Politecnico di Torino

SEEDS

SpacE Exploration and Development Systems

Astrodynamics

III Edition 2007 - 08 - Ver. 2.0.1

Authors:

Guido Colasurdo
Dipartimento di Meccanica e Aeronautica
Sapienza Università di Roma

and

Giulio Avanzini
Dipartimento di Ingegneria Aeronautica e Spaziale
Politecnico di Torino

Contents

1 Two–Body Orbital Mechanics **1**

1.1 Birth of Astrodynamics: Kepler's Laws 1
1.2 Newton's Laws of Motion 2
1.3 Newton's Law of Universal Gravitation 3
1.4 The n–Body Problem 4
1.5 Equation of Motion in the Two-Body Problem 5
1.6 Potential Energy 6
1.7 Constants of the Motion 7
1.8 Trajectory Equation 8
1.9 Conic Sections 8
1.10 Relating Energy and Semi-major Axis 9

2 Two-Dimensional Analysis of Motion **11**

2.1 Reference Frames 11
2.2 Velocity and acceleration components 12
2.3 First-Order Scalar Equations of Motion 12
2.4 Perifocal Reference Frame 13
2.5 Flight Path Angle 14
2.6 Elliptical Orbits 15
2.6.1 Geometry of an Elliptical Orbit 15
2.6.2 Period of an Elliptical Orbit 16
2.7 Time–of–Flight on the Elliptical Orbit 16
2.8 Extension to hyperbola and parabola 18
2.9 Circular and Escape Velocity, Hyperbolic Excess Speed 18
2.10 Cosmic Velocities 19

3 Three-Dimensional Analysis of Motion **21**

3.1 The Measurement of Time 21
3.2 Non–rotating Frames 23
3.3 Classical Orbital Elements 24
3.4 Modified Equinoctial Orbital Elements 25
3.5 Determining the orbital elements 26
3.6 Determining Spacecraft Position and Velocity 27
3.7 Coordinate Transformation 27
3.8 Derivative in a Rotating Reference Frame 29
3.9 Topocentric Reference Frame 30
3.10 Satellite Ground Track 32
3.11 Ground Visibility 32

4 Earth Satellites **37**
4.1 Introduction . . . 37
4.2 The cost of thrusting in space . . . 37
4.3 Energetic aspects of the orbit injection . . . 38
4.4 Injection errors . . . 40
4.5 Perturbations . . . 41
4.5.1 Special perturbations . . . 42
4.5.2 General perturbations . . . 42
4.6 Perturbed satellite orbits . . . 42
4.6.1 Lunar and solar attraction . . . 43
4.6.2 Asphericity of the Earth . . . 44
4.6.3 Aerodynamic drag . . . 45
4.6.4 Radiation pressure . . . 46
4.6.5 Electromagnetic effects . . . 46
4.7 Geographical constraints . . . 46
4.8 Practical orbits . . . 47

5 Orbital Manoeuvres **51**
5.1 Introduction . . . 51
5.2 One-impulse manoeuvres . . . 51
5.2.1 Adjustment of perigee and apogee height . . . 52
5.2.2 Simple rotation of the line-of-apsides . . . 52
5.2.3 Simple plane change . . . 53
5.2.4 Combined change of apsis altitude and plane orientation . . . 54
5.3 Two-impulse manoeuvres . . . 54
5.3.1 Change of the time of periapsis passage . . . 56
5.3.2 Transfer between circular orbits . . . 56
5.3.3 Hohmann transfer . . . 57
5.3.4 Noncoplanar Hohmann transfer . . . 58
5.4 Three–impulse manoeuvres . . . 59
5.4.1 Bielliptic transfer . . . 59
5.4.2 Three-impulse plane change . . . 60
5.4.3 Three-impulse noncoplanar transfer between circular orbits . . . 62

6 Lunar trajectories **63**
6.1 The Earth-Moon system . . . 63
6.2 Simple Earth-Moon trajectories . . . 63
6.3 The patched-conic approximation . . . 67
6.3.1 Geocentric leg . . . 67
6.3.2 Selenocentric leg . . . 68
6.3.3 Noncoplanar lunar trajectories . . . 71
6.4 Restricted Circular Three–Body Problem . . . 72
6.4.1 System Dynamics . . . 72
6.4.2 Jacobi Integral . . . 75
6.4.3 Jacobi Integral in terms of absolute velocity . . . 77
6.4.4 Lagrangian libration points . . . 78
6.4.5 Lagrange Point stability . . . 81
6.4.6 Surfaces of zero velocity . . . 83
6.5 Summary and practical considerations . . . 85

7 (Introduction to) Interplanetary trajectories **91**
7.1 The Solar system 91
7.1.1 Sun and planets 91
7.1.2 Other bodies 93
7.1.3 Orbital resonance 94
7.2 Interplanetary missions 94
7.2.1 Generalities 94
7.2.2 Elongation and sinodic period 95
7.3 Heliocentric transfer 95
7.3.1 Ideal planets' orbits 96
7.3.2 Accounting for real planets' orbits 96
7.4 Earth departure trajectory 97
7.5 Arrival to the target planet 98

Chapter 1

Two–Body Orbital Mechanics

In this chapter some basics concepts of Classical Mechanics will be reviewed in order to develop an analytical solution for the two–body problem from Newton's Law of Universal Gravitation. The constants of the motion and orbit parameters will be introduced while deriving the trajectory equation.

1.1 Birth of Astrodynamics: Kepler's Laws

Since the time of Aristotle the motion of planets was thought as a combination of smaller circles moving on larger ones. The official birth of modern Astrodynamics was in 1609, when the German mathematician Johannes Kepler published his first two laws of planetary motion. The third followed in 1619. From *Astronomia Nova* we read

... sequenti capite, ubi simul etiam demonstrabitur, nullam Planetae relinqui figuram Orbitæ, praeterquam perfecte ellipticam; conspirantibus rationibus, a principiis Physicis, derivatis, cum experientia observationum et hypotheseos vicariæ hoc capite allegata.

and

Quare ex superioribus, sicut se habet CDE area ad dimidium temporis restitutorii, quod dicatur nobis 180 gradus: sic CAG, CAH areae ad morarum in CG et CH diuturnitatem. Itaque CGA area fiet mensura temporis seu anomaliae mediae, quae arcui eccentrici CG respondet, cum anomalia media tempus metiatur.

The third law was formulater in "Harmonice Mundi"

Sed res est certissima exactissimaque, quod proportio quae est inter binorum quorumcunque Planetarum tempora periodica, sit praecise sesquialtera proportionis mediarum distantiarum, id est Orbium ipsorum

In modern terms, the three laws can be formulated as follows:

First Law: The orbit of each planet is an ellipse with the Sun at a focus.

Second Law: The line joining the planet to the Sun sweeps out equal areas in equal times.

Third Law: The square of the period of a planet is proportional to the cube of its mean distance from the Sun.

Kepler derived a geometrical and mathematical description of the planets' motion from the accurate observations of his professor, the Danish astronomer Tycho Brahe. Kepler first described the orbit of a planet as an ellipse; the Sun is at one focus and there is no object at the other. In the second law he foresees the conservation of angular momentum: since the distance

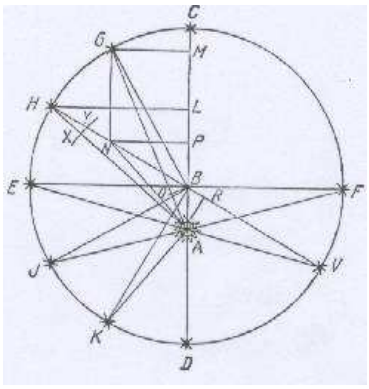

Figure 1.1: **Kepler's second law.**

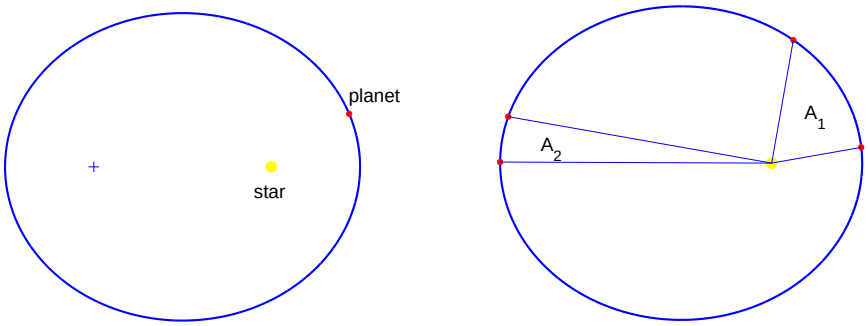

Figure 1.2: **Sketch of Kepler's first and second laws.**

of a planet from the Sun varies and the area being swept remains constant, a planet has variable velocity, that is, the planet moves faster when it falls towards the Sun and is accelerated by the solar gravity, whereas it will be decelerated on the way back out. The semi-major axis of the ellipse can be regarded as the average distance between the planet and the Sun, even though it is not the time average, as more time is spent near the apocenter than near pericenter. Not only the length of the orbit increases with distance, also the orbital speed decreases, so that the increase of the sidereal period is more than proportional.

Kepler's first law can be extended to objects moving at greater than escape velocity (e.g., some comets); they have an open parabolic or hyperbolic orbit rather than a closed elliptical one. Yet the Sun lies on the focus of the trajectory, inside the "bend" drawn by the celestial body. Thus, all of the conic sections are possible orbits. The second law is also valid for open orbits (since angular momentum is still conserved), but the third law is inapplicable because the motion is not periodic. Also, Kepler's third law needs to be modified when the mass of the orbiting body is not negligible compared to the mass of the central body. However, the correction is fairly small in the case of the planets orbiting the Sun. A more serious limitation of Kepler's laws is that they assume a two-body system. For instance, the Sun-Earth-Moon system is far more complex, and for calculations of the Moon's orbit, Kepler's laws are far less accurate than the empirical method invented by Ptolemy hundreds of years before.

Kepler was able to provide only a description of the planetary motion, but paved the way to Newton, who first gave the correct explanation fifty years later.

1.2 Newton's Laws of Motion

In Book I of his Principia (1687) Newton introduces the three *Axiomata sive Leges Motus*:

Lex I - Corpus omne perseverare in statu suo quiscendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare.
Lex II - Mutationem motus proportionalem esse vi motrici impressæ, et fieri secundum lineam rectam qua vis illa imprimitur.
Lex III - Actioni contrariam semper et æ qualem esse reactionem: sive corporum duorum actiones in se mutuo semper esse æquales et in partes contrarias dirigi.

The laws of motion can be translated into the following English version:

First Law: Every object continues in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by forces impressed upon it.

Second Law: The rate of change of momentum is proportional to the force impressed and is in the same direction as that force.

Third Law: To every action there is always opposed an equal reaction..

The first law requires the identification of an inertial system with respect to which it is possible to define the absolute motion of the object. The second law can be expressed mathematically as

$$\vec{\boldsymbol{F}} = \frac{\mathrm{d}\vec{\boldsymbol{p}}}{\mathrm{d}t}$$

where $\vec{\boldsymbol{F}}$ is the resultant of the forces acting on the object and $\vec{\boldsymbol{p}} = m\vec{\boldsymbol{v}}$ is its momentum. For a constant mass m, it is

$$\vec{\boldsymbol{F}} = m\vec{\boldsymbol{a}}$$

where $\vec{\boldsymbol{a}} = \mathrm{d}\vec{\boldsymbol{v}}/\mathrm{d}t$ is the acceleration of the mass measured in an inertial reference frame. A different equation applies to a variable-mass system (as for instance a rocket launcher where the variation of the mass per unit time, $\dot{m}$, is sizable). In that case the equation of motion becomes

$$\vec{\boldsymbol{F}} = m\vec{\boldsymbol{a}} + \dot{m}\vec{\boldsymbol{v}}$$

The third law permits to deal with a dynamical problem by using an equilibrium equation. Its scope is however wider: for instance, the presence of the action $\vec{\boldsymbol{F}}$ on m implies an action $-\vec{\boldsymbol{F}}$ on another portion of an ampler system.

1.3 Newton's Law of Universal Gravitation

In the same book Newton enunciated the law of Universal Gravitation by stating that two bodies, the masses of which are$M$ and m, respectively, attract one another along the line joining them with a force proportional to the product of their masses and inversely proportional to the square of the distance between them, that is,

$$F = G\frac{Mm}{r^2}$$

where $G = 6.673\ 10^{-11}\ \mathrm{m^3kg^{-1}s^{-2}}$ is the Universal Gravitational Constant. The law is easily extended from point masses to bodies with a spherical symmetry (see Mengali, 2001, §1.3), but it is fairly accurate also for body of arbitrary shape, when their distance is several orders of magnitude larger than their own dimension.

