249\}^{\mathrm{T}} \times 10^3\; m/s \tag{203b}$$

Find $\mathbf{r}_\mathcal{I}(t)$ and $\mathbf{v}_\mathcal{I}(t)$ at $t = t_0 + 1$ hr.

Solution. We follow the step-by-step approach of the implicit solution method below:

I. First some pre-computations:

$$r_0 = \|\mathbf{r}_{\mathcal{I}}(0)\| = 1.0209 \times 10^6\ m \tag{204a}$$
$$v_0 = \|\mathbf{v}_{\mathcal{I}}(0)\| = 9.9005 \times 10^3\ m/s \tag{204b}$$

II. Compute integrals of motion: there are three of them as follows:

a. **angular momentum vector.**

$$\mathbf{h}_{\mathcal{I}} = \mathbf{r}_{\mathcal{I}}(0) \times \mathbf{v}_{\mathcal{I}}(0) = \{3.3139,\ -3.3139,\ 8.8140\}^{\mathrm{T}} \times 10^{10}\ m^2/s \tag{205}$$

and its magnitude,

$$h = \|\mathbf{h}_{\mathcal{I}}\| = 9.9825 \times 10^{10}\ m^2/s \tag{206}$$

finally, the unit vector along $\mathbf{h}$, expressed in the $\mathcal{I}$ frame:

$$\hat{\mathbf{i}}_{h,\mathcal{I}} = \frac{\mathbf{h}_{\mathcal{I}}}{h} = \{0.33197,\ -0.33197,\ 0.88259\}^{\mathrm{T}} \tag{207}$$

b. **eccentricity vector.**

$$\mathbf{c}_{\mathcal{I}} = \mathbf{v}_{\mathcal{I}}(0) \times \mathbf{h}_{\mathcal{I}} - \frac{\mu}{r_0}\mathbf{r}_{\mathcal{I}}(0) = \{-2.9421,\ 4.4104, 2.7643\}^{\mathrm{T}} \times 10^{14}\ m^3/s \tag{208}$$

it magnitude,

$$c = \|\mathbf{c}_{\mathcal{I}}\| = 5.979 \times 10^{14}\ m^3/s \tag{209}$$

unit vector along $\mathbf{c}$, expressed in $\mathcal{I}$,

$$\hat{\mathbf{i}}_{e,\mathcal{I}} = \frac{\mathbf{c}_{\mathcal{I}}}{c} = \{-0.4921,\ 0.7376,\ 0.4623\}^{\mathrm{T}} \tag{210}$$

We can also determine the unit vector $\hat{\mathbf{i}}_{y,\mathcal{I}}$ to complete the triad $\mathcal{O} = \{F, \hat{\mathbf{i}}_e, \hat{\mathbf{i}}_y, \hat{\mathbf{i}}_h\}$ as follows

$$\hat{\mathbf{i}}_{y,\mathcal{I}} = \frac{\hat{\mathbf{i}}_{h,\mathcal{I}} \times \hat{\mathbf{i}}_{e,\mathcal{I}}}{\|\hat{\mathbf{i}}_{h,\mathcal{I}} \times \hat{\mathbf{i}}_{e,\mathcal{I}}\|} = \{-0.8047,\ -0.5880,\ 0.0815\}^{\mathrm{T}} \tag{211}$$

c. **energy.**

$$\varepsilon = \frac{v_0^2}{2} - \frac{\mu}{r_0} = +9.9650 \times 10^6\ m^2/s^2 \tag{212}$$

(first indication of a hyperbolic orbit: positive energy)

III. Compute orbital elements and $\mathbf{R}_{\mathcal{O}\mathcal{I}}$.

a. **semi-major axis.**

$$a = -\frac{\mu}{2\varepsilon} = -2.000 \times 10^7\ km \quad \text{[hyperbola!]} \tag{213}$$

b. **eccentricity.**

$$e = \frac{c}{\mu} = 1.5 \quad \text{[hyperbola!]} \tag{214}$$

c. **inclination.**

$$i = \cos^{-1} i_{h_3} = 0.4887\ rad \sim 28\,\mathrm{deg} \tag{215}$$

d. **longitude of ascending node.**

$$\Omega \mapsto \begin{cases} c\Omega = -i_{h_2}/si = 0.7071 \\ s\Omega = i_{h_1}/si = 0.7071 \end{cases} ; \quad \Rightarrow \Omega = 0.7854\ rad \sim 45\,\mathrm{deg} \tag{216}$$

e. **argument of periapsis.**

$$\omega \;\mapsto \begin{cases} c\omega \;=\; i_{y_3}/si = 0.1736 \\ s\omega \;=\; i_{e_3}/si = 0.9849 \end{cases} \quad ; \quad \Rightarrow \omega \;=\; 1.3963\, rad \;\sim\; 80\,\text{deg} \tag{217}$$

f. **true anomaly @ t_0.** We use method 3 as described in Sec.(10.1). First determine $\cos f_0$ using the equation of orbit,

$$\cos f_0 \;=\; \frac{1}{e}\left(\frac{a(1-e^2)}{r_0} - 1\right) \;=\; 0.9659 \tag{218}$$

Next isolate the correct quadrant by determining the sign of $\dot{r}(0)$,

$$\text{sgn}(\dot{r}(0)) \;=\; \text{sgn}(\mathbf{r}(0)\cdot\mathbf{v}(0)) \underset{\text{Eq.(203)}}{=} +1 \tag{219}$$

Therefore, $f_0 \in [0, \pi]$. Using Eq.(218), $f_0 \;=\; 15\,\text{deg}$.

g. **rotation matrix $\mathbf{R}_{\mathcal{O}\mathcal{I}}$.** We have

$$\mathbf{R}_{\mathcal{O}\mathcal{I}} \;=\; \begin{Bmatrix} \hat{\mathbf{i}}^{\mathrm{T}}_{e,\mathcal{I}} \\ \\ \hat{\mathbf{i}}^{\mathrm{T}}_{y,\mathcal{I}} \\ \\ \hat{\mathbf{i}}^{\mathrm{T}}_{h,\mathcal{I}} \end{Bmatrix} \;=\; \begin{bmatrix} -0.4921 & 0.7376 & 0.4623 \\ \\ -0.8047 & -0.5880 & 0.0815 \\ \\ 0.3320 & -0.3320 & 0.8826 \end{bmatrix} \tag{220}$$

IV. Solve hyperbolic Kepler's equation for current angular position in orbit. First determine the hyperbolic eccentric anomaly at t_0

$$F_0 \;=\; 2\tanh^{-1}\left(\sqrt{\frac{e-1}{e+1}}\tan\frac{f_0}{2}\right) \;=\; 0.11789\, rad \tag{221}$$

Thus the mean hyperbolic anomaly at t_0 is

$$M_h(t_0) \;=\; e\sinh F_0 - F_0 \;=\; 0.059355\, rad \tag{222}$$

Now, $t = t_0 + 3600$. Thus the hyperbolic Kepler's equation becomes:

$$M_h(t) \;=\; \underbrace{M_h(0) \;+\; \frac{\sqrt{\mu}}{p^{3/2}}(t-t_0)(e^2-1)^{3/2}}_{0.8629} \;=\; e\sinh F \;-\; F \tag{223}$$

Setting a threshold of $\epsilon \;=\; 10^{-6}\, rad$, we obtain $\Delta F^{(5)} \;=\; -3.1753\times 10^{-7}$, which meets the required threshold. The corresponding approximation is

$$F(t) \approx F^{(5)} \;=\; 1.0725\, rad \tag{224}$$

Thus the current true anomaly is

$$f \;=\; 2\tan^{-1}\left(\sqrt{\frac{e+1}{e-1}}\tanh\frac{F}{2}\right) \;=\; 1.6624\, rad \;\sim\; 95.3\,\text{deg} \tag{225}$$

V. Determine $\mathbf{r}_{\mathcal{I}}(t)$ and $\mathbf{v}_{\mathcal{I}}(t)$. First we use the equation of orbit to determine $r(t)$

$$r(t) \;=\; \frac{a(1-e^2)}{1+e\cos f} \;=\; 2.8973\times 10^7\; m \tag{226}$$

It is interesting to note that the radius magnitude of the object has grown by a full order of magnitude in just a single hour of flight (compare Eqs.(204a) and (226))! Then, use vis-viva to get $v(t)$

$$v(t) \;=\; \sqrt{2\left(-\frac{\mu}{2a}+\frac{\mu}{r(t)}\right)} \;=\; 6.888\times 10^{3}\ m/s \tag{227}$$

In accordance with Kepler's second law, the object has slowed down, as it should with increasing distance from the focus. We can now obtain $\mathbf{r}_{\mathcal{O}}(t)$

$$\mathbf{r}_{\mathcal{O}}(t) \;=\; r(t)\begin{Bmatrix}\cos f\\ \sin f\\ 0\end{Bmatrix} \;=\; \begin{Bmatrix}-0.26489\\ 2.8852\\ 0\end{Bmatrix}\times 10^{7}\ m \tag{228}$$

Also,

$$\mathbf{v}_{\mathcal{O}}(t) \;=\; \begin{Bmatrix}\dot r\cos f - r\dot f\sin f\\ \dot r\sin f + r\dot f\cos f\\ 0\end{Bmatrix} \tag{229}$$

where $\dot f \;=\; h/r^2 \;=\; 1.1892\times 10^{-4}\ rad/s$. Also, $\dot r \;=\; \pm\sqrt{v(t)^2 \;-\; r^2\dot f^2} \;=\; \pm 6.7802\times 10^{3}\ m/s$. Since $f\sim 95$ deg (object is *above* the semi-major axis), $\dot r > 0 \Rightarrow \dot r \;=\; +6.7802\times 10^{3}\ m/s$. Thus

$$\mathbf{v}_{\mathcal{O}}(t) \;=\; \begin{Bmatrix}-4.0509\\ 6.4368\\ 0\end{Bmatrix}\times 10^{3}\ m/s \tag{230}$$

Finally,

$$\mathbf{r}_{\mathcal{I}}(t) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\mathbf{r}_{\mathcal{O}}(t) \;=\; \begin{Bmatrix}-2.1916\\ -1.8917\\ 0.11274\end{Bmatrix}\times 10^{7}\ m \tag{231}$$

and,

$$\mathbf{v}_{\mathcal{I}}(t) \;=\; \mathbf{R}^{\mathrm{T}}_{\mathcal{O}\mathcal{I}}\mathbf{v}_{\mathcal{O}}(t) \;=\; \begin{Bmatrix}-3.1869\\ -6.7726\\ -1.3481\end{Bmatrix}\times 10^{3}\ m/s \tag{232}$$

14 F & G Solution of Keplerian Motion

So far, we have considered two methods of solution for the two body problem, namely explicit numerical integration and implicit solution via the Kepler's equation. The main advantage of explicit integration is its straightforward approach and universality in terms of applicability to all types of conic sections. Its disadvantage however is that it must be used with stringent error control given the typically long time-durations of integration involved, coupled with nonlinearity in the governing dynamics (i.e. fundamental equation of Keplerian motion). It is essentially a time marching procedure that can lead to significant accumulation of error.

On the other hand, the implicit method is based on the intergals of motion of the two body problem, and (quiet elegantly) though the use of 'orbital elements,' reduces the original six dimensional nonlinear ODE system to a *single* algebraic equation (Kepler's equation, or Barker's equation in case of parabolas). This leads to a direct mapping of the initial states $(\mathbf{r}_0, \mathbf{v}_0)$ into the current states $(\mathbf{r}(t), \mathbf{v}(t))$ without the need for time-marching. There is no associated error accumulation and accuracy can be controlled very precisely. The disadvantage is that this approach is not universal and we must use a different route for the different conic sections.

In this part of the notes, we will consider the so-called $F-G$ solution (a.k.a $F-G$ mapping) of the two body-problem, which in essence is also an implicit solution scheme.

- The **key idea** behind $F - G$ mapping is the following:

  A. Two-body motion is planar, i.e. the plane containing the position ($\mathbf{r}$) and velocity ($\mathbf{v}$) vectors is time-invariant, defined by the angular momentum vector $\mathbf{h} = \mathbf{r} \times \dot{\mathbf{r}}$.

  B. Any two non-coincidental vectors in the $(\mathbf{r}, \mathbf{v})$ orbital plane can be used as *basis vectors* to represent a general vector in the orbital plane. It is not required for these basis vectors to be orthogonal so long as they are not coincidental (i.e. lie along the same line). This fact comes from the basic fundamentals of linear algebra.

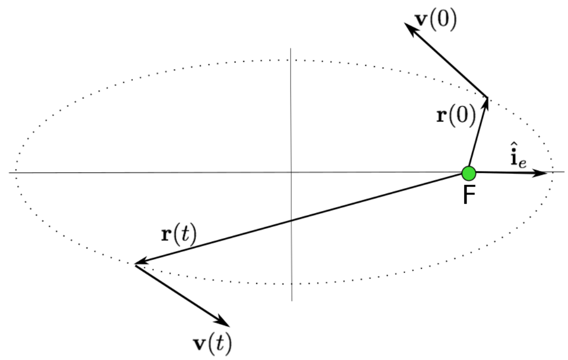


Figure 30: Underlying Logic: $F - G$ mapping

- Fig.(30) presents an illustration: the radius and velocity vectors at the initial and current times are shown. By virtue of time-invariance of the angular momentum vector, both these pairs lie in the same plane. Note that the $(\mathbf{r}(t), \mathbf{v}(t))$ pair is not coincidental at any time t; otherwise we would have a zero angular momentum vector, $\mathbf{h} = \mathbf{r} \times \mathbf{v}$, which is nonsensical.

  Therefore, following the above discussion, the pair $(\mathbf{r}(0), \mathbf{v}(0))$ can be used as a basis for representing the position and velocity vectors at all other times. In particular, we can write

$$\mathbf{r}(t) = F\mathbf{r}(0) + G\mathbf{v}(0) \tag{233}$$

  where, F and G are scalars, and may be identified as *coordinates* (or components) of the vector $\mathbf{r}(t)$ along the basis vectors $\mathbf{r}(0)$ and $\mathbf{v}(0)$. The above relationship is the "big-idea": you must first realize that $\mathbf{r}(0)$ and $\mathbf{v}(0)$ together form a valid basis for all vectors in the orbital plane; then Eq.(233) follows from basic linear algebra. Differentiating Eq.(233) (inertially) with respect to time, we get

$$\mathbf{v}(t) = \dot{F}\mathbf{r}(0) + \dot{G}\mathbf{v}(0) \tag{234}$$

  Note that in going from Eq.(233) to Eq.(234), we exploited the fact that $\mathbf{r}(0)$ and $\mathbf{v}(0)$ are *known* initial conditions and thus constants. Clearly, $\dot{F}$ and $\dot{G}$ are the coordinates of the vector $\mathbf{v}(t)$ in the $(\mathbf{r}(0), \mathbf{v}(0))$ basis.

- Equations (233) and (234) together form the template for the $F - G$ solution of Keplerian motion. The unknowns (to be determined) are the four scalar coordinates F, G, $\dot{F}$ and $\dot{G}$.

- Please note that while Fig.(30) was drawn assuming the underlying orbit to be elliptical, both facts (A) and (B) listed above in page 1 are valid for two-body motion *in general* and thus represent a universal solution approach. Simply put, Eqs.(233) and (234) hold for all types of conic sections.

- At this point, it is worthwhile to pause and ask - why are we doing all this, or; why is the above logic, even though acceptable and actually quite graceful, a good way to solve the two-body problem?

  First of all, observe that in comparison to the implicit method that we studied in great detail, the $F-G$ approach appears to be significantly more burdensome - in the former we reduced the problem to one of solving a nonlinear algebraic equation for a *single* scalar, namely the eccentric anomaly, true anomaly or hyperbolic eccentric anomaly (for ellipses, parabolas and hyperbolas respectively). However, in the proposed $F-G$ method, we must solve for four scalar unknowns: F, G, $\dot{F}$ and $\dot{G}$.

  At the same time, recall that researchers were in the quest for finding an exact solution to the two body problem. The implicit method *almost* pulled the rabbit out of the hat, hitting a bottleneck at the Kepler's equation. If it were possible to somehow obtain exact expressions for the scalars in the $F-G$ approach, increasing the burden from one to four would be completely justified!

  Even if we suppose that no exact solutions may be found, the $F-G$ approach reduces the number unknowns from six in the original fundamental equation to four, which is an improvement over the explicit scheme. Therefore at face value, the newly proposed $F-G$ method appears to fall somewhere in between explicit numerical integration (six nonlinear ODEs) and the implicit method involving Kepler's equation (one nonlinear algebraic equation).

- Let us now try and determine expressions for the unknown scalar variables F, G, $\dot{F}$ and $\dot{G}$. Clearly, these are functions of time:

$$F = F(t); \quad \dot{F} = \dot{F}(t); \tag{235a}$$

$$G = G(t); \quad \dot{G} = \dot{G}(t) \tag{235b}$$

- The initial conditions for the above scalars can be obtained by evaluating Eqs.(233) and (234) at $t = t_0$:

$$\mathbf{r}(0) = F_0\mathbf{r}(0) + G_0\mathbf{v}(0) \tag{236a}$$

$$\mathbf{v}(0) = \dot{F}_0\mathbf{r}(0) + \dot{G}(0)\mathbf{v}(0) \tag{236b}$$

  Simply comparing coefficients on the left and right hand sides of the above two equations, we get

$$F_0 = 1; \quad G_0 = 0 \tag{237a}$$

$$\dot{F}_0 = 0; \quad \dot{G}_0 = 1 \tag{237b}$$

- It is also not very hard to show that F and G follow a fundamental-equation like ODE dynamics:

$$\ddot{F} = -\frac{\mu}{r^3}F \tag{238a}$$

$$\ddot{G} = -\frac{\mu}{r^3}G \tag{238b}$$

- F and G are sometimes called Lagrange coefficients.