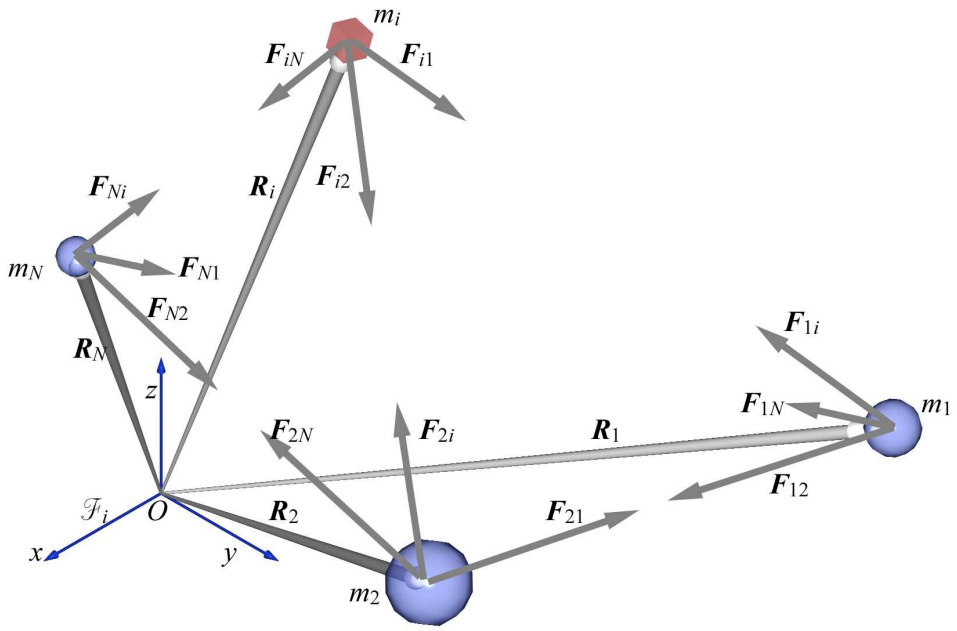

Figure 1.3: **The n–body problem.**

1.4 The n–Body Problem

If n bodies of mass m_i, $i = 1, 2, ..., n$ move in an empty space under the action of their mutual gravitational attraction only, the determination of their motion represents the so–called n–body problem. An inertial reference frame $\mathcal{F}_i(O; x, y, z)$ is required to apply Newton's second law of dynamics to each one of the n bodies. The position of the i–th one is given by the vector $\vec{\boldsymbol{R}}_i$. Assuming that each body is point–mass element, the problem is described by n second-order non-linear vector differential equations

$$\ddot{\vec{\boldsymbol{R}}}_i = -\sum_{\substack{j=1 \\ j \neq i}}^{N} G\frac{m_j}{R_{ij}^3}(\vec{\boldsymbol{R}}_i - \vec{\boldsymbol{R}}_j)$$

where $R_{ij} = ||\vec{\boldsymbol{R}}_i - \vec{\boldsymbol{R}}_j||$.

In general the n–body problem is said not to admit any analytical solution for $n > 2$, inasmuch as the set of ordinary differential equations that rules the motion has only 10 first integrals (Bruns demonstrated this theorem in 1887). Nonetheless, in 1912, the Finnish mathematician K.F. Sundman derived a series solution in powers of $t^{1/3}$ for the $n = 3$ (the so–called 3–body problem).

The n–body problem is of paramount importance in the design of space missions. The analysis of a space mission studies the motion of a spacecraft, which is the n–th body in a system that comprises other $n - 1$ celestial bodies. The simultaneous knowledge of the position of all the n bodies is necessary, as the mutual attraction between each pair of bodies depends upon their relative position: the n–body problem is the ideal model for the system.

The problem can be expressed in the form

$$\ddot{\vec{\boldsymbol{R}}}_i = \frac{1}{m_i}\left(\vec{\boldsymbol{F}}_{g,i} + \vec{\boldsymbol{F}}_{o,i}\right)$$

where

$$\vec{\boldsymbol{F}}_{g,i} = -\sum_{\substack{j=1 \\ j\neq i}}^{N} G\frac{m_j}{R_{ij}^3}(\vec{\boldsymbol{R}}_i - \vec{\boldsymbol{R}}_j)$$

is the sum of the actions on m_i of all the other $n-1$ bodies, which are assumed to have point masses (or, equivalently, a spherically symmetric mass distribution), and behaves according to Newton's law of gravitation. The additional term $\vec{\boldsymbol{F}}_{o,i}$ takes into account that

- usually the bodies are not spherically symmetric;
- the spacecraft may eject mass to obtain thrust;
- other forces (atmospheric drag, solar pressure, electromagnetic forces, etc.) act on the spacecraft;

so that $\vec{\boldsymbol{F}}_{o,i}$ represents the resultant force on m_i due to these perturbations.

The interest often lies on the motion relative to a specific celestial body, for instance, the first one. The corresponding equation can be subtracted from all the others, and the vector equations relevant for the problem of describing the motion about m_1 become $n-1$.

Moreover, the spacecraft mass is usually negligible and cannot perturb the motion of the celestial bodies. In this case the problem is called *restricted* n–body problem, where the term restricted means that the interest lies on the motion of a body so small that it does not perturb the orbits of the other $n-1$. Such orbits are regularly computed by astronomers and provided to the Astrodynamics community in the form of ephemeris. This means that the position of the $n-2$ celestial bodies m_i, $i = 2, 3, ..., n-1$ is obtained from observations and it is assumed known.

The remaining vector second–order differential equation, describing the spacecraft motion relative to the first body, is transformed into a system of six scalar first–order differential equations, that can be solved numerically. An analytical solution can be only obtained if one assumes the presence of just one celestial body, and neglects all the other actions on the spacecraft, which is acted upon only by the point–mass gravitational attraction of the primary body (see Cornelisse, 1979, §15.1 for more details).

When the resticted 3–body problem is dealt with, no explicit analytical solution is known, but some important analytical results can be obtained in the circular case, when the two primary bodies m_1 and m_2 are on a relative cricular orbit.

1.5 Equation of Motion in the Two-Body Problem

Consider a system composed of two bodies of masses m_1 and m_2, $m_1 > m_2$; the bodies are spherically symmetric and no forces other than gravitation are present. Assume an inertial frame $\mathcal{F}_i$; vectors $\vec{\boldsymbol{R}}_1$ and $\vec{\boldsymbol{R}}_2$ describe the positions of masses m_1 and m_2, respectively. The position of m_2 relative to m_1 is

$$\vec{\boldsymbol{r}} = \vec{\boldsymbol{R}}_2 - \vec{\boldsymbol{R}}_1$$

so that their mutual distance is $r = ||\vec{\boldsymbol{r}}||$ and the following equations hold for the relative velocity and acceleration,

$$\dot{\vec{\boldsymbol{r}}} = \dot{\vec{\boldsymbol{R}}}_2 - \dot{\vec{\boldsymbol{R}}}_1 \ ; \quad \ddot{\vec{\boldsymbol{r}}} = \ddot{\vec{\boldsymbol{R}}}_2 - \ddot{\vec{\boldsymbol{R}}}_1$$

where all the derivatives are taken in a (at least approximately) inertial frame.

The law of Universal Gravitation is used to express the force acting on each mass, so that the second law of dynamics can be used to describe separately the motion of each mass,

$$m_1 \ddot{\vec{\boldsymbol{R}}}_1 = G\frac{m_1 m_2}{r^3}\vec{\boldsymbol{r}} \; ; \quad m_2 \ddot{\vec{\boldsymbol{R}}}_2 = -G\frac{m_1 m_2}{r^3}\vec{\boldsymbol{r}}$$

By subtracting the second equation from the first (after an obvious simplification concerning masses) one obtains

$$\ddot{\vec{\boldsymbol{r}}} = -G\frac{m_1 + m_2}{r^3}\vec{\boldsymbol{r}}$$

In most practical systems one of the masses is several order of magnitude larger than the other one. The best balanced 2–body system in the Solar system is the planet Pluto with its satellite Caron, which has a mass equal to 1/7 that of the planet. The Earth–Moon system is the second best balanced pair, where the mass of the Moon is only 0.0123 that of the Earth.

As a consequence, it is usually possible to let $m_1 = M$ and $m_2 = m$, with $M \gg m$, so that the Equation of Relative Motion can be rewritten in the following simple form:

$$\ddot{\vec{\boldsymbol{r}}} = -\frac{\mu}{r^3}\vec{\boldsymbol{r}}$$

where $\mu = GM$ is the primary body gravitational parameter. Some numerical values of mu are reported in Tab. 1.1

It is worthwhile to remark that

- the relative position $\vec{\boldsymbol{r}}$ is measured in a non-rotating frame which is not rigorously inertial;
- the relative motion is practically independent of the mass of the secondary body, when $m \ll M$; a unit mass will be assumed in the following.

1.6 Potential Energy

The mechanical work done against the force of gravity to move the secondary body from position 1 to position 2

$$\mathcal{L} \;=\; \int_1^2 \frac{\mu}{r^3}\vec{\boldsymbol{r}} \cdot \mathrm{d}\vec{\boldsymbol{s}} \;=\; \int_1^2 \frac{\mu}{r^2}\mathrm{d}r \;=\; -\frac{\mu}{r_2} + \frac{\mu}{r_1} \;=\; \mathcal{U}_2 - \mathcal{U}_1$$

does not depend on the actual trajectory from 1 to 2, so that one deduces that the gravitational field is conservative. The work can be expressed as the difference between the values of the potential energy at point 1 and 2, the dependence of which on the radius (*i.e.* the distance of m from the M) can be expressed in the form

$$\mathcal{U} = -\frac{\mu}{r} + C$$

Table 1.1: Values of the gravitational parameter μ.

Sun	1.327	10^{11}	$\mathrm{km}^3\ \mathrm{s}^{-2}$
Earth	3.986	10^5	$\mathrm{km}^3\ \mathrm{s}^{-2}$
Moon	4.903	10^3	$\mathrm{km}^3\ \mathrm{s}^{-2}$

The value of the arbitrary constant C is conventionally assumed to be zero in Astrodynamics, that is, the maximum value of the potential energy is zero, when the spacecraft is at infinite distance from the primary body, otherwise it is negative, and equal to

$$\mathcal{U} = -\frac{\mu}{r}$$

1.7 Constants of the Motion

By taking the scalar product of the equation of motion with the velocity vector $\vec{\boldsymbol{v}} = \dot{\vec{\boldsymbol{r}}}$,

$$\dot{\vec{\boldsymbol{r}}} \cdot \ddot{\vec{\boldsymbol{r}}} = -\frac{\mu}{r^3}\left(\dot{\vec{\boldsymbol{r}}} \cdot \vec{\boldsymbol{r}}\right)$$

and writing everything on the left–hand side by using Eq. (1.6) in Appendix, one gets

$$\vec{\boldsymbol{v}} \cdot \dot{\vec{\boldsymbol{v}}} + \frac{\mu}{r^3}\left(\dot{\vec{\boldsymbol{r}}} \cdot \vec{\boldsymbol{r}}\right) = v\dot{v} + \frac{\mu}{r^3} r\dot{r}$$

that is,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{v^2}{2} - \frac{\mu}{r}\right) = 0$$

During the motion the specific mechanical energy

$$\mathcal{E} = \frac{v^2}{2} - \frac{\mu}{r}$$

that is, the sum of kinetic and potential energy, is constant, even though it can be continuously transferred from the kinetic to the potential form, and vice versa. This result is the obvious consequence of the conservative nature of the gravitational force, which is the only action on the spacecraft in the two-body problem.

The vector product of $\vec{\boldsymbol{r}}$ with the equation of motion gives

$$\vec{\boldsymbol{r}} \times \dot{\vec{\boldsymbol{r}}} = \mathbf{0} \;\Rightarrow\; \frac{\mathrm{d}}{\mathrm{d}t}\left(\vec{\boldsymbol{r}} \times \dot{\vec{\boldsymbol{r}}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}}\right) = \mathbf{0}$$

where Eq. (1.3) was applied.

Therefore, the angular momentum

$$\vec{\boldsymbol{h}} = \vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}}$$

is a constant vector. Since $\vec{\boldsymbol{h}}$ is normal to the orbital plane, the spacecraft motion remains in the same plane. This result is not surprising: the unique action is radial, no torque acts on the satellite, and the angular momentum is constant.

Finally, the cross product of the equation of motion with the (constant) angular momentum vector $\vec{\boldsymbol{h}}$ gives

$$\ddot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{h}} + \frac{\mu}{r^3}\vec{\boldsymbol{r}} \times \vec{\boldsymbol{h}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{h}}\right) + \frac{\mu}{r^3}\vec{\boldsymbol{r}} \times \left(\vec{\boldsymbol{r}} \times \dot{\vec{\boldsymbol{r}}}\right) = 0$$

By using Eqs. (1.6) and (1.8), one gets

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{h}}\right) + \frac{\mu}{r^3}\left[\left(\vec{\boldsymbol{r}} \cdot \dot{\vec{\boldsymbol{r}}}\right)\vec{\boldsymbol{r}} - \left(\vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{r}}\right)\dot{\vec{\boldsymbol{r}}}\right] &= \\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{h}}\right) + \mu\left(\frac{\dot{r}}{r^2}\vec{\boldsymbol{r}} - \frac{\dot{\vec{\boldsymbol{r}}}}{r}\right) &= \\
\frac{\mathrm{d}}{\mathrm{d}t}\left(\dot{\vec{\boldsymbol{r}}} \times \vec{\boldsymbol{h}}\right) - \mu\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\vec{\boldsymbol{r}}}{r}\right) &= 0
\end{aligned} \tag{1.1}$$

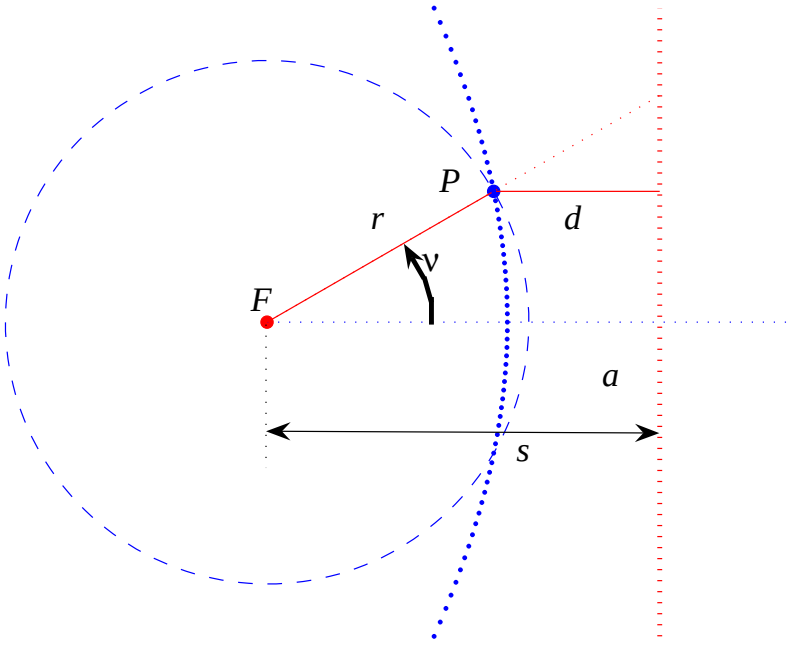

Figure 1.4: **Building a conic section.**

As a consequence, the vector

$$\vec{\boldsymbol{e}} = \frac{\vec{\boldsymbol{v}} \times \vec{\boldsymbol{h}}}{\mu} - \frac{\vec{\boldsymbol{r}}}{r}$$

is another constant of the spacecraft motion in the two-body problem.