14.1 Solution for F and G

We are in the quest for the four scalar time-varying unknowns described above. It turns out that given the initial conditions $(\mathbf{r}(0), \mathbf{v}(0))$, it is possible to obtain an algebraic mapping to the current values of F and G, i.e. $F(t)$ and $G(t)$. In other words, we will *not* need to solve an ordinary differential equation. This is the good news. The bad news is that there is no escaping the inherent nonlinearity in the two-body dynamics and we will still need to solve the Kepler's equation.

- To begin, recall our expression for the position and velocity vectors in the orbital plane:

$$\mathbf{r}_{\mathcal{O}}(t) = x\,\hat{\mathbf{i}}_e + y\,\hat{\mathbf{i}}_y \tag{239a}$$
$$\mathbf{v}_{\mathcal{O}}(t) = \dot{x}\,\hat{\mathbf{i}}_e + \dot{y}\,\hat{\mathbf{i}}_y \tag{239b}$$

Evaluating the above expressions at t_0,

$$\mathbf{r}_{\mathcal{O}}(t_0) = x_0\hat{\mathbf{i}}_e + y_0\hat{\mathbf{i}}_y \tag{240a}$$
$$\mathbf{v}_{\mathcal{O}}(t_0) = \dot{x}_0\hat{\mathbf{i}}_e + \dot{y}_0\hat{\mathbf{i}}_y \tag{240b}$$

Substituting Eqs.(239) and (240) in Eq.(233), we get

$$\begin{Bmatrix} x \\ y \end{Bmatrix} = F\begin{Bmatrix} x_0 \\ y_0 \end{Bmatrix} + G\begin{Bmatrix} \dot{x}_0 \\ \dot{y}_0 \end{Bmatrix} \tag{241}$$

Essentially what we did above is that we expressed Eq.(233) in the orbital reference frame and used our previously used coordinatizations of the $\mathbf{r}$ and $\mathbf{v}$ vectors in this frame (Eqs.(239)). Since we are working in the orbital reference frame, the third coordinate (out of plane) is irrelevant and we have dropped it in the above Eq.(241) (the third component involves a trivial $0 = 0$ type of equation) .

Separating out the two algebraic equations above,

$$x = F\,x_0 + G\,\dot{x}_0 \tag{242a}$$
$$y = F\,y_0 + G\,\dot{y}_0 \tag{242b}$$

Keep in mind that we are trying to solve for the coefficients F and G. Therefore, the above two equations can be re-written in matrix form as

$$\begin{Bmatrix} x \\ y \end{Bmatrix} = \underbrace{\begin{bmatrix} x_0 & y_0 \\ \dot{x}_0 & \dot{y}_0 \end{bmatrix}}_{\boldsymbol{T}}\begin{Bmatrix} F \\ G \end{Bmatrix} \tag{243}$$

where the matrix appearing above is identified as the transformation matrix $\boldsymbol{T}$, which when inverted, leads to the following expressions for F and G as desired:

$$\begin{Bmatrix} F \\ G \end{Bmatrix} = \frac{1}{(x_0\dot{y}_0 - \dot{x}_0 y_0)}\begin{bmatrix} \dot{y}_0 & -\dot{x}_0 \\ -y_0 & x_0 \end{bmatrix}\begin{Bmatrix} x \\ y \end{Bmatrix} \tag{244}$$

The denominator in the above equation does not look too good, until you realize that $h = \|\mathbf{h}\| = \|\mathbf{h}_{\mathcal{O}}\| = \|\mathbf{r}_{\mathcal{O}}(0) \times \mathbf{v}_{\mathcal{O}}(0)\| \underset{\text{Eq.(240)}}{=} (x_0\dot{y}_0 - \dot{x}_0 y_0)$! Therefore Eq.(243) opens up to give

$$F = \frac{1}{h}\left(\dot{y}_0\,x - \dot{x}_0\,y\right) \tag{245a}$$
$$G = \frac{1}{h}\left(-y_0\,x + x_0\,y\right) \tag{245b}$$

Of course, the above is not the final answer because both x and y are unknown (we only know the initial conditions x_0, y_0, $\dot{x}_0$, $\dot{y}_0$). Unfortunately, this is the point where the F and G approach also hits a bottleneck, because we need expressions for the current states x and y and for this, the Kepler's equation must be solved...

- Eqs.(245) derived above are universal, i.e. valid for all conic types. However, in our study of the implicit solution, we have seen that the Kepler's equation is valid only for ellipses. For parabolas the Barker's equation and for hyperbolas the hyperbolic Kepler's equation must be solved to obtain (x, y). Therefore, from this point on, we must specialize our analysis in order to treat various types of two-body orbits.

14.2 $F - G$ solution for ellipses.

- So far, we have obtained expressions for F and G in terms of the current states (x, y) and initial states $(x_0, y_0, \dot{x}_0, \dot{y}_0)$. For ellipses, we have already come across the following expressions for (x, y) in terms of the eccentric anomaly, E:

$$x = a(\cos E - e) \tag{246a}$$
$$y = a\sqrt{1-e^2}\sin E = b\sin E \tag{246b}$$

Also,

$$\dot{x} = \frac{-\sqrt{\mu a}}{r}\sin E \tag{246c}$$
$$\dot{y} = \frac{h}{r}\cos E \tag{246d}$$

Substitute the above expressions into Eqs.(245) to obtain

$$F = \frac{1}{h}\left[\left(\frac{h}{r_0}\cos E_0\right)(a(\cos E - e)) + \left(\frac{\sqrt{\mu a}}{r_0}\sin E_0\right)\left(a\sqrt{1-e^2}\sin E\right)\right] \tag{247a}$$

$$= \frac{1}{h}\left[\frac{ah}{r_0}\cos E_0(\cos E - e) + \frac{a\sqrt{\mu a(1-e^2)}}{r_0}\sin E_0 \sin E\right] \tag{247b}$$

$$= \frac{1}{h}\left[\frac{ah}{r_0}\cos E_0(\cos E - e) + \frac{ah}{r_0}\sin E_0 \sin E\right] \tag{247c}$$

$$= \frac{a}{r_0}\left[\underbrace{\cos E_0 \cos E + \sin E_0 \sin E}_{\cos(E-E_0)} - e\cos E_0\right] \tag{247d}$$

$$= \frac{a}{r_0}\left[\cos(E - E_0) - e\cos E_0\right] \tag{247e}$$

Of course, the current eccentric anomaly (E) must be obtained by solving the Kepler's equation.

Now, define

$$\widehat{E} = E - E_0 \tag{248}$$

Also recall that

$$r = a(1 - e\cos E) \tag{249}$$

Therefore, $r_0 = a(1 - e\cos E_0)$ and

$$e\cos E_0 = (1 - r_0/a) \tag{250}$$

Using Eq.(250) in Eq.(247e),

$$F = \frac{a}{r_0}\left[\cos\widehat{E} - \left(1 - \frac{r_0}{a}\right)\right] \tag{251}$$

Or,

$$\boxed{F = 1 - \frac{a}{r_0}\left(1 - \cos\widehat{E}\right)} \tag{252}$$

where r_0 is known from the initial conditions and $\widehat{E} = (E - E_0)$ is determined by solving the Kepler's equation.

- Similarly for G we have from Eq.(245b),

$$\begin{aligned}
G &= \frac{1}{h}\left(-y_0\, x + x_0\, y\right) \\
&= \frac{1}{h}\left[\left(-a\sqrt{1-e^2}\sin E_0\right)\left(a(\cos E - e)\right) \;+\; \left(a(\cos E_0 - e)\right)\left(a\sqrt{1-e^2}\sin E\right)\right] && (253a)\\
&= \frac{a^2\sqrt{1-e^2}}{h}\left[\underbrace{-\sin E_0\cos E + \sin E\cos E_0}_{\sin(E-E_0)} - e(\sin E - \sin E_0)\right] && (253b)\\
&= \frac{a^2\sqrt{1-e^2}}{\sqrt{\mu a(1-e^2)}}\left[\sin(E-E_0) - e(\sin E - \sin E_0)\right] && (253c)\\
&= \sqrt{\frac{a^3}{\mu}}\left[\sin\widehat{E} - e(\sin E - \sin E_0)\right] && (253d)
\end{aligned}$$

Now recall that $n = 2\pi/P = \sqrt{\mu/a^3}$, $M = E - e\sin E$ and $M_0 = E_0 - e\sin E_0$. Also,

$$M - M_0 \;=\; n(t-t_0) \;=\; \underbrace{(E-E_0)}_{\widehat{E}} - e(\sin E - \sin E_0) \tag{254}$$

thereby, $e(\sin E - \sin E_0) \;=\; \widehat{E} - n(t-t_0)$. Using Eq.(254) in Eq.(253d),

$$\begin{aligned}
G &= \frac{1}{n}\left[\sin\widehat{E} \;-\; \widehat{E} \;+\; n(t-t_0)\right] && (255a)\\
&= (t-t_0) \;+\; \frac{1}{n}\left(\sin\widehat{E} \;-\; \widehat{E}\right) && (255b)
\end{aligned}$$

Or,

$$\boxed{G \;=\; (t-t_0) \;+\; \sqrt{\frac{a^3}{\mu}}\left(\sin\widehat{E} \;-\; \widehat{E}\right)} \tag{256}$$

It can also be shown that in terms of σ (defined in Eq.(108)), the above relationship can be written as:

$$\boxed{G \;=\; \frac{a\sigma_0}{\sqrt{\mu}}\left(1 - \cos\widehat{E}\right) \;+\; r_0\sqrt{\frac{a}{\mu}}\sin\widehat{E}} \tag{257}$$

where, $\sigma_0 \;=\; \mathbf{r}(0)\cdot\mathbf{v}(0)/\sqrt{\mu}$ (following Eq.(108)).