1.8 Trajectory Equation

The dot product

$$\vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{e}} = \frac{\vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{v}} \times \vec{\boldsymbol{h}}}{\mu} - \frac{\vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{r}}}{r} = \frac{\vec{\boldsymbol{r}} \times \vec{\boldsymbol{v}} \cdot \vec{\boldsymbol{h}}}{\mu} - r = \frac{h^2}{\mu} - r$$

allows one to determine the shape of the orbit. Letting ν be the angle between the (constant) vector $\vec{\boldsymbol{e}}$ and the relative position vector $\vec{\boldsymbol{r}}$, it is $\vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{e}} = re\cos\nu$, so that the shape of the orbit in polar coordinates centered in M is given by

$$r = \frac{h^2/\mu}{1 + e\cos\nu}$$

The radius attains its minimum value for $\nu = 0$ and the constant vector $\vec{\boldsymbol{e}}$ is therefore directed from the central body to the periapsis.

1.9 Conic Sections

A conic is the locus of points P such that the ratio of their distance r from a given point F (*focus*) to their distance d from a given line a (*directrix*) is a positive constant e (*eccentricity*), $r/d = e$. Letting s be the distance between the focus and the directrix it is (see Fig. 1.9)

$$d = s - r\cos\nu \quad \Rightarrow \quad r = e\,(s - r\cos\nu)$$

By collecting r on the r.h.s., one gets

$$r = \frac{es}{1 + e\cos\nu}$$

which is the equation of a conic section, written in polar coordinates with the origin at the focus. The parameter $p = se$ is called the *semi–latus rectum*. It is the distance between the focus and the point P of the conic section for $\nu = \pi/2$.

By comparison with the trajectory equation, one deduces that

- in the two-body problem the spacecraft moves along a conic section that has the primary body in its focus; Kepler's first law is demonstrated and extended from ellipses to any type of conics;
- the semi-latus rectum p of the trajectory is related to the angular momentum of the spacecraft ($p = h^2/\mu$)
- the eccentricity of the conic section is the magnitude of $\vec{\boldsymbol{e}}$, which is named *eccentricity vector*.

The name of the curves called conic sections derives from the fact that they can be obtained as the intersections of a plane with a right circular cone. If the plane cuts across a half-cone, the section is an ellipse; one obtains a circle, if the plane is normal to the axis of the cone, and a parabola in the limit case of a plane parallel to a generatrix of the cone. The two branches of a hyperbola are obtained when the plane cuts both the half-cones. Degenerate conic sections (a point, one or two straight lines) arise if the plane passes through the vertex of the cone.

1.10 Relating Energy and Semi-major Axis

One of the constant of the motion, the angular momentum, has been proved to be simply related to the semi-latus rectum. Also the specific energy $\mathcal{E}$ will now be related to a geometrical parameter of the conics.

The constant values of the angular momentum and total energy can be evaluated at any point of the trajectory, in particular at the periapsis, where the spacecraft velocity is orthogonal to the radius vector and

$$h = r_P v_P \; ; \quad \mathcal{E} = \frac{v_P^2}{2} - \frac{\mu}{r_P} = \frac{h^2}{2r_P^2} - \frac{\mu}{r_P}$$

If the orbit is a closed ellipse, the pericentre r_P and the apocentre r_A (*i.e.* the closest and farthest distance of m from the focus where M lies) are given, respectively, by $r_P = p/(1+e)$ and $r_A = p/(1-e)$, so that the semimajor axis a of the ellipse is

$$a = \frac{r_P + r_A}{2} = \frac{p}{1-e^2} \;\Rightarrow\; p = a(1-e^2) = \frac{h^2}{\mu}$$

It is thus possible to write

$$h^2 = \mu a(1-e^2) \quad \text{and} \quad r_P = a(1-e).$$

By substituting these expressions in that of the specific energy at periapsis, one obtains

$$\mathcal{E} = \frac{\mu a(1-e^2)}{2a^2(1-e)^2} - \frac{\mu}{a(1-e)} = \mu\left(\frac{1+e}{2a(1-e)} - \frac{1}{a(1-e)}\right) = -\frac{\mu}{2a}$$

A very simple relationship exists between the specific mechanical energy and the semi-major axis.

The last relationship between geometrical parameter of the conics and the constant of the motion is obtained by extracting the eccentricity e from the above expression of h^2, that is,

$$e = \sqrt{1 - \frac{h^2}{\mu a}} = \sqrt{1 + 2\mathcal{E}\frac{h^2}{\mu^2}}.$$

It should be noted how degenerate conics have zero angular momentum and therefore unit eccentricity.

Appendix: Vector Operations

The result of some vector operations is given in this appendix.

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}} = a^2 \tag{1.2}$$

$$\vec{\boldsymbol{a}} \times \vec{\boldsymbol{a}} = \mathbf{0} \tag{1.3}$$

$$\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} = \vec{\boldsymbol{b}} \times \vec{\boldsymbol{a}} \tag{1.4}$$

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}} = \vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}} \tag{1.5}$$

Equation (1.5) is intuitively proved by considering that either side represents the volume of the parallelepiped the sides of which are given by the vectors $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$, and $\vec{\boldsymbol{c}}$. In a more rigorous way, the scalar triple product of three vectors

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}} = \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

is a pseudoscalar and would reverse sign under inversion of two rows. Therefore

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}} = \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}} \times \vec{\boldsymbol{a}} = \vec{\boldsymbol{c}} \cdot \vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}$$

By equating the first and the last expression, and remembering that the dot product commutes, one gets the result.

$$\vec{\boldsymbol{a}} \cdot \dot{\vec{\boldsymbol{a}}} = a\dot{a} \;\Leftarrow \frac{\mathrm{d}}{\mathrm{d}t}\left(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\left(a^2\right) \tag{1.6}$$

Equation (1.6) is very important in orbital mechanics and rocket propulsion: when a vector elementary increment $\mathrm{d}\vec{\boldsymbol{a}}$ is considered, its component parallel to $\vec{\boldsymbol{a}}$ increases the vector magnitude, whereas the perpendicular component just rotates $\vec{\boldsymbol{a}}$, keeping its magnitude constant.

As for the vector triple product, the following results hold:

$$\left(\vec{\boldsymbol{a}} \times \vec{\boldsymbol{b}}\right) \times \vec{\boldsymbol{c}} = \left(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}}\right)\vec{\boldsymbol{b}} - \left(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}}\right)\vec{\boldsymbol{a}} \tag{1.7}$$

$$\vec{\boldsymbol{a}} \times \left(\vec{\boldsymbol{b}} \times \vec{\boldsymbol{c}}\right) = \left(\vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{c}}\right)\vec{\boldsymbol{a}} - \left(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}}\right)\vec{\boldsymbol{b}} \tag{1.8}$$

Chapter 2

Two-Dimensional Analysis of Motion

2.1 Reference Frames

The angular momentum vector is constant in the two-body problem. This means that the position and velocity vectors always lie on the same plane , perpendicular to $\vec{\boldsymbol{h}}$, *i.e.* the trajectory lies in a fixed *orbit plane* $\mathcal{P}$. A two-dimensional analysis of the motion is usefully carried out in the orbit plane after introducing a non-rotating reference frame $\mathcal{F}_i = (M; \hat{\boldsymbol{n}}, \hat{\boldsymbol{m}}, \hat{\boldsymbol{k}})$, centered in the center of the primary body M, where $\hat{\boldsymbol{n}}$ can be chosen arbitrarily on $\mathcal{P}$, while

$$\hat{\boldsymbol{m}} = \frac{\vec{\boldsymbol{h}} \times \hat{\boldsymbol{n}}}{h} \; ; \quad \hat{\boldsymbol{k}} = \frac{\vec{\boldsymbol{h}}}{h}$$

The distance r from the main body and the anomaly θ define the spacecraft position vector

$$\vec{\boldsymbol{r}} = r\cos\theta\hat{\boldsymbol{n}} + r\sin\theta\hat{\boldsymbol{m}}$$

A rotating frame $\mathcal{F}_o = (m; \hat{\boldsymbol{i}}, \hat{\boldsymbol{j}}, \hat{\boldsymbol{k}})$ is centered on the spacecraft and defined by the right-handed set of unit vectors

$$\hat{\boldsymbol{i}} = \frac{\vec{\boldsymbol{r}}}{r} \; ; \quad \hat{\boldsymbol{j}} = \frac{\vec{\boldsymbol{h}} \times \hat{\boldsymbol{i}}}{h} \; ; \quad \hat{\boldsymbol{k}} = \frac{\vec{\boldsymbol{h}}}{h}.$$

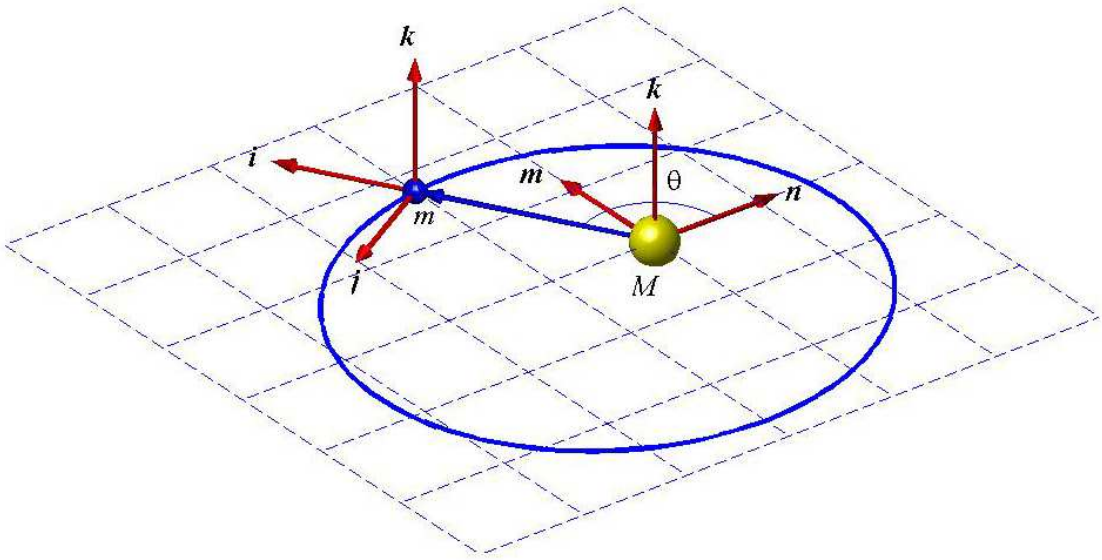

Figure 2.1: **Non–rotating and orbit frames.**

Velocity and acceleration components can be expressed in this frame. From Figure 2.1 it is evident that

$$\hat{\boldsymbol{i}} = \cos\theta\hat{\boldsymbol{n}} + \sin\theta\hat{\boldsymbol{m}} \; ; \quad \hat{\boldsymbol{j}} = -\sin\theta\hat{\boldsymbol{n}} + \cos\theta\hat{\boldsymbol{m}}$$

and, taking the time derivative,

$$\dot{\hat{\boldsymbol{i}}} = \dot{\theta}\hat{\boldsymbol{j}} \; ; \quad \dot{\hat{\boldsymbol{j}}} = -\dot{\theta}\hat{\boldsymbol{i}}$$

2.2 Velocity and acceleration components

The velocity vector

$$\vec{\boldsymbol{v}} = v_r\hat{\boldsymbol{i}} + v_\theta\hat{\boldsymbol{j}}$$

is described by its radial and tangential components, obtained by computing the first derivative of the position vector

$$\vec{\boldsymbol{v}} = \frac{\mathrm{d}\vec{\boldsymbol{r}}}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(r\hat{\boldsymbol{i}}) = \dot{r}\hat{\boldsymbol{i}} + r\dot{\hat{\boldsymbol{i}}} = \dot{r}\hat{\boldsymbol{i}} + r\dot{\theta}\hat{\boldsymbol{j}}$$

One obtains

$$v_r = \dot{r} \; ; \quad v_\theta = r\dot{\theta}$$

In a similar way the derivative of the velocity vector

$$\vec{\boldsymbol{a}} = \frac{\mathrm{d}\vec{\boldsymbol{v}}}{\mathrm{d}t} = \frac{\mathrm{d}^2\vec{\boldsymbol{r}}}{\mathrm{d}t^2} = (\ddot{r} - r\dot{\theta})\hat{\boldsymbol{i}} + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\hat{\boldsymbol{j}}$$

provides the radial and tangential components of the acceleration vector

$$a_r = \ddot{r} - r\dot{\theta} \; ; \quad a_\theta = r\ddot{\theta} + 2\dot{r}\dot{\theta}$$

2.3 First-Order Scalar Equations of Motion

The integration of the first–order scalar equations of motion is required in those cases when, together with the gravity pull from the primary body, there are other forces acting on the spacecraft, such as gravitational perturbations due to other celestial bodies or the thrust of the engine used to modify the orbit. A set of variables suitable for defining the state of the spacecraft is given by the radius r, anomaly θ, radial (v_r) and tangential (v_θ) velocity components. Neglecting other perturbations and assuming that the motion of the spacecraft is controlled by an engine that produces an acceleration $\vec{\boldsymbol{a}}_T$ rotated of an angle ψ away from the radial direction (Fig. 2.3), the vector equation of motion

$$\ddot{\vec{\boldsymbol{r}}} = -\frac{\mu}{r^3}\vec{\boldsymbol{r}} + \vec{\boldsymbol{a}}_T$$

can be split into its radial and tangential components,

$$\begin{aligned} a_r &= \ddot{r} - r\dot{\theta}^2 = -\frac{\mu}{r^2} + a_T\sin\psi \\ a_\theta &= r\ddot{\theta} + 2\dot{r}\dot{\theta} = a_T\cos\psi \end{aligned}$$