- To finish, we need the expressions for $\dot{F}$ and $\dot{G}$ for use in Eq.(234). These can be obtained by differentiating Eqs.(252) and (256) respectively. First differentiate Eq.(252) to get

$$\dot{F} \underset{\text{Eq.(252)}}{=} -\frac{a}{r_0}\sin\widehat{E}\;\dot{\widehat{E}} \tag{258}$$

but, $\dot{\widehat{E}} \;=\; (\dot{E} \;-\; \dot{E}_0) \;=\; \dot{E}$. Therefore,

$$\dot{\widehat{E}} \;=\; \dot{E} \;=\; na/r \qquad \text{(See Eq.(159))} \tag{259}$$

Thus, $\dot{F} = -na^2 \sin \widehat{E}/(rr_0)$, or,

$$\boxed{\dot{F} = -\frac{\sqrt{\mu a}}{rr_0} \sin \widehat{E}} \tag{260}$$

where, $r = a(1 - e\cos E)$ (current radius magnitude). Purely in terms of the variable $\widehat{E}$, it can be shown (without proof here) that

$$r = a(1 - e\cos E) = a + (r_0 - a)\cos \widehat{E} + \sigma_0 \sqrt{a} \sin \widehat{E} \tag{261}$$

- Finally, we have for $\dot{G}$,

$$\dot{G} \underset{\text{Eq.(256)}}{=} 1 + \sqrt{\frac{a^3}{\mu}} \left(\cos \widehat{E} - 1\right) \dot{\widehat{E}} \tag{262a}$$

$$\underset{\text{Eq.(259)}}{=} 1 + \cancel{\sqrt{\frac{a^3}{\mu}}} \frac{\cancel{n}a}{r} \left(\cos \widehat{E} - 1\right) \tag{262b}$$

Or,

$$\boxed{\dot{G} = 1 - \frac{a}{r}\left(1 - \cos \widehat{E}\right)} \tag{263}$$

as before, $r = a(1 - e\cos E)$ or, as given in Eq.(261).

- Equations (252), (256), (260) and (263) together comprise the $F - G$ solution via Eqs.(233) and (234). Remember these equations are valid only for elliptical orbits, so you must precompute the eccentricity to confirm that they are applicable for your particular problem.

14.3 F-G solution for hyperbolas.

- Since the exact closed form solution of Keplerian motion is known for parabolic orbits, we will not consider them here. For hyperbolas, Table (6) provides (without proof) the corresponding equations for hyperbolic orbits. You can find the derivations in a standard text (e.g. Curtis, Battin etc).
- To begin, we point out an imminent terminology nightmare – the symbol F in the current context has been used for the Lagrange coefficient as it appears in Eq.(233). However, in Sec.(13.3) (see Eq.(194)), F was also defined as the hyperbolic eccentric anomaly. Therefore, to avoid (justified!) confusion, we will *redefine* the Lagrange coefficients for hyperbolas as follows:

$$\mathcal{F} \triangleq \text{Lagrange coeff. } F \text{ for hyperbolas} \tag{264a}$$

$$\mathcal{G} \triangleq \text{Lagrange coeff. } G \text{ for hyperbolas} \tag{264b}$$

$$\dot{\mathcal{F}} \triangleq \text{Lagrange coeff. } \dot{F} \text{ for hyperbolas} \tag{264c}$$

$$\dot{\mathcal{G}} \triangleq \text{Lagrange coeff. } \dot{G} \text{ for hyperbolas} \tag{264d}$$

- With the definitions of Eq.(264), Table (6) provides the $F - G$ (rather $\mathcal{F} - \mathcal{G}$) solution for hyperbolas. Note the similarity of the expressions with their elliptical counterparts.
- For hyperbolas, the $\mathcal{F} - \mathcal{G}$ solution can now be written as:

$$\mathbf{r}(t) = \mathcal{F}\,\mathbf{r}(0) + \mathcal{G}\,\mathbf{v}(0) \tag{265}$$

$$\mathbf{v}(t) = \dot{\mathcal{F}}\,\mathbf{r}(0) + \dot{\mathcal{G}}\,\mathbf{v}(0) \tag{266}$$

Elliptic orbits	Hyperbolic orbits
Define E = *eccentric anomaly* $\widehat{E} = E - E_0$	Define F = *hyperbolic eccentric anomaly* $\widehat{F} = F - F_0$
$F = 1 - \frac{a}{r_0}\left(1-\cos\widehat{E}\right)$	$\mathcal{F} = 1 - \frac{a}{r_0}\left(1-\cosh\widehat{F}\right)$
$G = (t-t_0) + \sqrt{\frac{a^3}{\mu}}\left(\sin\widehat{E} - \widehat{E}\right)$ Also, $G = \dfrac{a\sigma_0}{\sqrt{\mu}}\left(1-\cos\widehat{E}\right) + r_0\sqrt{\dfrac{a}{\mu}}\sin\widehat{E}$	$\mathcal{G} = \dfrac{a\sigma_0}{\sqrt{\mu}}\left(1-\cosh\widehat{F}\right) + r_0\sqrt{\dfrac{-a}{\mu}}\sinh\widehat{F}$
$\dot{F} = -\dfrac{\sqrt{\mu a}}{rr_0}\sin\widehat{E}$	$\dot{\mathcal{F}} = -\dfrac{\sqrt{-\mu a}}{rr_0}\sinh\widehat{F}$
$\dot{G} = 1 - \frac{a}{r}\left(1-\cos\widehat{E}\right)$	$\dot{\mathcal{G}} = 1 + \frac{a}{r}\left(\cosh\widehat{F} - 1\right)$
$r = a(1-e\cos E)$ Also, $r = a + (r_0 - a)\cos\widehat{E} + \sigma_0\sqrt{a}\sin E_0$	$r = a(1-e\cosh F)$ Also, $r = -a + (r_0+a)\cosh\widehat{F} + \sigma_0\sqrt{-a}\sinh\widehat{F}$

Table 6: Lagrange coefficients for ellipses and hyperbolas

15 Universal Solution of the Two-Body Problem

Table (7) is a summary of the methods we have considered so far for solving the two body problem. Essentially, we see that the methods that do not involve time integration must know *before-hand* what kind of orbit (in terms of eccentricity) we are dealing with. Based on that information, we then must employ a different set of equations for each case. Clearly, it would be nice to have a method that has all the "good properties" listed above, i.e. is accurate, applies to all types of orbits (ellipses, parabolas, hyperbolas alike) and does not involve time marching. Such a method does exist and is called the "universal solution of the TBP".

Method	The Good	The Bad
Explicit integration	Applies to all types of orbits (*universal*)	Error accumulation
Implicit method	Accurate, no time integration	Not universal
F-G solution	Accurate, no time integration	Not universal

Table 7: Pros and cons of methods so far

- The key idea behind this approach is to define a new "time-like" variable, χ defined as follows

$$\sqrt{\mu}dt \doteq rd\chi \tag{267}$$

Eq.(267) is called the Sundman transformation.

- Note that χ does not actually have the units of time; it is only a *time-like* variable. Let us try to get a feel for what it is. Recall that for an ellipse,

$$\frac{dE}{dt} = \sqrt{\frac{\mu}{a}}\cdot\frac{1}{r} \underset{(267)}{=} \frac{1}{\sqrt{a}}\frac{d\chi}{dt} \tag{268}$$

Thus it follows (since a is a constant),

$$\chi = \sqrt{a}\,(E - E_0) \qquad \text{(ellipses)} \tag{269}$$

Similarly, it is easy to show that for a hyperbola,

$$\chi = \sqrt{-a}\,(F - F_0) \tag{270}$$

and for a parabola,

$$\chi = \sqrt{p}\,(\tan\frac{f}{2} - \tan\frac{f_0}{2}) \tag{271}$$

- Putting Eq.(268)-(271) together, we get

$$\sqrt{\mu}\,dt = r\begin{cases} d\left(\sqrt{p}\tan\frac{f}{2}\right) \\ d(\sqrt{a}\,E) \\ d(\sqrt{-a}\,F)\end{cases} = r d\chi \tag{272}$$

  The above equation is the motivation for the definition of the universal variable, χ.