Remembering that

$$v_r = \dot{r} \;\Rightarrow\; \dot{v}_r = \ddot{r} \; ; \quad v_\theta = r\dot{\theta} \;\Rightarrow\; \dot{v}_\theta = \dot{r}\dot{\theta} + r\ddot{\theta}$$

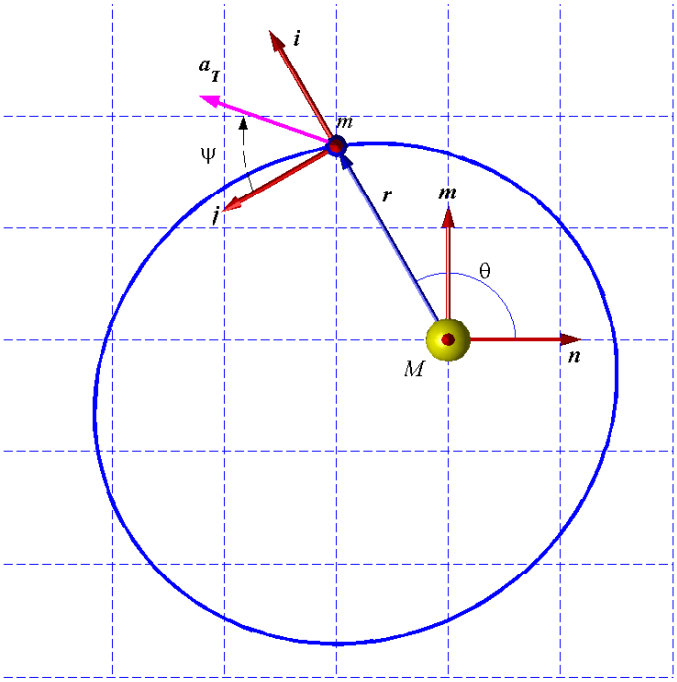

Figure 2.2: **Problem representation in the orbit plane.**

one easily obtains the system of first–order ordinary differential equations describing the evolution of the state of the spacecraft

$$\begin{aligned} \dot{r} &= v_r \\ \dot{\theta} &= \frac{v_\theta}{r} \\ \dot{v}_r &= \frac{v_\theta^2}{r} - \frac{\mu}{r^2} + a_T \sin\psi \\ \dot{v}_\theta &= -\frac{v_r v_\theta}{r} + a_T \cos\psi \end{aligned}$$

The integration can be carried out when one knows the thrust law (*i.e.*, magnitude and direction of the thrust acceleration vector) and four integration constants (*e.g.*, position and velocity components at the initial time). Note that no integration is required for Keplerian trajectories: four constants are sufficient to completely describe the spacecraft motion in the orbit plane.

2.4 Perifocal Reference Frame

The eccentricity vector (*i.e.*, the direction of the semi-major axis) is constant for Keplerian trajectories; as a consequence di angle ω between the unit vector $\hat{\boldsymbol{n}}$ and the periapsis $\hat{p}_1 = \vec{\boldsymbol{e}}/e$ is also constant ($\hat{p}_1 \cdot \hat{\boldsymbol{n}} = \cos\omega$). As a consequence, the *perifocal reference frame*, $\mathcal{F}_P = (S; \hat{\boldsymbol{p}}_1, \hat{\boldsymbol{p}}_2, \hat{\boldsymbol{p}}_3)$, becomes a suitable choice for the non-rotating frame. The related right-handed set of unit vectors is given by

$$\hat{p}_1 = \frac{\vec{\boldsymbol{e}}}{e} \; ; \quad \hat{p}_2 = \frac{\vec{\boldsymbol{h}} \times \vec{\boldsymbol{p}}_1}{h} \; ; \quad \hat{p}_3 = \frac{\vec{\boldsymbol{h}}}{h} = \hat{\boldsymbol{k}}$$

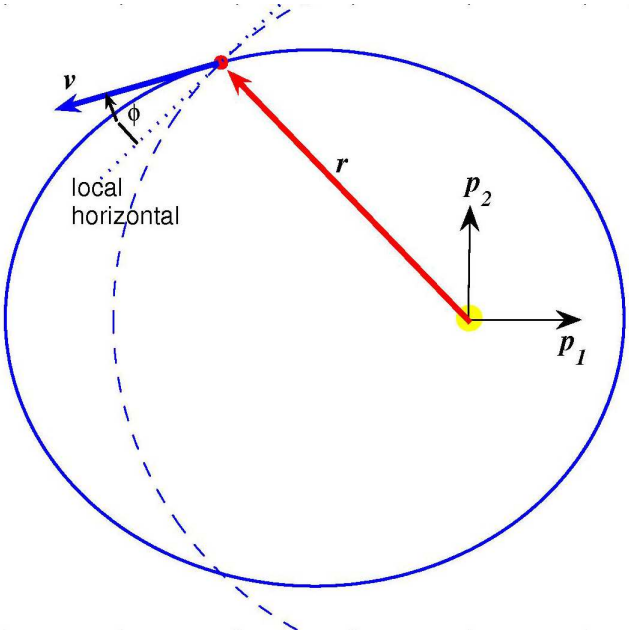

Figure 2.3: **Perifocal frame and flight–path angle.**

where $\hat{p}_1$ and $\hat{p}_2$ lie on the orbit plane $\mathcal{P}$. When $\hat{p}_1$ and $\hat{p}_2$ replace $\hat{n}$ and $\hat{m}$, respectively, all the equations in the previous sections still hold, but the *true anomaly* ν replaces the angle θ:

$$\theta = \nu + \omega \; ; \quad \dot{\nu} = \dot{\theta} \; ; \quad \ddot{\nu} = \ddot{\theta}$$

2.5 Flight Path Angle

By combining the expression for angular momentum with the trajectory equation,

$$h = r v_\theta \; ; \quad r = \frac{h^2/\mu}{1 + e\cos\nu}$$

it is possible to express the tangential velocity as

$$v_\theta = \frac{\mu}{h}(1 + e\cos\nu)$$

Thus the trajectory equation may be written in the form

$$r(1 + e\cos\nu) = p = \text{const.}$$

By taking the time derivative of the above equation, one gets the following equation:

$$\dot{r}(1 + e\cos\nu) - re\dot{\nu}\sin\nu = 0$$

so that the flight path angle φ can be written as

$$\tan\varphi = \frac{v_r}{v_\theta} = \frac{\dot{r}}{r\dot{\nu}} = \frac{e\sin\nu}{1 + e\cos\nu}$$

As a consequence the radial component of the spacecraft velocity can be expressed in the form

$$v_r = v_\theta \tan\varphi = \frac{\mu}{h} e \sin\nu$$

By using the above equations, the velocity components in the perifocal reference frame are given by

$$v_{p_1} = v_r\cos\nu - v_\theta\sin\nu = -\frac{\mu}{h}\sin\nu \; ; \quad v_{p_2} = v_r\sin\nu + v_\theta\cos\nu = \frac{\mu}{h}(e + \cos\nu)$$

2.6 Elliptical Orbits

2.6.1 Geometry of an Elliptical Orbit

An ellipse can be defined as the locus of points such that the sum of their distances from two fixed point, the *focii*, is constant:

$$r_1 + r_2 = 2a.$$

The maximum width of the ellipse is the major axis, and its length is $2a$, while the maximum width at the centre in the direction perpendicular to the major axis is the minor axis, $2b$. The distance between the focii is $2c$.

From the expression of an ellipse's equation in polar coordinates

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

and defining its geometrical characteristics as in Fig. 2.6.1, it is easy to derive the following relations:

$$\begin{aligned} r_P &= p/(1+e) \\ r_A &= p/(1-e) \\ 2a &= r_P + r_A = 2p/(1-e^2) \\ 2c &= r_A - r_P = 2pe/(1-e^2) = 2ae \\ b^2 &= a^2 - c^2 = a^2(1-e^2) = ap \end{aligned}$$

As a consequence, the area of the ellipse

$$A = \pi ab$$

can be rewritten as

$$A = \pi a^2\sqrt{(1-e^2)} = \pi a\sqrt{ap}$$

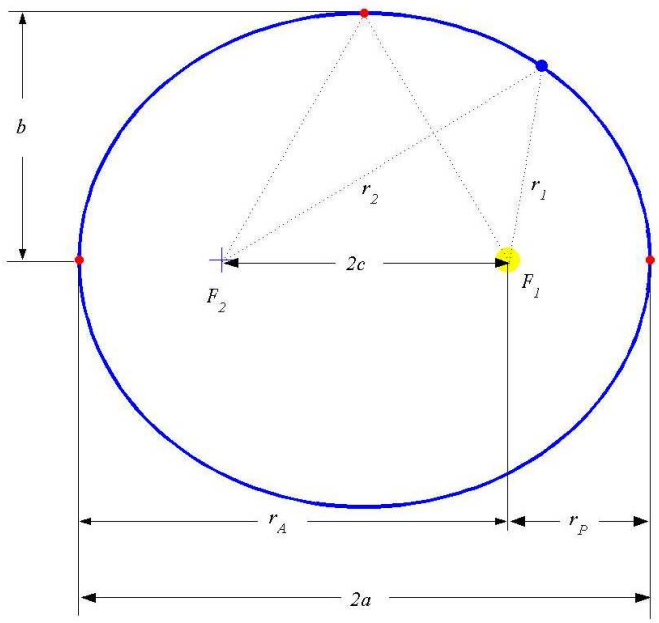

Figure 2.4: **Geometry of an ellipse.**

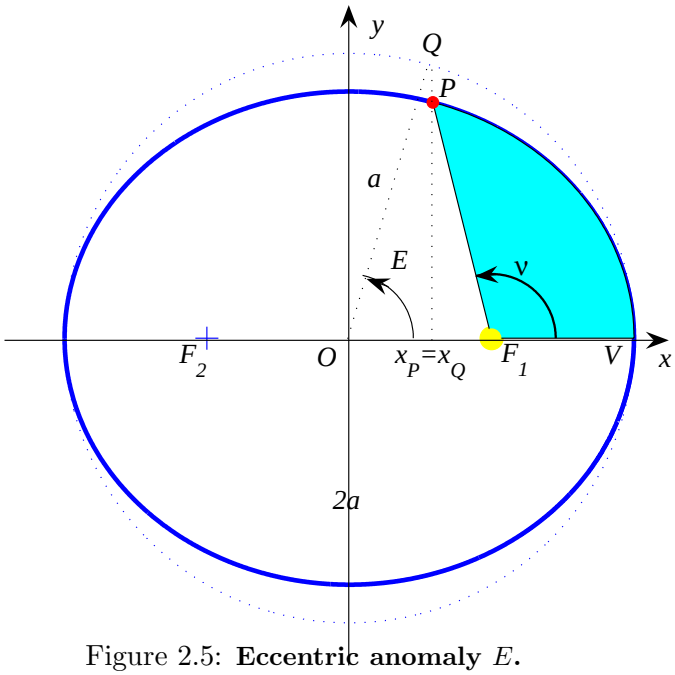

Figure 2.5: **Eccentric anomaly** E**.**

2.6.2 Period of an Elliptical Orbit

The differential element of area swept out by the radius vector as it moves through an angle $\mathrm{d}\nu$ is

$$\mathrm{d}A = \frac{1}{2}r^2\mathrm{d}\nu \quad \Rightarrow \quad \frac{\mathrm{d}A}{\mathrm{d}t} = \frac{1}{2}r^2\dot{\nu} = \frac{1}{2}rv_\theta = \frac{h}{2} = \text{const}$$

thus demonstrating Keplers third law. The constant value of $\mathrm{d}A/\mathrm{d}t$ is evaluated by assuming that the radius vector sweeps out the entire area of the ellipse,

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\pi ab}{\mathcal{T}}$$

which permits to obtain the period $\mathcal{T}$ of an elliptical orbit:

$$\mathcal{T} = \frac{2\pi ab}{h} = 2\pi\sqrt{\frac{a^2b^2}{\mu p}} = 2\pi\sqrt{\frac{a^3}{\mu}}.$$

2.7 Time–of–Flight on the Elliptical Orbit

The time-of-flight t_P from the periapsis V to any point P of the elliptic orbit is reduced to the evaluation of the area $A(FVP)$ swept out by the radius vector,

$$t_P = \mathcal{T}\frac{A(FVP)}{\pi ab}$$

If one introduces an auxiliary circle of radius a, centered in the center of the ellipse O, the point Q on the circle with the same abscissa as the point P is distinguished by its eccentric anomaly E. By using the Cartesian equations of ellipse and circle

$$\frac{x_P^2}{a^2} + \frac{y_P^2}{b^2} = 1 \; ; \quad \frac{x_Q^2}{a^2} + \frac{y_Q^2}{a^2} = 1$$

(where the axes are centered in O, as in Fig. 2.7) one realizes that, for the same abscissa ($x_P = x_Q$) the ratio of the y coordinates y_P/y_Q is constant and equal to b/a (the same conclusion also applies to the segments F_1P and F_1Q). Therefore the shaded area FVP is given by

$$A(FVP) = \frac{b}{a} A(FVQ)$$

The area A(FVQ) is equal to that of the circular sector OVQ minus the triangle OFQ,

$$\begin{aligned} A(FVQ) = A(OVQ) - A(OFQ) &= \frac{E}{2\pi}(\pi a^2) - \frac{1}{2} ca \sin E \\ &= \frac{E}{2} a^2 - \frac{1}{2} e a^2 \sin E \\ &= \frac{a^2}{2} \left(E - e \sin E \right). \end{aligned}$$

As a consequence, the time–of–flight from V to P is

$$t_P = 2\pi \sqrt{\frac{a^3}{\mu}} \frac{1}{\pi ab} \frac{b}{a} \frac{1}{2} (a^2) \left(E - e \sin E \right)$$

that is,

$$t_P = \sqrt{\frac{a^3}{\mu}} \left(E - e \sin E \right) = \frac{1}{n} M_1$$

where, according to Kepler, the *mean motion* n and the *mean anomaly* M are defined as

$$n = \sqrt{\frac{\mu}{a^3}} \; ; \quad M = E - e \sin E.$$

The angle E is related to the actual spacecraft position by the equation

$$a \cos E = ea + r \cos \nu$$

and by replacing either r or ν, one obtains, respectively,

$$\cos E = \frac{a - r}{ea} \; ; \quad \cos E = \frac{e + \cos \nu}{1 + e \cos \nu}$$

The correct quadrant for E is obtained by considering that ν and E are always in the same half-plane: when ν is between 0 and π, so is E.