- In order to begin analysis, we would like to derive equations relating r and χ. Start with the energy integral ($v^2/2 - \mu/r = -\mu/2a$). Define

$$\alpha \doteq \frac{1}{a} \tag{273}$$

  such that

$$v^2 = 2\mu\left(\frac{1}{r} - \frac{\alpha}{2}\right) \tag{274}$$

- Next, we take derivatives:

  A. We have $r^2 = \mathbf{r}\cdot\mathbf{r}$. Its χ derivatives are:

$$\frac{dr^2}{d\chi} = 2r\frac{dr}{d\chi} = 2\mathbf{r}\cdot\frac{d\mathbf{r}}{d\chi} \tag{275a}$$

$$= 2\mathbf{r}\cdot\frac{d\mathbf{r}}{dt}\frac{dt}{d\chi} \tag{275b}$$

$$= 2\mathbf{r}\cdot\mathbf{v}\,\frac{dt}{d\chi} \tag{275c}$$

$$\underset{(267)}{=} 2\mathbf{r}\cdot\mathbf{v}\,\frac{r}{\sqrt{\mu}} \tag{275d}$$

$$= 2\cancel{\sqrt{\mu}}\sigma\frac{r}{\cancel{\sqrt{\mu}}} \qquad \left(\text{Recall } \sigma = \frac{\mathbf{r}\cdot\mathbf{v}}{\sqrt{\mu}}\right) \tag{275e}$$

  From Eqs.(275a) and (275e), we get

$$\frac{dr}{d\chi} = \sigma \tag{276}$$

  Take one more derivative:

$$\frac{d}{d\chi}\left(\frac{dr}{d\chi}\right) = \frac{d\sigma}{d\chi} \tag{277}$$

  Or,

$$\frac{d^2r}{d\chi^2} = \frac{d\sigma}{d\chi} = \frac{d\sigma}{dt}\frac{dt}{d\chi} \tag{278a}$$

$$= \frac{d}{dt}\left(\frac{\mathbf{r}\cdot\mathbf{v}}{\sqrt{\mu}}\right)\frac{r}{\sqrt{\mu}} \tag{278b}$$

$$= \frac{1}{\mu}\left(\mathbf{v}\cdot\mathbf{v} + \mathbf{r}\cdot\dot{\mathbf{v}}\right) r \tag{278c}$$

$$= \frac{r}{\mu}\left(v^2 + \mathbf{r}\cdot\left(-\frac{\mu}{r^3}\mathbf{r}\right)\right) \tag{278d}$$

$$= \frac{r}{\mu}\left(v^2 - \frac{\mu}{r}\right) \tag{278e}$$

$$\underset{(274)}{=} \frac{r}{\mu}\left(2\frac{\mu}{r} - \mu\alpha - \frac{\mu}{r}\right) \tag{278f}$$

$$\underset{(274)}{=} \frac{r}{\mu}\left(\frac{\mu}{r} - \mu\alpha\right) \tag{278g}$$

$$= 1 - r\alpha \tag{278h}$$

Or,

$$\frac{d^2 r}{d\chi^2} + r\alpha = 1 \tag{279}$$

Equivalently from Eq.(278a),

$$\frac{d\sigma}{d\chi} + r\alpha = 1 \tag{280}$$

Amazingly, Eqs.(279) and (280) are linear, (albeit non-homogeneous) ODEs with constant coefficients! So, replacing t with χ results in a linear ODE!

B. Lets take one more derivative,

$$\frac{d^3 r}{d\chi^3} = \frac{d}{d\chi}\left(\frac{d^2 r}{d\chi^2}\right) \underset{(279)}{=} \frac{d}{d\chi}(1 - r\alpha) \tag{281a}$$

$$= -\alpha\frac{dr}{d\chi} \tag{281b}$$

$$\underset{(276)}{=} -\alpha\sigma \tag{281c}$$

Now also

$$\frac{d^3 r}{d\chi^3} = \frac{d^2}{d\chi^2}\left(\frac{dr}{d\chi}\right) \underset{(276)}{=} \frac{d^2\sigma}{d\chi^2} \tag{282}$$

Combining Eqs.(282) and (281c),

$$\frac{d^2\sigma}{d\chi^2} + \alpha\sigma = 0 \tag{283}$$

Yet another linear (this one homogeneous!) ODE with constant coefficients for σ.

C. In a manner similar to the analysis shown above, starting with Eq.(267) and differentiating three times, it can be shown that,

$$\frac{d^4 t}{d\chi^4} + \alpha\frac{dt^2}{d\chi^2} = 0 \tag{284}$$

D. Finally, look at the derivatives of $\mathbf{r}$.

$$\frac{d\mathbf{r}}{d\chi} = \frac{d\mathbf{r}}{dt}\frac{dt}{d\chi} = \mathbf{v}\frac{r}{\sqrt{\mu}} \tag{285}$$

Second derivative:

$$\frac{d^2\mathbf{r}}{d\chi^2} \underset{(285)}{=} \frac{d}{d\chi}\left(\frac{r}{\sqrt{\mu}}\mathbf{v}\right) \tag{286a}$$

$$= \frac{1}{\sqrt{\mu}}\left[\frac{dr}{d\chi}\mathbf{v} + r\frac{d\mathbf{v}}{d\chi}\right] \tag{286b}$$

$$\underset{(276)}{=} \frac{1}{\sqrt{\mu}}\left[\sigma\mathbf{v} + r\frac{d\mathbf{v}}{dt}\frac{dt}{d\chi}\right] \tag{286c}$$

$$\underset{(267)}{=} \frac{1}{\sqrt{\mu}}\left[\sigma\mathbf{v} + r\left(-\frac{\mu}{r^3}\mathbf{r}\right)\frac{r}{\sqrt{\mu}}\right] \tag{286d}$$

$$= \frac{\sigma}{\sqrt{\mu}}\mathbf{v} - \frac{\mathbf{r}}{r} \tag{286e}$$

Third derivative:

$$\frac{d^3\mathbf{r}}{d\chi^3} = \frac{d}{d\chi}\left(\frac{d^2\mathbf{r}}{d\chi^2}\right) \underset{(286e)}{=} \frac{d}{d\chi}\left[\frac{\sigma}{\sqrt{\mu}}\dot{\mathbf{r}} - \frac{\mathbf{r}}{r}\right] \tag{287a}$$

$$= \frac{1}{\sqrt{\mu}}\left(\frac{d\sigma}{d\chi}\dot{\mathbf{r}} + \sigma\frac{d\dot{\mathbf{r}}}{d\chi}\right) + \left(-\frac{1}{r}\frac{d\mathbf{r}}{d\chi} + \frac{\mathbf{r}}{r^2}\frac{dr}{d\chi}\right) \tag{287b}$$

$$\underset{((280),(285),(279))}{=} \left(\frac{1}{\sqrt{\mu}}(1-r\alpha)\,\dot{\mathbf{r}} + \frac{\sigma}{\sqrt{\mu}}\frac{d\dot{\mathbf{r}}}{dt}\frac{dt}{d\chi}\right) + \left(-\frac{1}{r}\frac{r}{\sqrt{\mu}}\mathbf{v} + \frac{1}{r^2}\sigma\mathbf{r}\right) \tag{287c}$$

$$\underset{(267)}{=} \left(\frac{1}{\sqrt{\mu}}(1-r\alpha)\,\dot{\mathbf{r}} + \frac{\sigma}{\sqrt{\mu}}\left(-\frac{\mu}{r^3}\mathbf{r}\right)\frac{r}{\sqrt{\mu}}\right) + \left(-\frac{1}{r}\frac{r}{\sqrt{\mu}}\mathbf{v} + \frac{1}{r^2}\sigma\mathbf{r}\right) \tag{287d}$$

$$= \cancel{\frac{\dot{\mathbf{r}}}{\sqrt{\mu}}} - \frac{r\alpha}{\sqrt{\mu}}\dot{\mathbf{r}} - \cancel{\frac{\sigma\mathbf{r}}{r^2}} - \cancel{\frac{\mathbf{v}}{\sqrt{\mu}}} + \cancel{\frac{\sigma}{r^2}\mathbf{r}} \tag{287e}$$

$$\underset{(285)}{=} -\alpha\frac{d\mathbf{r}}{d\chi} \tag{287f}$$

Thus,

$$\frac{d^3\mathbf{r}}{d\chi^3} + \alpha\frac{d\mathbf{r}}{d\chi} = 0 \tag{288}$$

It is not a surprise by now that the above is also a linear, homogeneous ODE with constant coefficients.

♣ Summary so far: Having replaced t with χ, we have managed to obtain a sequence of linear constant coefficient ODEs for t, σ, r and $\mathbf{r}$. Moreover, these are decoupled (look at Eqs.(279),(283),(284) and (288)). Put together, these are:

$$\frac{d^2 r}{d\chi^2} + r\alpha = 1 \tag{289a}$$

$$\frac{d^2\sigma}{d\chi^2} + \alpha\sigma = 0 \tag{289b}$$

$$\frac{d^3\mathbf{r}}{d\chi^3} + \alpha\frac{d\mathbf{r}}{d\chi} = 0 \tag{289c}$$

$$\frac{d^4 t}{d\chi^4} + \alpha\frac{dt^2}{d\chi^2} = 0 \tag{289d}$$

15.1 Solution of the Universal Equations

Remember, we are ultimately looking for $\mathbf{r}(t)$ and $\mathbf{v}(t)$. We begin with Eq.(289b). This is a linear oscillator and can be solved easily as

$$\sigma(\chi) = A\cos\omega\chi + B\sin\omega\chi \tag{290}$$

Even though the above is a perfectly valid (and compact) solution for $\sigma(\chi)$, we chose to write it in power series form as follows

$$\sigma(\chi) = \sum_{k=0}^{\infty} a_k \chi^k \tag{291}$$

In order to obtain the coefficients a_k, we substitute the proposed solution (Eq.(291)) into the governing equation (289b) and match coefficients to get

$$a_{k+2} = -\frac{\alpha a_k}{(k+1)(k+2)} \tag{292}$$

Clearly, the above is a recursion and holds for $k = 0, 1, 2, \ldots$. Thus we have

$$\sigma(\chi) = a_0\left(1 - \frac{\alpha\chi^2}{2!} + \frac{(\alpha\chi^2)^2}{4!} - \ldots\right) + a_1\chi\left(1 - \frac{\alpha\chi^2}{3!} + \frac{(\alpha\chi^2)^2}{5!} + \ldots\right) \tag{293}$$

To move further, define the first two of the so-called *universal functions*:

$$U_0(\alpha;\chi) \doteq 1 - \frac{\alpha\chi^2}{2!} + \frac{(\alpha\chi^2)^2}{4!} - \frac{(\alpha\chi^2)^3}{6!} + \ldots \tag{294}$$

$$U_1(\alpha;\chi) \doteq \chi\left(1 - \frac{\alpha\chi^2}{3!} + \frac{(\alpha\chi^2)^2}{5!} - \frac{(\alpha\chi^2)^3}{7!} + \ldots\right) \tag{295}$$

To give us

$$\sigma(\chi) = a_0 U_0(\alpha;\chi) + a_1 U_1(\alpha;\chi) \tag{296}$$

As mentioned above, $U_0(\alpha;\chi)$ and $U_1(\alpha;\chi)$ are known as universal functions. Looking at Eqs.(294) and (295) together, note that

$$U_1(\alpha;\chi) = \int_0^{\chi} U_0(\alpha;\chi)d\chi \tag{297}$$

We can easily extend the sequence of Eqs.(294), (295) and (297) to obtain an infinite sequence of universal functions, $U_n(\alpha;\chi)$, generally defined as

$$U_n(\alpha;\chi) \doteq \chi^n\left(\frac{1}{n!} - \frac{\alpha\chi^2}{(n+2)!} + \frac{(\alpha\chi^2)^2}{(n+4)!} - \frac{(\alpha\chi^2)^3}{(n+6)!} + \cdots\right) \tag{298}$$

The above sequence of functions satisfies numerous identities, some of which are given below

1. $$U_k(\alpha;\chi) = \int_0^{\chi} U_{k-1}(\alpha;\chi)\,d\chi \tag{299}$$

2. $$U_{k-1}(\alpha;\chi) = \frac{dU_k(\alpha;\chi)}{d\chi} \tag{300}$$

   except for $U_0(\alpha;\chi)$, for which

   $$\frac{dU_0(\alpha;\chi)}{d\chi} = -\alpha U_1(\alpha;\chi) \tag{301}$$

3. $$U_k(\alpha;\chi) + \alpha U_{k+2}(\alpha;\chi) = \frac{\chi^k}{k!} \tag{302}$$

Another important property of the sequence $\{U_k(\alpha;\chi)\}_{k=0}^{\infty}$ is that it is a linearly independent set. In other words,

$$c_0U_0 + c_1U_1 + c_2U_2 + \ldots + c_NU_N = 0 \Leftrightarrow c_0, c_1, c_2, \ldots c_N = 0 \tag{303}$$

for all N. This is crucial in what follows and can be proved using the concept of a *Wronskian* (proof not important here). Now, as mentioned above, numerous other identities can be derived, but in the current context, the above are the most useful to us.

Now, in view of Eqs.(289a) and (289c), we propose a "third-order solution" for the magnitude, r as follows:

$$r(\chi) = c_0U_0(\alpha;\chi) + c_1U_1(\alpha;\chi) + c_2U_2(\alpha;\chi) \tag{304}$$

where $c_0 - c_2$ are coefficients to be determined. The "order" selection is motivated by the fact that given its governing equation, it is not expected that $r(\chi)$ would admit contributions from universal functions with leading terms beyond χ^2. In order to determine the coefficients, we utilize the initial conditions:

- First, note the following *convention*: @ $t = t_0$, we set $\chi = 0$. This comes from Eq.(267). As a result,
$$r(\chi = 0) = r(t_0) = r_0 \tag{305}$$
Using Eq.(305) in (304), we have
$$r(\chi = 0) \;=\; \underset{\text{(known)}}{r_0} \;=\; c_0U_0(\alpha;0) + c_1U_1(\alpha;0) + c_2U_2(\alpha;0) \tag{306}$$
However, from Eq.(298), $U_0(\alpha; 0) = 1$, $U_1(\alpha; 0) = 0$, $U_2(\alpha; 0) = 0$. In fact, $U_k(\alpha; 0) = 0 \;\; \forall\, k \geq 1$. Sub this information in Eq.(306),
$$r_0 \;=\; c_0 \cdot 1 \;=\; c_0 \tag{307}$$

- Next, differentiate Eq.(304) to get
$$\frac{dr}{d\chi} \;=\; c_0\frac{dU_0}{d\chi} + c_1\frac{dU_1}{d\chi} + c_2\frac{dU_2}{d\chi} \tag{308a}$$
$$\underset{(300),(301)}{=} -\alpha c_0U_1(\alpha;\chi) + c_1U_0(\alpha;\chi) + c_2U_1(\alpha;\chi) \tag{308b}$$
$$\sigma \;\underset{(276)}{=}\; (c_2 - \alpha c_0)U_1(\alpha;\chi) + c_1U_0(\alpha;\chi) \tag{308c}$$
Thus at $\chi = 0$ (t_0),
$$\underset{\text{(known)}}{\sigma_0} \;\underset{(298)}{=}\; c_1 \tag{309}$$

- Finally, differentiate Eq.(304) one more time:
$$\frac{d^2r}{d\chi^2} \;\underset{(308c)}{=}\; (c_2 - \alpha c_0)\frac{dU_1}{d\chi} + c_1\frac{dU_0}{d\chi} \tag{310a}$$
$$\underset{(300),(301)}{=} (c_2 - \alpha c_0)U_0(\alpha;\chi) - \alpha c_1U_1(\alpha;\chi) \tag{310b}$$
$$(1 - r\alpha) \;\underset{(289a)}{=}\; (c_2 - \alpha c_0)U_0(\alpha;\chi) - \alpha c_1U_1(\alpha;\chi) \tag{310c}$$
Thus @ t_0 $(\chi = 0)$,
$$(1 - \alpha r_0) \;\underset{(298)}{=}\; (c_2 - \alpha c_0)\cdot 1; \tag{311a}$$
$$\text{i.e., } c_2 \;=\; 1 - \alpha r_0 + \alpha c_0 \tag{311b}$$
$$\underset{(307)}{=}\; 1 \tag{311c}$$

- Collecting Eqs.(307),(309), (311c) and (304),
$$r(\chi) \;=\; r_0U_0(\alpha;\chi) + \sigma_0U_1(\alpha;\chi) + U_2(\alpha;\chi) \tag{312}$$

To assure yourself that $U_3(\alpha;\chi), U_4(\alpha;\chi)$ etc. do not contribute, let $r(\chi) \;=\; c_0U_0 + c_1U_1 + c_2U_2 + c_3U_3 + \ldots$ and following the same procedure as above, show that $c_3 = c_4 = \ldots = 0$. Proof for c_3 is given below:
Let $r(\chi) \;=\; c_0U_0 + c_1U_1 + c_2U_2 + c_3U_3$. Thus we have

$$\frac{dr}{d\chi} \;=\; \sigma \;=\; (c_2 - c_0\alpha)U_1 + c_1U_0 + c_3U_2 \tag{313a}$$
$$\frac{d^2r}{d\chi^2} \;=\; (1 - \alpha r) \;=\; (c_2 - c_0\alpha)U_0 + (c_3 - c_1\alpha)U_1 \tag{313b}$$
$$\frac{d^3r}{d\chi^3} \;\underset{(313b)}{=}\; -\alpha\frac{dr}{d\chi} \;=\; -\alpha\sigma \;=\; -\alpha(c_2 - c_0\alpha)U_1 + (c_3 - c_1\alpha)U_0 \tag{313c}$$

Thus at t_0 ($\chi = 0$),

$$
\begin{aligned}
r_0 &= c_0 \cdot 1 + 0 \Rightarrow c_0 = r_0 && (314a)\\
\sigma_0 &= c_1 \cdot 1 + 0 \Rightarrow c_1 = \sigma_0 && (314b)\\
(1 - \alpha r_0) &= (c_2 - c_0\alpha) \cdot 1 + 0 \Rightarrow c_2 = 1 - r_0\alpha + c_0\alpha = 1 && (314c)\\
-\alpha\sigma_0 &= (c_3 - c_1\alpha) \cdot 1 - 0 = c_3 = -\alpha\sigma_0 + c_1\alpha = 0 \text{ (as desired!)} && (314d)
\end{aligned}
$$

♣ **Time out.** What is going on here? We replaced t with "χ" and all of a sudden we find ourselves solving simple *linear* ordinary differential equations! Where did all the "difficulties" of the two-body problem disappear? Is there a catch? Could it have been so simple all along?