For the evaluation of the time-of-flight to point P_2 from some general point P_1, which is not periapsis, the following relation holds:

$$t_{1\to 2} = t_{P_1} - t_{P_2} = n(M_2 - M_1)$$

that is,

$$t_{1\to 2} = \sqrt{\frac{a^3}{\mu}} \left[E_2 - E_1 - e \left(\sin E_2 - \sin E_1 \right) \right]$$

Some care is required if the spacecraft passes k times through periapsis, as $2k\pi$ must be added to E_2.

2.8 Extension to hyperbola and parabola

The time–of–flight t_P from the periapsis V to any point P of a hyperbolic trajectory can be expressed as

$$t_P = \sqrt{-\frac{a^3}{\mu}}\,(e\sinh F - F)$$

where the *hyperbolic eccentric anomaly* F can be related to either r or ν,

$$\cosh F = \frac{a-r}{ea}\ ;\quad \cosh F = \frac{e+\cos\nu}{1+e\cos\nu}$$

Note how the above definitions of F closely resemble those of E, with the substitution of the goniometric one with the corresponding hyperbolic one.

The time–of–flight on a parabolic trajectory is

$$t_P = \frac{1}{2\sqrt{\mu}}\left(pD + \frac{1}{3}D^3\right)$$

where the *parabolic eccentric anomaly* D is defined as

$$D = \sqrt{p}\tan\left(\frac{\nu}{2}\right)$$

2.9 Circular and Escape Velocity, Hyperbolic Excess Speed

On a circular orbit, the radius $r = r_c =$ is constant. Dot multiplication of the vector equation of motion by the radial unit vector $\hat{\boldsymbol{\imath}}$ provides

$$a_r = -r\dot{\theta}^2 = -\frac{\mu}{r_c^2}$$

The *circular velocity* v_c is the speed necessary to place a spacecraft on a circular orbit, provided that a correct direction (*i.e.* $v_r = 0$) is obtained. In this case it is

$$v_c = \sqrt{\frac{\mu}{r_c}}$$

Note that the greater the radius of the circular orbit, the smaller the speed required to keep the spacecraft on this orbit.

The speed which is just sufficient to allow an object coasting to an infinite distance from the intial distance r_0 is called *escape velocity* v_e. The specific mechanical energy

$$\mathcal{E} = -\frac{\mu}{2a} = \frac{v^2}{2} - \frac{\mu}{r} = \frac{v_e^2}{2} - \frac{\mu}{r_0} = 0$$

must be zero, as $v \to 0$ at $r \to \infty$. The above equation implies $a = \infty$, so that the minimum-energy escape trajectory is a parabola. One obtains

$$v_e = \sqrt{2\frac{\mu}{r_0}}$$

On the same trajectory the escape velocity is not constant: the farther away the spacecraft is from the central body, the less speed it takes to escape the remainder of the gravitational field.

If a spacecraft has a greater speed than the escape velocity, the residual speed at a very large distance from the central body is defined *hyperbolic excess velocity* v_∞. Its value is obtained from the specific mechanical energy

$$\mathcal{E} = -\frac{\mu}{2a} = \frac{v^2}{2} - \frac{\mu}{r} = \frac{v_\infty^2}{2}$$

2.10 Cosmic Velocities

Cosmic velocities are the theoretical minimum speeds that a spacecraft must reach on leaving the ground and

1. orbit the Earth, $v_{\rm I}$;
2. escape its gravitational pull, $v_{\rm II}$;
3. leave the solar system, $v_{\rm III}$;
4. hit the Sun, $v_{\rm IV}$.

The first and second cosmic velocities correspond, respectively, to the circular and escape velocities, computed using the gravitational parameter of the Earth ($\mu_E = 398600$ km^{3}s^{-2}2) and the Earth's mean radius ($r_E = 6371$ km),

$$v_{\rm I} = \sqrt{\frac{\mu_E}{r_E}} = 7.91 \text{ km s}^{-1} \; ; \quad v_{\rm II} = \sqrt{2\frac{\mu_E}{r_E}} = 11.18 \text{ km s}^{-1}$$

The circular and escape velocities are then computed using the solar gravitational parameter ($\mu_S = 1.327\ 10^{11}$ km^{3}s^{-2}) and the mean Sun-Earth distance ($R_E = 149.5\ 10^6$ km). The Earth's mean velocity while orbiting the Sun, v_E, and the escape velocity from the Sun, $v_{\rm esc}$, for an initial distance equal to the Earth's orbit mean radius, R_E, are given, respectively, by

$$v_E = \sqrt{\frac{\mu_S}{R_E}} = 29.79 \text{ km s}^{-1} \; ; \quad v_{\rm esc} = \sqrt{2\frac{\mu_S}{R_E}} = 42.13 \text{ km s}^{-1}$$

The former is an averaged value of the Earth heliocentric velocity, which is almost constant as the Earth's orbit eccentricity is very small. The latter is the heliocentric velocity just sufficient to escape the solar attraction, for a spacecraft which has left the sphere of influence of the Earth, but is yet at distance R_E from the Sun. As the spacecraft heliocentric velocity is the sum of the vectors $\vec{\boldsymbol{v}}_E$ (velocity of the Earth with respect to the Sun) and $\vec{\boldsymbol{v}}_\infty$ (hyperbolic excess speed of the spacecraft when leaving the sphere of influence of the Earth), these are parallel, if the escape from the Sun is achieved along the tangent to the Earth's orbit in the same direction as the Earth's motion, that is

$$v_{\rm esc} = v_E + v_\infty \;\Rightarrow\; v_\infty = v_{\rm esc} - v_E = 12.34 \text{ km s}^{-1}$$

with the minimum velocity relative to the Earth. The corresponding velocity on leaving the ground (i.e., the third cosmic velocity) is obtained using energy conservation,

$$\frac{v_{\rm III}^2}{2} - \frac{\mu}{r_E} = \frac{v_\infty^2}{2} \;\Rightarrow\; v_{\rm III} = \sqrt{v_\infty^2 + v_{\rm II}^2} = 16.65 \text{ km s}^{-1}.$$

A similar procedure is used to evaluate the fourth cosmic velocity, where in the heliocentric reference frame the spacecraft must be motionless when leaving the sphere of influence of the Earth in order to hit the theoretically point-mass Sun falling down on it along the vertical in the heliocentric system. Assuming a launch along the tangent to the Earth's orbit in the opposite direction with respect to the Earth's motion, one obtains

$$v_{\rm esc} = 0 = v_\infty - v_E \quad \Rightarrow \quad v_\infty = v_E = 29.79 \text{ km s}^{-1}$$

so that

$$v_{\rm IV} = \sqrt{v_\infty^2 + v_{\rm II}^2} = 31.82 \text{ km s}^{-1}.$$

Table 2.1: **Orbit types**

v	$\mathcal{E}$	a	v_∞	e	**shape**
$< v_e$	< 0	> 0	—	< 1	ellipse
$= v_e$	$= 0$	∞	0	$= 1$	parabola
$> v_e$	> 0	< 0	> 0	> 1	hyperbola

Appendix: Orbit types

The type of conic orbit is strictly related to the energetic and geometrical parameters. An analysis of the following equations

$$\mathcal{E} = \frac{v^2}{2} - \frac{\mu}{r} = \frac{v^2}{2} - \frac{v_e^2}{2} = -\frac{\mu}{2a} = \frac{v_\infty^2}{2} \; ; \quad e = \sqrt{1 + 2\mathcal{E}\frac{h^2}{\mu^2}}$$

produces the conditions presented in the Tab. 2.1. The fulfillment of any of the reported conditions is a sure sign of orbit type.

By increasing the spacecraft velocity, the trajectory changes in a regular way. Assume that a spacecraft is launched horizontally with velocity v_L from the surface of an atmosphere–free Earth. A trajectory external to the surface is impossible before the launch velocity reaches the first cosmic velocity v_I; the trajectory would be an internal ellipse, with the launch point as the apocenter, if the Earth had its mass concentrated in the center point. Starting from a degenerate ellipse ($e = 1$, $a = r_E/2$), the eccentricity progressively decreases as $v_L is increased$ until a circular orbit around the Earth for $r = r_E$ is achievable for $v_L = v_\mathrm{I}$. By increasing v_L further, the orbit is again elliptical, but the launch point is at the pericenter. The larger v_L, the greater is the eccentricity, until for $v_L = v_\mathrm{II}$ the spacecraft in injected into a parabolic trajectory. The following trajectories are hyperbolae of increasing energy, until, for infinite launch velocity, the spacecraft moves on a degenerate hyperbola ($e = \infty$), *i.e.*, a straight line tangent to the Earth surface. In all this sequence, the semi-latus rectum p increases in a very regular way.

Chapter 3

Three-Dimensional Analysis of Motion

A space mission occurs in a three-dimensional environment and the spacecraft motion must be analyzed accordingly. At the same time, it is necessary to define the measurement of time in an unambiguous way, as many phenomena related to space travel happen on very different time–scales (minutes to hours for manoeuvres, days to months for orbit perturbations, several years for the operational life of the satellite). This chapter is devoted to the presentation of time and space reference frames.

3.1 The Measurement of Time

The *sidereal day* is the time D_S required for the Earth to rotate once on its axis relative to the "fixed stars". The time between two successive transits of the Sun across the same meridian is called *apparent solar day*. Two solar days have not exactly the same length because the Earth's axis is not exactly perpendicular to the ecliptic plane and the Earth's orbit is elliptic. The Earth has to turn slightly more than a complete rotation relative to the fixed stars, as the Earth has traveled 1/365th of the way on its orbit in one day.

The mean solar day (24 h or 86400 s) is defined by assuming that the Earth is on a circular orbit with the same period as the actual one, and its axis is perpendicular to the orbit plane.

The constant length of the sidereal day is obtained by considering that during a complete orbit around the Sun (1 year) the Earth performs one revolution more about its axis relative to the fixed stars than it doeas with respect to the Sun. One obtains that $D_S = 23^{\mathrm{h}}\ 56^{\mathrm{m}}\ 04^{\mathrm{s}} = 86164$ s. Therefore the angular velocity of the Earth motion around its axis is $\omega_\oplus = 2\pi/D_S = 2\pi \cdot 1.0027379$ rad/day $= 7.2921 \cdot 10^{-5}$ rad s^{-1}.

The local mean solar time on the Greenwich meridian is called Greenwich Mean Time (GMT), Universal Time (UT), or Zulu (Z) time (slight differences in their definition are here neglected).

The Julian calendar was introduced by Julius Caesar in 46 BC in order to approximate the tropical year and be synchronous with the seasons. The Julian calendar consisted of cycles of three 365–day years followed by a 366–day leap year. Hence the Julian year had on average 365.25 days; nevertheless it was a little too long, causing the vernal equinox to slowly drift backwards in the calendar year.

The Gregorian calendar, which presently is used nearly everywhere in the world, was decreed by Pope Gregory XIII, from whom it was named, on February 24$^{\mathrm{th}}$, 1582, in order to better approximate the length of a solar year, thus ensuring that the vernal equinox would be constantly close to a specific date. The calendar is based on a cycle of 400 years comprising 146,097 days; leap years are every fourth year, but are omitted in years divisible by 100 but not divisible by

400, giving a year of average length 365.2425 days. This value is very close to the 365.2424 days of the vernal equinox year (which is shorter than the sidereal year because of the vernal equinox precession). The last day of the Julian calendar was October 4$^{\text{th}}$, 1582 and this was followed by the first day of the Gregorian calendar, October 15$^{\text{th}}$, 1582. The deletion of ten days was not strictly necessary, but had the purpose of locating the vernal equinox on March 21$^{\text{st}}$.

The Julian day or Julian day number (JD) was introduced to map the temporal sequence of days onto a sequence of integers. The Julian day number makes the number of days between two given dates easier to determine, as it is sufficient to subtract one Julian day number from the other. The Julian day is the number of days that have elapsed since 12 noon GMT on Monday, January 1st, 4713 BC in the proleptic (i.e., extrapolated) Julian calendar. It should be noted that for astronomers a "day" begins at noon, also according to tradition, as midnight could not be accurately determined, before clocks. Also note that 4713 BC is -4712 using the astronomical year numbering, that has year 0, whereas Gregorian calendar directly passes from 1 BC to 1 AD.

The day from noon on January 1st, 4713 BC to noon on January 2nd, 4713 BC is counted as Julian day zero. The astronomical Julian date provides a complete measure of time by appending to the Julian day number the fraction of the day elapsed since noon (for instance, .25 means 18 o'clock).

Given a Julian day number JD, the modified Julian day number MJD is defined as MJD = JD − 2,400,000.5. This has two purposes:

- days begin at midnight rather than noon;
- for dates in the period from 1859 to about 2130 only five digits need to be used to specify the date rather than seven.