It turns out there **is** a catch. And its nothing but the very first equation in these developments, namely, Eq.(267). We were able to linearize the TBP using the ingenious transformation $\sqrt{\mu}dt = rd\chi$... but this transformation is *nonlinear* (RHS is the product of r and χ)! As a result, we will encounter the same difficulties as before when we try to inverse-transform all χ-domain quantities back into the *time* (t)-*domain*. This is yet another manifestation of the *no-free-lunch* theorem.

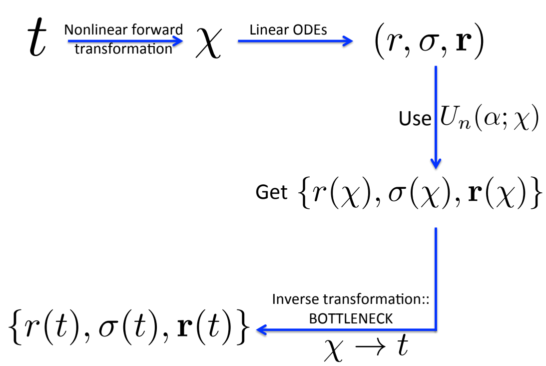


Figure 31: The bottleneck in the universal solution of TBP

The bottleneck, i.e. the inverse transformation ($\chi \to t$) is illustrated in Fig.(31). Clearly, the only thing left to do is to obtain this map. We could use Eq.(289d) for this purpose, but there is an easier way. Recall the original transformation: $\sqrt{\mu}dt = rd\chi$. Now, using Eq.(312),

$$
\begin{aligned}
\sqrt{\mu}dt &= (r_0U_0(\alpha;\chi) + \sigma_0U_1(\alpha;\chi) + U_2(\alpha;\chi))d\chi && (315a)\\
&= r_0U_0(\alpha;\chi)d\chi + \sigma_0U_1(\alpha;\chi)d\chi + U_2(\alpha;\chi)d\chi && (315b)
\end{aligned}
$$

But from Eqs.(300) and (301), $U_2d\chi = dU_3$, $U_1d\chi = dU_2$ and $U_0d\chi = dU_1$. Thus,

$$\sqrt{\mu}dt = r_0dU_1 + \sigma_0dU_2 + dU_3 \tag{316}$$

Integrating out,

$$\sqrt{\mu}(t - t_0) = r_0U_1(\alpha;\chi) + \sigma_0U_2(\alpha;\chi) + U_3(\alpha;\chi) \tag{317}$$

The above equation is the $t - \chi$ map and is aptly called the ***universal Kepler's equation***. Given the current time t, we must solve Eq.(317) to obtain $\chi(t)$. And that is the bottleneck, since Eq.(317) is nonlinear in χ. An iterative technique must be used, just like for the Kepler's equation corresponding to elliptical orbits.

Newton Raphson Iterations

- Initial guess: $\chi^{(0)}$. See example problem below on how to select a good initial guess.
- k^{th} correction.

$$\Delta\chi^{(k)} \;=\; \frac{\sqrt{\mu}(t-t_0) - [r_0 U_1(\alpha;\chi^{(k-1)}) + \sigma_0 U_2(\alpha;\chi^{(k-1)}) + U_3(\alpha;\chi^{(k-1)})]}{r_0 U_0(\alpha;\chi^{(k-1)}) + \sigma_0 U_1(\alpha;\chi^{(k-1)}) + U_2(\alpha;\chi^{(k-1)})} \tag{318}$$

- k^{th} update.

$$\chi^{(k)} \;=\; \chi^{(k-1)} + \Delta\chi^{(k)} \tag{319}$$

Repeat for $k = 1, 2, 3, \ldots, N$ until the desired error tolerance $|\Delta\chi^{(k)}|$ is met.

We are now in the position to obtain our final answer: $\mathbf{r}(t)$ and $\mathbf{v}(t)$. By solving the universal Kepler's equation (((317)), we already have $\chi(t)$, where t is given. All we need now is $\mathbf{r}(\chi)$ and $\mathbf{v}(\chi)$ so that we can find $\mathbf{r}(t) = \mathbf{r}(\chi(t))$ and $\mathbf{v}(t) = \mathbf{v}(\chi(t))$.

First determine $\mathbf{r}(\chi)$:

- Following the same logic as Eq.(304), we propose

$$\mathbf{r}(\chi) \;=\; \mathbf{z}_0 U_0(\alpha;\chi) + \mathbf{z}_1 U_1(\alpha;\chi) + \mathbf{z}_2 U_2(\alpha;\chi) \tag{320}$$

where, vector coefficients $\mathbf{z}_0$, $\mathbf{z}_1$ and $\mathbf{z}_2$ need to be determined from initial conditions.

- We have $\mathbf{r}(\chi = 0) = \mathbf{r}_0 = \mathbf{z}_0 \cdot 1 + 0$. Or,

$$\mathbf{z}_0 \;=\; \mathbf{r}_0 \tag{321}$$

- Differentiate Eq.(320) and use Eq.(285) to get

$$\frac{d\mathbf{r}(\chi)}{d\chi} \;=\; \mathbf{v}\frac{r}{\sqrt{\mu}} \;=\; -\alpha\mathbf{z}_0 U_1(\alpha;\chi) + \mathbf{z}_1 U_0(\alpha;\chi) + \mathbf{z}_2 U_1(\alpha;\chi) \tag{322}$$

Thus at t_0,

$$\mathbf{v}_0\frac{r_0}{\sqrt{\mu}} \;=\; \mathbf{z}_1 \cdot 1 + 0 = \mathbf{z}_1 \tag{323}$$

- Differentiate one more time to get

$$\frac{d^2\mathbf{r}(\chi)}{d\chi^2} = -\frac{\mathbf{r}}{r} + \frac{\sigma}{\sqrt{\mu}}\mathbf{v} \;=\; (\mathbf{z}_2 - \alpha\mathbf{z}_0)U_0(\alpha;\chi) - \alpha\mathbf{z}_1 U_1(\alpha;\chi) \tag{324}$$

Thus at t_0

$$-\frac{\mathbf{r}_0}{r_0} + \frac{\sigma_0}{\sqrt{\mu}}\mathbf{v}_0 \;=\; \mathbf{z}_2 - \alpha\mathbf{z}_0 \;=\; \mathbf{z}_2 - \alpha\mathbf{r}_0 \tag{325}$$

Or

$$\mathbf{z}_2 \;=\; \frac{\sigma_0}{\sqrt{\mu}}\mathbf{v}_0 + \left(\alpha - \frac{1}{r_0}\right)\mathbf{r}_0 \tag{326}$$

- Sub Eqs.(321), (323) and (326) into Eq.(320) to get

$$\mathbf{r}(\chi) \;=\; \mathbf{r}_0 U_0(\alpha;\chi) + \frac{r_0}{\sqrt{\mu}}\mathbf{v}_0 U_1(\alpha;\chi) + \left[\frac{\sigma_0}{\sqrt{\mu}}\mathbf{v}_0 + \left(\alpha - \frac{1}{r_0}\right)\mathbf{r}_0\right] U_2(\alpha;\chi) \tag{327}$$

The above is a universal solution for $\mathbf{r}(\chi)$! Note that the above can be rearranged to get

$$\mathbf{r}(\chi) \;=\; \mathbf{r}_0 \underbrace{\left[U_0(\alpha;\chi) + \left(\alpha - \frac{1}{r_0}\right) U_2(\alpha;\chi)\right]}_{F(\chi)} + \mathbf{v}_0 \underbrace{\left[\frac{r_0}{\sqrt{\mu}} U_1(\alpha;\chi) + \frac{\sigma_0}{\sqrt{\mu}} U_2(\alpha;\chi)\right]}_{G(\chi)} \tag{328}$$

which is a universal $F-G$ solution for $\mathbf{r}(\chi)$!! Very neat indeed. What is interesting is that without intending it, we found the universal solution of $\mathbf{r}(\chi)$ for the two-body problem in the $F-G$ form. We have

$$\mathbf{r}(\chi) \;=\; F(\chi)\mathbf{r}_0 \;+\; G(\chi)\mathbf{v}_0 \tag{329}$$

where

$$F(\chi) \;=\; U_0(\alpha;\chi) \;+\; \left(\alpha - \frac{1}{r_0}\right)U_2(\alpha;\chi) \tag{330a}$$

$$G(\chi) \;=\; \frac{r_0}{\sqrt{\mu}}U_1(\alpha;\chi) + \frac{\sigma_0}{\sqrt{\mu}}U_2(\alpha;\chi) \tag{330b}$$

In fact, Eq.(330a) can be simplified further. Use Eq.(302) for $k=0$, i.e. $U_0 + \alpha U_2 = 1$ such that

$$\begin{aligned} F(\chi) \;&=\; U_0 + \alpha U_2 - \frac{U_2}{r_0} \\ &=\; 1 - \frac{U_2(\alpha;\chi)}{r_0} \end{aligned} \tag{331}$$

Next we have $\mathbf{v}(\chi)$, which can be found via Eq.(285). The details are skipped here. See the summary and example problem below.