JD 2,400,000.5 designates the midnight between November 16$^{\text{th}}$ and 17$^{\text{th}}$, 1858; so day 0 in the system of modified Julian day numbers is November 17th, 1858.

The Julian day number (JD), which starts at noon UT on a specified date (D, M, Y) of the Gregorian or Julian calendar, can be computed using the following procedure (all divisions are integer divisions, in which remainders are discarded; astronomical year numbering is used, *e.g.*, 10 BC = -9). After computing

$$a = \frac{14 - M}{12} \; ; \quad y = Y + 4800 - a \; ; \quad m = M + 12a - 3$$

the Julian day number for a date in the Gregorian calendar is given by

$$\text{JD} = D + \frac{153m + 2}{5} + 365y + \frac{y}{4} - \frac{y}{100} + \frac{y}{400} - 32045 \; ;$$

for a date in the Julian calendar it is

$$\text{JD} = D + \frac{153m + 2}{5} + 365y + \frac{y}{4} - 32083.$$

To convert the other way round, for the Gregorian calendar compute

$$a = \text{JD} + 32044 \; ; \quad b = \frac{4a + 3}{146097} \; ; \quad c = a - \frac{146097b}{4}$$

or, for a date in the Julian calendar,

$$b = 0 \; ; \quad c = \text{JD} + 32082.$$

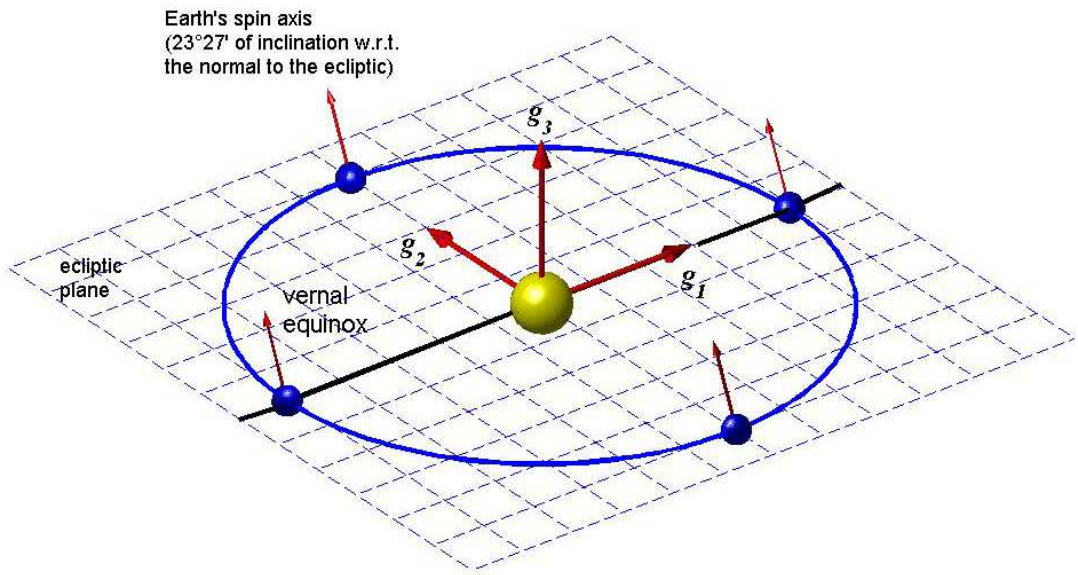

Figure 3.1: **Heliocentric frame.**

Then, for both calendars,

$$d = \frac{4c+3}{1461} \; ; \quad e = c - \frac{1461d}{4} \; ; \quad m = \frac{5e+2}{153}$$

and, finally,

$$D = e - \frac{153m+2}{5} + 1 \; ; \quad M = m + 3 - 12\frac{m}{10} \; ; \quad Y = 100b + d - 4800 + \frac{m}{10}.$$

3.2 Non–rotating Frames

In order to describe a space mission, it is necessary to provide, together with the time–frame for the sequence of events and evolution of the orbit, a suitable reference system and a corrisponent set of coordinates, that suits the application.

The first requirement for the spatial description of an orbit is a suitable (and at least approximately inertial) reference frame. A frame centered in the primary body (*i.e.* the most massive body in the system), with axes pointing a fixed direction with respect to the far stars (thus non–rotating) usually does well the job.

For trajectories around the Sun, the *heliocentric frame* $\mathcal{F}_H = (S; \hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \hat{\boldsymbol{g}}_3)$ (Fig. 3.2), with the x and y axes on the *ecliptic plane* (that is, the plane of the Earth's orbit about the Sun) and the z axis perpendicular to it, in the direction of the Earth's angular momentum, is an obvious choice. For Earth orbiting satellites, a *geocentric frame* $\mathcal{F}_G = (O; \hat{\boldsymbol{g}}_1, \hat{\boldsymbol{g}}_2, \hat{\boldsymbol{g}}_3)$ centered in the center of the Earth and based on the *equatorial plane* is preferred (Fig. 3.2), the equatorial plane being perpendicular to the Earth's spin axis. For both systems the third unit vector, namely $\hat{\boldsymbol{g}}_3$, is parallel to the z-axes and is therefore defined; its direction is towards the North. The x-axis is common to both system and is parallel to the *equinox line*, which is the line of intersection of the two fundamental planes, the ecliptic and the equatorial plane; its positive direction is from the Earth to the Sun on the first day of spring or the vernal equinox (see Fig. 3.2). The unit vector $\hat{\boldsymbol{g}}_1$ points towards the constellation Aries (the ram).

The Earth's axis of rotation actually exhibits a slow precession motion and the x–axis shifts westward with a period of 26000 years (a superimposed oscillation with a period of 18.6 years

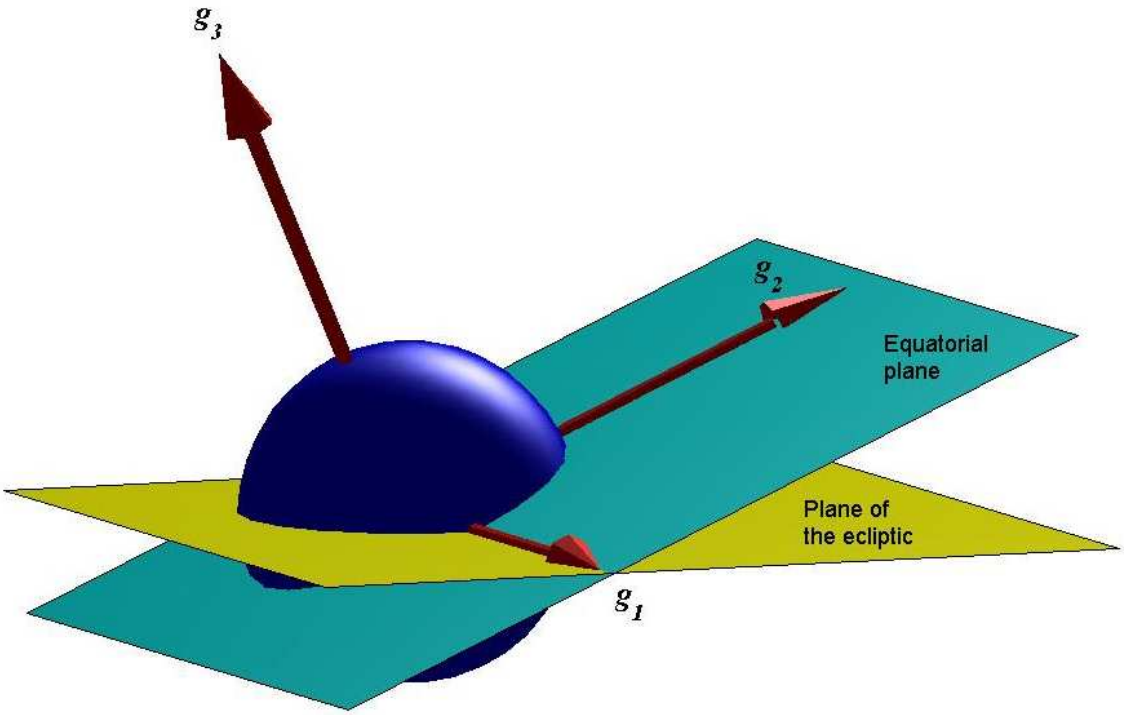

Figure 3.2: **Geocentric frame.**

derives from the nutation of the Earth's axis, which is due to the variable inclination of the lunar orbit on the ecliptic plane). For computations, the definition of both systems is based on the direction of the line–of–intersection at a specified date or *epoch*. In astronomy, an epoch is a moment in time for which celestial coordinates or orbital elements are specified; the current standard epoch is J2000.0, which is January 1st, 2000 at 12:00. The remaining unit vectors for both reference frames, named $\hat{\boldsymbol{g}}_2$, are uniquely defined by the requirement for both frames to be orthogonal and right–handed.

The spacecraft position can be described by the Cartesian components of its position vector with respect to the origin of the reference frame; the use of the distance ρ from the center and two angles is usually preferred. The declination δ is measured northward from the x–y plane; the right ascension α is measured on the fundamental plane, eastward from the vernal equinox direction.

We remind that once an orbiting object is dealt with, it is easy to express its position and velocity vectors in the perifocal frame $\mathcal{F}_P = (S; \hat{\boldsymbol{p}}_1, \hat{\boldsymbol{p}}_2, \hat{\boldsymbol{p}}_3 = \hat{\boldsymbol{k}})$, introduced in the previous Chapter. $\mathcal{F}_P$ is also a non–rotating, but its orientation depends on the particular orbit considered.

The issue of coordinate transformation will be dealt with in the Section 3.7, where the more general problem of the coordinate transformation is presented for the specific case of passage from perifocal to geocentric–equatorial components.

3.3 Classical Orbital Elements

A Keplerian trajectory is uniquely defined by 6 parameters. The initial values for the integration of the second-order vector equation of motion, *i.e.* the spacecraft position and velocity at epoch, could be used, but a different set of parameters, which provides an immediate description of the

trajectory, is preferable. The *classical orbital elements* are widely used for this purpose. Only 4 elements are necessary in the two–dimensional problem: three parameters describe size, shape, and direction of the line of apsides; the fourth is required to pinpoint the spacecraft position along the orbit at a particular time. The remaining 2 parameters describe the orientation of the orbital plane.

The classical orbital elements are

1. *eccentricity* e (shape of the orbit);
2. *semi–major axis* a or, equivalently, *semi–latus rectum* p (size of the orbit);
3. *inclination* i, that is, the angle between the third unit vector $\hat{\boldsymbol{g}}_3$ and the angular momentum $\vec{\boldsymbol{h}}$ (inclination of the orbit plane with respect to the base plane of the non–rotating reference frame);
4. *longitude of the ascending node*, Ω, that is, the angle in the fundamental plane measured eastward from the first coordinate axis of the non–rotating frame $\hat{\boldsymbol{g}}_1$ to the *ascending node*, *i.e.* the point where the spacecraft crosses the fundamental plane while moving in the northerly direction (orientation of the orbit plane);
5. *argument of periapsis* ω, that is the angle in the orbit plane between the ascending node and the periapsis, measured in the direction of the spacecraft motion (pericenter direction in the orbit plane);
6. *true anomaly* ν_0 at a particular time t_0 or epoch (spacecraft position); it is sometime replaced by *time of periapsis passage* T.

Some of the above parameters are not defined when either inclination or eccentricity are zero. Alternate parameters are

- *longitude of periapsis* $\Pi = \Omega + \omega$, which is defined when $i = 0$ and the ascending node is not defined;
- *argument of latitude at epoch* $u_0 = \omega + \nu_0$, which is defined when the orbit is circular, $e = 0$, and the periapsis is not defined;
- *true longitude at epoch* $\ell_0 = \Omega + \omega + \nu_0 = \Pi + \nu_0 = \Omega + u_0$, which remains defined when either $i = 0$ or $e = 0$.

3.4 Modified Equinoctial Orbital Elements

It should be noted that geostationary orbits are circular and equatorial, which means that both the eccentricity and the inclination are zero. One possible alternative choice is to use the *modified equinoctial orbital elements*. This is a set of orbital elements that are always univocally defined, for circular, elliptic, and hyperbolic orbits with any inclination, as the modified equinoctial equations exhibit no singularity for zero eccentricity and orbital inclinations equal to 0 and 90 degrees.

However, two of the components are singular for an orbital inclination of 180 degrees, that is, for *retrograde orbits*, that is, equatorial orbits flown in the direction opposite to Earth's spin motion. As these orbits are seldom used, this singularity does not represent a serious problem.

Relationship between modified equinoctial and classical orbital elements is given by the following set of equations:

1. $p = a(1 - e^2)$;

2. $f = e\cos(\omega + \Omega)$;
3. $g = e\sin(\omega + \Omega)$;
4. $h = \tan(i/2)\cos\Omega$;
5. $k = \tan(i/2)\sin\Omega$;
6. $\ell_0 = \Omega + \omega + \nu_0$.

The inverse relations that allow to recover the classical orbital elements from the modified ones are as follows:

semimajor axis: $a = \dfrac{p}{1 - f^2 - g^2}$

eccentricity: $e = \sqrt{f^2 + g^2}$

inclination: $i = 2\tan^{-1}\left(\sqrt{h^2 + k^2}\right)$
$= \tan^{-1}\left(2\sqrt{h^2 + k^2}, 1 - h^2 - k^2\right)$

argument of periapsis: $\omega = \tan^{-1}(g/f) - \tan^{-1}(k/h)$
$= \tan^{-1}(gh - fk, fh + gk)$

right ascension of the ascending node: $\Omega = \tan^{-1}(k, h)$

true anomaly at epoch: $\nu_0 = \ell_0 - \Omega - \omega$

argument of latitude at epoch: $u_0 = \omega + \nu_0$
$= \tan^{-1}(h\sin\ell_0 - k\cos\ell_0, h\cos\ell_0 + k\sin\ell_0)$

In the above equations the expression $\tan^{-1}(a, b)$ indicates the calculation of the four quadrant inverse tangent.

3.5 Determining the orbital elements

The orbital elements are easily found starting from the knowledge of the Cartesian components of position and velocity vectors $\vec{\boldsymbol{r}}$ and $\vec{\boldsymbol{v}}$ at a particular time t_0 in either non–rating reference frame defined in Section 3.2. One preliminarily computes the components of the constant vectors

$$\vec{\boldsymbol{h}} = \vec{\boldsymbol{r}} \times \boldsymbol{v} \quad \text{and} \quad \vec{\boldsymbol{e}} = \frac{\vec{\boldsymbol{v}} \times \vec{\boldsymbol{h}}}{\mu} - \frac{\vec{\boldsymbol{r}}}{r}$$

and therefore the unit vectors

$$\hat{\boldsymbol{k}} = \frac{\vec{\boldsymbol{h}}}{h} \; ; \quad \hat{\boldsymbol{i}} = \frac{\vec{\boldsymbol{r}}}{r} \; ; \quad \hat{\boldsymbol{n}} = \frac{\vec{\boldsymbol{g}}_3 \times \vec{k}}{||\vec{\boldsymbol{g}}_3 \times \vec{k}||} \; ; \quad \hat{\boldsymbol{p}}_1 = \frac{\vec{\boldsymbol{e}}}{e}$$

These vectors, centered in the center of mass of the primary body, define respectively the directions normal to the orbit plane, and towards the spacecraft, the ascending node, and the pericenter, respectively.

The orbital elements are thus given by

1. $p = h^2/\mu$
2. $e = ||\hat{\boldsymbol{e}}||$
3. $\cos i = \hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{g}}_3 = k_3$
4. $\cos\Omega = \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{g}}_1 = n_3$ $\quad (\Omega > \pi, \text{ if } n_2 < 0)$
5. $\cos\omega = \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{p}}_1$ $\quad (\omega > \pi, \text{ if } e_3 < 0)$
6. $\cos\nu_0 = \hat{\boldsymbol{i}} \cdot \hat{\boldsymbol{p}}_1$ $\quad (\nu_0 > \pi, \text{ if } \vec{\boldsymbol{r}} \cdot \vec{\boldsymbol{v}} < 0)$

In a similar way one evaluates the alternate parameters

- $\cos u_0 = \hat{\boldsymbol{i}} \cdot \hat{\boldsymbol{n}}$ $\quad (u_0 > \pi$, if $i_3 < 0)$
- $\cos \Pi = \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{g}}_1$ $\quad (Pi > \pi$, if $e_2 < 0)$
- $\cos \ell_0 = \hat{\boldsymbol{i}} \cdot \hat{\boldsymbol{g}}_1$ $\quad (\ell_0 > \pi$, if $i_2 < 0)$

The last two equations hold only for zero orbit inclination ($i = 0$).

3.6 Determining Spacecraft Position and Velocity

After the orbital elements have been obtained from the knowledge of $\vec{\boldsymbol{r}}$ and $\vec{\boldsymbol{v}}$ at a specified time, the problem of updating the spacecraft position and velocity is solved in the perifocal reference frame using the closed–form solution of the equation of motion. The position and velocity components for a specific value of the true anomaly ν are found using the equations presented in Section 2.5. The procedure in Section 2.7 permits the evaluation of the time of passage at the selected anomaly. The $\vec{\boldsymbol{r}}$ and $\vec{\boldsymbol{v}}$ components in the geocentric-equatorial (or heliocentric-ecliptic) frame can be obtained using a coordinate transformation.

3.7 Coordinate Transformation

The more general problem of coordinate transformation is presented here for the specific case of passage from perifocal to geocentric–equatorial components, but the approach can be extended to change of reference between any pair of mutually orthogonal triads of unit vectors in a three–dimensional space.

The same vector can be expressed in terms of components in different frames (*e.g.* $\mathcal{F}_P$ and $\mathcal{F}_G$) as

$$\begin{aligned} \vec{\boldsymbol{v}} &= v_{g1}\hat{\boldsymbol{g}}_1 + v_{g2}\hat{\boldsymbol{g}}_2 + v_{g3}\hat{\boldsymbol{g}}_3 \\ &= v_{p1}\hat{\boldsymbol{p}}_1 + v_{p2}\hat{\boldsymbol{p}}_2 + v_{p3}\hat{\boldsymbol{p}}_3 \end{aligned}$$

where $\boldsymbol{v}_P = (v_{p1}, v_{p2}, v_{p3})^T$ indicates the components of $\vec{\boldsymbol{v}}$ in $\mathcal{F}_P$ and $\boldsymbol{v}_G = (v_{g1}, v_{g2}, v_{g3})^T$ indicates the components of the same vector in $\mathcal{F}_P$. The characteristics of the vector (magnitude and direction) are not affected by the change of frame in which its component are represented.

Taking the dot product with the unit vector $\hat{\boldsymbol{g}}_1$ provides the component

$$v_{g1} = v_{p1}\hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{g}}_1 + v_{p2}\hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{g}}_1 + v_{p3}\hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{g}}_1.$$

By repeating the same procedure with $\hat{\boldsymbol{g}}_2$ and $\hat{\boldsymbol{g}}_3$, one obtains the transformation matrix

$$\begin{pmatrix} v_{g1} \\ v_{g2} \\ v_{g3} \end{pmatrix} = \begin{bmatrix} \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{g}}_1 & \hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{g}}_1 & \hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{g}}_1 \\ \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{g}}_2 & \hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{g}}_2 & \hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{g}}_2 \\ \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{g}}_3 & \hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{g}}_3 & \hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{g}}_3 \end{bmatrix} \begin{pmatrix} v_{p1} \\ v_{p2} \\ v_{p3} \end{pmatrix}$$

or, in more compact form

$$\boldsymbol{v}_G = \boldsymbol{T}_{GP}\boldsymbol{v}_P$$

where $\boldsymbol{T}_{GP}$ is the *coordinate transformation matrix*. Quite obviously it is also

$$\boldsymbol{v}_P = \boldsymbol{T}_{PG}\boldsymbol{v}_G = (\boldsymbol{T}_{GP})^{-1}\boldsymbol{v}_G$$

but by applying the procedure used to obtain $\boldsymbol{T}_{GP}$ to the inverse transformation from geocentric–equatorial to perifocal components, the inverse matrix ${\boldsymbol{T}_{GP}}^{-1}$ is simply given by the transpose

matrix $\boldsymbol{T}_{GP}{}^T$. This is however a general property of the transformations between orthogonal basis.

Each element of $\boldsymbol{T}_{GP}$ is the dot product between two unit vectors, that is, the cosine of the angle between them, hence the name *cosine matrix* often used to define the coordinate transformation matrix $\boldsymbol{T}_{GP}$. As an example, T_{11} can be evaluated by means of the law of cosines (see Appendix) for the sides of the spherical triangle defined by the unit vectors $\hat{\boldsymbol{n}}$, $\hat{\boldsymbol{g}}_1$, and $\hat{\boldsymbol{p}}_1$, that provides

$$T_{11} = cos\Omega \cos\omega - \sin\Omega \sin\omega \sin i$$

The procedure is however cumbersome. It is better to split the transformation into a sequence of three *elementary rotations* (that is, rotations about one of the coordinate axes), by introducing two intermediate (auxiliary) reference frames:

- $\mathcal{F}_{G'} = (\hat{\boldsymbol{g}}'_1, \hat{\boldsymbol{g}}'_2, \hat{\boldsymbol{g}}'_3)$, based on the equatorial plane, with $\hat{\boldsymbol{g}}'_1 = \hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{g}}'_3 = \hat{\boldsymbol{g}}_3$;
- $\mathcal{F}_{P'} = (\hat{\boldsymbol{p}}'_1, \hat{\boldsymbol{p}}'_2, \hat{\boldsymbol{p}}'_3)$, based on the orbital plane, with $\hat{\boldsymbol{p}}'_1 = \hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{p}}'_3 = \hat{\boldsymbol{p}}_3$.

The rotation angle for each elementary rotation is referred to as an Euler angle. In the present case, starting from the geocentric–equatorial frame, the first rotation that takes $\mathcal{F}_G$ onto $\mathcal{F}_{G'}$ (that is, the unit vector $\hat{\boldsymbol{g}}_1$ onto $\hat{\boldsymbol{g}}'_1 = \hat{\boldsymbol{n}}$) is about the third axis, $\hat{\boldsymbol{g}}_3$ and its amplitude is Ω. The coordinate transformation is expressed in matrix form as

$$\boldsymbol{v}_G = \boldsymbol{T}_{GG'} \boldsymbol{v}_{G'} = \boldsymbol{T}' \boldsymbol{v}_{G'}$$

where the coordinate transformation matrix is easily obtained,

$$\boldsymbol{T}_{GG'} = \boldsymbol{T}' = \boldsymbol{T}_3(\Omega) = \begin{bmatrix} \hat{\boldsymbol{g}}'_1 \cdot \hat{\boldsymbol{g}}_1 & \hat{\boldsymbol{g}}'_2 \cdot \hat{\boldsymbol{g}}_1 & \hat{\boldsymbol{g}}'_3 \cdot \hat{\boldsymbol{g}}_1 \\ \hat{\boldsymbol{g}}'_1 \cdot \hat{\boldsymbol{g}}_2 & \hat{\boldsymbol{g}}'_2 \cdot \hat{\boldsymbol{g}}_2 & \hat{\boldsymbol{g}}'_3 \cdot \hat{\boldsymbol{g}}_2 \\ \hat{\boldsymbol{g}}'_1 \cdot \hat{\boldsymbol{g}}_3 & \hat{\boldsymbol{g}}'_2 \cdot \hat{\boldsymbol{g}}_3 & \hat{\boldsymbol{g}}'_3 \cdot \hat{\boldsymbol{g}}_3 \end{bmatrix} = \begin{bmatrix} \cos\Omega & -\sin\Omega & 0 \\ \sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

since $\hat{\boldsymbol{g}}_3 = \hat{\boldsymbol{g}}'_3$ are perpendicular to the first and second unit vectors of both frames and the dot products $\hat{\boldsymbol{g}}_1 \cdot \hat{\boldsymbol{g}}'_1 = \hat{\boldsymbol{g}}_2 \cdot \hat{\boldsymbol{g}}'_2 = \cos\Omega$. Finally, the angle between $\hat{\boldsymbol{g}}'_1$ and $\hat{\boldsymbol{g}}_2$ is $\pi/2 - \Omega$, such that $\cos(\pi/2 - \Omega) = \sin\Omega$, while that between $\hat{\boldsymbol{g}}'_2$ and $\hat{\boldsymbol{g}}_1$ is $\pi - \Omega$, so that $\cos(\pi - \Omega) = -\sin\Omega$.

The second rotation takes place about $\hat{\boldsymbol{g}}'_1 = \hat{\boldsymbol{n}} = \hat{\boldsymbol{p}}'_1$ and takes $\hat{\boldsymbol{g}}'_3 = \hat{\boldsymbol{g}}_3$ onto $\hat{\boldsymbol{p}}'_3 = \hat{\boldsymbol{p}}_3 = \vec{\boldsymbol{h}}/h$. Its amplitude i is the orbit inclination, and it can be shown that the coordinate transformation is ruled by the equation

$$\boldsymbol{v}'_G = \boldsymbol{T}_{G'P'} \boldsymbol{v}_{P'} = \boldsymbol{T}'' \boldsymbol{v}_{P'}$$

where

$$\boldsymbol{T}_{G'G''} = \boldsymbol{T}'' = \boldsymbol{T}_1(i) = \begin{bmatrix} \hat{\boldsymbol{p}}'_1 \cdot \hat{\boldsymbol{g}}'_1 & \hat{\boldsymbol{p}}'_2 \cdot \hat{\boldsymbol{g}}'_1 & \hat{\boldsymbol{p}}'_3 \cdot \hat{\boldsymbol{g}}'_1 \\ \hat{\boldsymbol{p}}'_1 \cdot \hat{\boldsymbol{g}}'_2 & \hat{\boldsymbol{p}}'_2 \cdot \hat{\boldsymbol{g}}'_2 & \hat{\boldsymbol{p}}'_3 \cdot \hat{\boldsymbol{g}}'_2 \\ \hat{\boldsymbol{p}}'_1 \cdot \hat{\boldsymbol{g}}'_3 & \hat{\boldsymbol{p}}'_2 \cdot \hat{\boldsymbol{g}}'_3 & \hat{\boldsymbol{p}}'_3 \cdot \hat{\boldsymbol{g}}'_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos i & -\sin i \\ 0 & \sin i & \cos i \end{bmatrix}$$

The last rotation takes place about the new third axis $\hat{\boldsymbol{p}}'_3 = \hat{\boldsymbol{p}}_3 = \vec{\boldsymbol{h}}/h$ and takes $\hat{\boldsymbol{g}}'_1 = \hat{\boldsymbol{n}} = \hat{\boldsymbol{p}}'_1$ onto $\hat{\boldsymbol{p}}_1 = \vec{\boldsymbol{e}}/e$. Again the coordinate transformation can be expressed in matrix form,

$$\boldsymbol{v}_{P'} = \boldsymbol{T}_{P'P} \boldsymbol{v}_P = \boldsymbol{T}''' \boldsymbol{v}_P$$

and the coordinate transformation matrix for rotation about the third axis is similar to the previous one (but for the amplitude of the rotation), so that

$$\boldsymbol{T}_{P'P} = \boldsymbol{T}''' = \boldsymbol{T}_3(\omega) = \begin{bmatrix} \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{p}}'_1 & \hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{p}}'_1 & \hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{p}}'_1 \\ \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{p}}'_2 & \hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{p}}'_2 & \hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{p}}'_2 \\ \hat{\boldsymbol{p}}_1 \cdot \hat{\boldsymbol{p}}'_3 & \hat{\boldsymbol{p}}_2 \cdot \hat{\boldsymbol{p}}'_3 & \hat{\boldsymbol{p}}_3 \cdot \hat{\boldsymbol{p}}'_3 \end{bmatrix} = \begin{bmatrix} \cos\omega & -\sin\omega & 0 \\ \sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrix multiplication provide the overall transformation matrix:

$$\boldsymbol{v}_G = \boldsymbol{T}_{GG'}\boldsymbol{v}_{G'} = \boldsymbol{T}'\boldsymbol{v}_{G'} = \boldsymbol{T}'(\boldsymbol{T}''\boldsymbol{v}_{P'}) = \boldsymbol{T}'[\boldsymbol{T}''(\boldsymbol{T}'''\boldsymbol{v}_P)] = (\boldsymbol{T}'\boldsymbol{T}''\boldsymbol{T}''')\boldsymbol{v}_P$$

so that the coordinate transformation matrix takes the form

$$\boldsymbol{T}_{GP} = \begin{bmatrix} \cos\Omega\cos\omega - \sin\Omega\sin\omega\cos i & -\cos\Omega\sin\omega - \sin\Omega\cos\omega\cos i & \sin\Omega\sin i \\ \sin\Omega\cos\omega + \cos\Omega\sin\omega\cos i & -\sin\Omega\sin\omega + \cos\Omega\cos\omega\cos i & -\cos\Omega\sin i \\ \sin\omega\sin i & \cos\omega\sin i & \cos i \end{bmatrix}$$

Euler's eigenaxis rotation theorem states that it is possible to rotate any frame onto another one with a simple rotation around an axis $\hat{\boldsymbol{a}}$ that is fixed in both frames, called the *Euler's rotation axis* or *eigenaxis*, the direction cosines of which are the same in the two considered frame (a simple algebraic proof of Euler's theorem is given in the appendix). Three parameters (one for the rotation angle plus two for the axis orientation) define the rotation.