Summary. Given t_0, $\mathbf{r}(t_0) = \mathbf{r}_0$, $\mathbf{v}(t_0) = \mathbf{v}_0$ and t, determine $\mathbf{r}(t)$ and $\mathbf{v}(t)$.

Step 1. Solve the universal Kepler's equation to find $\chi(t)$.

Step 2. Find $\mathbf{r}(\chi(t))$ via Eq.(329) and $\mathbf{v}(\chi(t))$ as given below

$$\mathbf{v}(\chi(t)) \;=\; \dot{F}(\chi)\mathbf{r}_0 + \dot{G}(\chi)\mathbf{v}_0 \tag{332}$$

where, $\dot{F}(\chi)$ and $\dot{G}(\chi)$ are

$$\dot{F}(\chi) \;=\; -\frac{\sqrt{\mu}}{r r_0}U_1(\alpha;\chi) \tag{333a}$$

$$\dot{G}(\chi) \;=\; 1 - \frac{1}{r}U_2(\alpha;\chi) \tag{333b}$$

where, $r = r(\chi)$ was given in Eq.(312).

❣ **Example** A spacecraft was found to have the following initial position and velocity vectors:

$$\begin{aligned} \mathbf{r}_0 \;&=\; [20,000, -105,000, -19,000]' \times 10^3 \; m; \text{and,} \\ \mathbf{v}_0 \;&=\; [0.9, -3.4; -1.5]' \times 10^3 \; m/s \end{aligned}$$

Determine its position and velocity vectors two hours later.

Solution. We will use the universal approach, i.e. not bother with determining the characteristics of the orbit the spacecraft is in. We only need α. Based on given information, $r_0 = \|\mathbf{r}_0\| = 1.0856 \times 10^8 \; m$; $v_0 = \|\mathbf{v}_0\| = 3.8236 \times 10^3 \; m/s$ and $\sigma_0 = \frac{\mathbf{r}_0 \cdot \mathbf{v}_0}{\sqrt{\mu}} \;=\; 2.0210 \times 10^4 \; m^{1/2}$. Thus energy $\varepsilon = v_0^2/2 - \mu/r_0 = -\mu\alpha/2$ $= +3.6384 \times 10^6 \; m^2/s^2$; such that $\alpha = -2\varepsilon/\mu = -1.8256 \times 10^{-8} \; m^{-1}$. Of course, this is a hyperbola, but we will not use this fact.

Step 1. Solve the universal Kepler's equation. We have $t = t_0 +$ 2 hours $= (t_0 + 7200)\; s$. Without loss of generality, let $t_0 = 0$, such that $t = 7200 \; s$. The Kepler's equation to be solved is:

$$7200\sqrt{\mu} \;=\; r_0 U_1(\alpha;\chi) + \sigma_0 U_2(\alpha;\chi) + U_3(\alpha;\chi) \tag{334}$$

where, χ is the unknown. To begin the iterations, we could assume $\chi^{(0)} = 0$. But a better guess can be made as follows: Recall Eq.(267): $\sqrt{\mu}dt = rdt$. If we assume that r is a constant and equals r_0 (circular orbit assumption), then $\sqrt{\mu}dt = r_0 dt$ and integrating, $\sqrt{\mu}(t - t_0) = r_0\chi$; which can be solved to obtain the initial guess,

$$\chi^{(0)} = \frac{\sqrt{\mu}}{r_0}\Delta t \tag{335}$$

In the present problem, we get $\chi^{(0)} = 1.3241 \times 10^3\ m^{1/2}$. We set **a threshold of** 10^{-9} for the error and perform Newton-Raphson iterations as laid out in Eqs.(318)-(319). See Table (8) (only four decimal places are shown).

k	$\Delta\chi^{(k)}$	$\chi^{(k)}$
0	-	1.3241×10^3
1	-1.3701×10^2	1.1871×10^3
2	-1.6955×10^0	1.1854×10^3
3	-2.5631×10^{-4}	1.1854×10^3
4	-7.2487×10^{-12}	1.1854×10^3

Table 8: Newton-Raphson Iterations for Universal Kepler's Equation

Clearly, the tolerance is met after $k = 4$ iterations and the final answer is $\chi(t) \approx 1.1854 \times 10^3\ m^{1/2}$.

Step 2. Use the $F - G$ solution to determine $\mathbf{r}(t) = \mathbf{r}(\chi(t))$. Using Eq.(312) with $\chi(t) = 1.1854 \times 10^3\ m^{1/2}$, we get $r(\chi) = 1.3472 \times 10^8\ m$. Also, following Eqs.(331) and (330b):

$$\begin{aligned} F(\chi) &= 1 - \frac{U_2(\alpha;\chi)}{r_0} = 9.9351 \times 10^{-1} \\ G(\chi) &= \frac{1}{\sqrt{\mu}}\left(r_0 U_1(\alpha;\chi) + \sigma_0 U_2(\alpha;\chi)\right) = 7.1861 \times 10^3 \end{aligned}$$

See notes below on evaluation of the universal functions. Thus;

$$\begin{aligned} \mathbf{r}(t) &= \mathbf{r}_0 F(\chi) + \mathbf{v}_0 G(\chi) \\ &= [2.6338 \times 10^7; -1.2875 \times 10^8; -2.9656 \times 10^7]'\ m \end{aligned}$$

Step 3. Compute $\mathbf{v}(\chi(t))$. We have

$$\begin{aligned} \dot{F}(\chi) &= -\frac{\sqrt{\mu}}{r r_0} U_1(\alpha;\chi) = -1.6250 \times 10^{-6} \\ \dot{G}(\chi) &= 1 - \frac{U_2(\alpha;\chi)}{r} = 9.9477 \times 10^{-1} \end{aligned}$$

Thus

$$\begin{aligned} \mathbf{v}(t) &= \mathbf{r}_0 \dot{F}(\chi) + \mathbf{v}_0 \dot{G}(\chi) \\ &= [8.6280 \times 10^2; -3.2116 \times 10^3; -1.4613 \times 10^3]'\ m/s \end{aligned}$$

♣ **How were the universal functions appearing above computed?**

Recall that

$$U_n(\alpha;\chi) \;=\; \chi^n\left(\frac{1}{n!} - \frac{(\alpha\chi^2)}{(n+2)!} + \frac{(\alpha\chi^2)^2}{(n+4)!} - \dots\right)$$

which is an infinite sum! First, $U_0(\alpha;\chi)$ and $U_1(\alpha;\chi)$ are computed using the above expression by including enough terms in the summation to meet a specified tolerance, say 10^{-9}. This can be implemented by writing a while-loop which is executed until the magnitude of contribution from each additional term is less than the specified threshold.

Then, for $U_2(\alpha;\chi)$ and higher, use the recursive identity of Eq.(302), i.e. $U_n + \alpha U_n = \chi^n/n!$, such that

$$U_0(\alpha;\chi) + \alpha U_2(\alpha;\chi) \;=\; 1 \Rightarrow U_2(\alpha;\chi) = \frac{1 - U_0(\alpha;\chi)}{\alpha} \tag{336}$$

$$U_1(\alpha;\chi) + \alpha U_3(\alpha;\chi) \;=\; \chi \Rightarrow U_3(\alpha;\chi) = \frac{\chi - U_1(\alpha;\chi)}{\alpha} \tag{337}$$

♣ **Alternate initial guess:** It turns out that if instead of the "good initial guess" given in Eq.(335), the "dumb guess" of $\chi^{(0)} = 0$ was used, the threshold of $\|\Delta\chi^{(k)}\| \leq 10^{-9}$ was met after five (as opposed to four) iterations. In other words, in this case, the "good" guess helped, but not by a lot.