Equivalently three Euler elementary, non coplanar rotations are sufficient to take the first frame onto the second one. There are $3 \times 2 \times 2 = 12$ possible sequences of Euler rotations. The order in which they are performed is not irrelevant, as matrix multiplication is not commutative. Singularities, e.g., when the orbit is circular or equatorial, cannot be avoided unless a fourth parameter is introduced.

3.8 Derivative in a Rotating Reference Frame

Consider a base reference frame $\mathcal{F}_B$ (which is not necessarily fixed) and another, rotating one $\mathcal{F}_R$, defined by the unit vectors $\hat{\boldsymbol{\imath}}$, $\hat{\boldsymbol{\jmath}}$, and $\hat{\boldsymbol{k}}$, rotating with angular velocity $\vec{\boldsymbol{\omega}}$ with respect to the base frame.

The time derivative of a generic vector

$$\vec{\boldsymbol{f}} = f_1\hat{\boldsymbol{\imath}} + f_2\hat{\boldsymbol{\jmath}} + f_3\hat{\boldsymbol{k}}$$

with respect to the base coordinate system is

$$\left.\frac{\mathrm{d}\vec{\boldsymbol{f}}}{\mathrm{d}t}\right|_B = \dot{f}_1\hat{\boldsymbol{\imath}} + \dot{f}_2\hat{\boldsymbol{\jmath}} + \dot{f}_3\hat{\boldsymbol{k}} + f_1\dot{\hat{\boldsymbol{\imath}}} + f_2\dot{\hat{\boldsymbol{\jmath}}} + f_3\dot{\hat{\boldsymbol{k}}}$$

It is easily proved that

$$\dot{\hat{\boldsymbol{\imath}}} = \vec{\boldsymbol{\omega}} \times \hat{\boldsymbol{\imath}} \; ; \quad \dot{\hat{\boldsymbol{\jmath}}} = \vec{\boldsymbol{\omega}} \times \hat{\boldsymbol{\jmath}} \; ; \quad \dot{\hat{\boldsymbol{k}}} = \vec{\boldsymbol{\omega}} \times \hat{\boldsymbol{k}}$$

and therefore

$$\left.\frac{\mathrm{d}\vec{\boldsymbol{f}}}{\mathrm{d}t}\right|_B = \left.\frac{\mathrm{d}\vec{\boldsymbol{f}}}{\mathrm{d}t}\right|_R + \vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{f}}$$

where the subscript R denotes a time derivative as it appears to an observer moving with the rotating frame. This rule is independent of the physical meaning of vector $\vec{\boldsymbol{f}}$.

In particular, when one takes the first and second time derivative of the position vector $\vec{\boldsymbol{r}}$ in an inertial frame $\mathcal{F}_I$, it is

$$\vec{\boldsymbol{v}}|_I = \left.\frac{\mathrm{d}\vec{\boldsymbol{r}}}{\mathrm{d}t}\right|_I = \left.\frac{\mathrm{d}\vec{\boldsymbol{r}}}{\mathrm{d}t}\right|_R + \vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{r}} = \vec{\boldsymbol{v}}|_R + \vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{r}}$$

and

$$\vec{\boldsymbol{a}}|_I = \left.\frac{\mathrm{d}\vec{\boldsymbol{v}}|_I}{\mathrm{d}t}\right|_I = \left.\frac{\mathrm{d}\vec{\boldsymbol{v}}_I}{\mathrm{d}t}\right|_R + \vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{v}}|_I = \left.\frac{\mathrm{d}^2\vec{\boldsymbol{r}}}{\mathrm{d}t^2}\right|_R + 2\vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{v}}|_R + \vec{\boldsymbol{\omega}} \times (\vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{r}})$$

The second equation of dynamics $m\vec{\boldsymbol{a}}|_I = \vec{\boldsymbol{f}}$, when written in a rotating frame becomes

$$m \left.\frac{\mathrm{d}^2\vec{\boldsymbol{r}}}{\mathrm{d}t^2}\right|_R = \vec{\boldsymbol{f}} + m2\vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{v}}|_R + m\vec{\boldsymbol{\omega}} \times (\vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{r}})$$

where, together with the resultant of the applied forces $\vec{\boldsymbol{f}}$, two apparent forces

$$\vec{\boldsymbol{f}}_{co} = -2m\vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{v}}|_R \quad \text{and} \quad \vec{\boldsymbol{f}}_{cf} = -m\vec{\boldsymbol{\omega}} \times (\vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{r}})$$

(named Coriolis and centrifugal force, respectively) must be added.

3.9 Topocentric Reference Frame

The launch of a satellite or a radar observation is made from a point on the Earth surface. The propulsive effort or the measured signal is connected to a rotating reference frame centered on the launch pad or radar location. The obvious choice for the fundamental plane is the local horizontal plane and as a consequence the z–axis is parallel to the local vertical , and points to the zenith. The x–axis points southward along the local meridian, and the y–axis eastward along the parallel. The right-handed set of unit vectors $\hat{\boldsymbol{t}}_1$, $\hat{\boldsymbol{t}}_2$, and $\hat{\boldsymbol{t}}_3$ defines the *topocentric frame* $\mathcal{F}_T$, centered on a location (*topos* in ancient greek) on the Earth's surface.

Representation of vectors in Topocentric coordinates

The vectors $\vec{\boldsymbol{\rho}}$ and $\vec{\boldsymbol{w}}$ are used to express position and velocity relative to the topocentric frame, that is, as they appear to an observer fixed to this frame. Spherical polar coordinates, expressed as magnitude of the vector ρ and two angles (elevation φ above the horizon and azimuth ψ measured clockwise from the North) are often preferred to the Cartesian components.

By indicating with the symbols $\boldsymbol{\rho}_T = (\rho_1, \rho_2, \rho_3)^T$ and $\boldsymbol{w}_T = (w_1, w_2, w_3)^T$ the components of the (relative) position and velocity vectors $\vec{\boldsymbol{\rho}}$ and $\vec{\boldsymbol{w}}$ in $\mathcal{F}_T$, respectively, it is

$$\begin{array}{rclcrcl} \rho_1 &=& -\rho\cos\varphi_\rho\cos\psi_\rho & ; & w_1 &=& -w\cos\varphi_w\cos\psi_w \\ \rho_2 &=& \rho\cos\varphi_\rho\sin\psi_\rho & ; & w_2 &=& w\cos\varphi_w\sin\psi_w \\ \rho_3 &=& \rho\sin\varphi_\rho & ; & w_3 &=& w\sin\varphi_w \end{array}$$

Nevertheless, the energy achieved by the spacecraft with the launch is directly related to the distance from the central body,

$$\vec{\boldsymbol{r}} = \vec{\boldsymbol{r}}_\oplus + \vec{\boldsymbol{\rho}} = \rho_1\hat{\boldsymbol{t}}_1 + \rho_2\hat{\boldsymbol{t}}_2 + (r_\oplus + \rho_3)\hat{\boldsymbol{t}}_1$$

and the absolute velocity (*i.e.*, the velocity with respect to a non-rotating reference frame such as the set of geocentric–equatorial coordinates)

$$\vec{\boldsymbol{v}} = \mathrm{d}\vec{\boldsymbol{r}}/\mathrm{d}t|_G = \vec{\boldsymbol{w}} + \vec{\boldsymbol{\omega}}_\oplus \times \vec{\boldsymbol{r}} = \vec{\boldsymbol{w}} + \vec{\boldsymbol{u}}$$

where $\vec{\boldsymbol{u}} = \omega_\oplus r\cos\delta\hat{\boldsymbol{t}}_2 \approx \omega_\oplus r_\oplus\cos\delta\hat{\boldsymbol{t}}_2$ decreases from the maximum value $u_{eq} = 464.6m/s$ at the equator, to zero at either pole. The components of the absolute velocity in the topocentric frame are therefore

$$\begin{array}{rclcrcl} v_1 &=& -w\cos\varphi_w\cos\psi_w & ; & v &=& \sqrt{v_1^2 + v_2^2 + v_3^2} \\ v_2 &=& w\cos\varphi_w\sin\psi_w + u & ; & \sin\phi_w &=& v_3/v \\ v_3 &=& w\sin\varphi_w & ; & tan\phi_w &=& -v_2/v_1 \end{array}$$

One should note that azimuth and elevation angles of the absolute velocity (ψ_v and ϕ_v) are different from the same angles (ψ_w and ϕ_w) for the relative velocity, that is, the angles as seen by an observer looking at the moving spacecraft from the topocentric frame.

As far as vector magnitudes are concerned, it is $w - u \leq v \leq w + u$, and the launch strategy aims at obtaining v (the effective velocity for orbital motion) larger than w (the rocket velocity attainable by means of the propellant only). In order to obtain the maximum benefit from the Earth rotation, u should be as high as possible and $v = w + u$. The last requirement implies $w_1 = w_3 = 0$, that is, $\phi_w = 0$ and $\psi_w = \pi/2$.

Coordinate transformation matrix from topocentric to geocentric equatorial frame

The estimation of the orbit parameters of a satellite requires the simultaneous knowledge of position and velocity vectors in the non–rotating, geocentric equatorial frame at a given time t_0. If these vectors are evaluated from a tracking station GS on the ground, with geographical coordinates Lat and Lo, they will be known in a topocentric frame centered in GS, and a coordinate transformation from topocentric to geocentric equatorial frame is thus necessary.

It must be remembered that the topocentric frame rotates together with the Earth, so that, first of all, it is necessary to evaluate the right ascension of the ground station itself, that is, the angular displacement between the vernal equinox unit vector $\hat{\boldsymbol{g}}_1$, and the plane of the meridian passing through the ground station. This can be evaluated as the sum of the right ascension of the Greenwich meridian GM plus the longitude of GS,

$$\alpha = \alpha_G + \mathrm{Lo}$$

where

$$\alpha_G = \alpha_{G_0} + \omega_\oplus (t - t_0)$$

The angle α_{G_0} at t_0 is often provided as the *Greenwich sidereal time* at 0:00 of January, 1st of a specified year. In such a case, α_{G_0} is close to 100 deg, with less than 1 deg oscillations due to the non–integer number of the days in a year.

As the equatorial plane is the base plane of the geocentric equatorial frame, the declination of GS is equal to its latitude, that is δ =Lat.

The elementary rotations that take $\mathcal{F}_G$ onto $\mathcal{F}_T$ are a rotation *alpha* about $\hat{\boldsymbol{g}}_3$, followed by a rotation $\pi/2 - \delta$ about $\hat{\boldsymbol{t}}_2$, that is

$$\boldsymbol{T}_{GT} = \boldsymbol{T}_3(\alpha)\boldsymbol{T}_2(\pi/2 - \delta)$$

where the symbol $\boldsymbol{T}_i(\ )$ indicates, as in the previous paragraph, the coordinate transformation matrix for an elementary rotation about the i-th axis. Being

$$\boldsymbol{T}_3(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$\boldsymbol{T}_2(\pi/2 - \delta) = \begin{bmatrix} \cos(\pi/2 - \delta) & 0 & \sin(\pi/2 - \delta) \\ 0 & 1 & 0 \\ -\sin(\pi/2 - \delta) & 0 & \cos(\pi/2 - \delta) \end{bmatrix} = \begin{bmatrix} \sin\delta & 0 & \cos\delta \\ 0 & 1 & 0 \\ -\cos\delta & 0 & \sin\delta \end{bmatrix}$$

one gets

$$\boldsymbol{T}_{GT} = \begin{bmatrix} \cos\alpha\sin\delta & -\sin\alpha & \cos\alpha\cos\delta \\ \sin\alpha\sin\delta & \cos\alpha & \sin\alpha\cos\delta \\ -\cos\delta & 0 & \sin\delta \end{bmatrix}$$

3.10 Satellite Ground Track

Information on overflown regions and satellite visibility from an observer on the ground requires the knowledge of the motion of the satellite relative to the Earth's surface. This motion results from the composition of the Keplerian motion of the satellite with the Earth rotation on its axis. The track of a satellite on the surface of a spherical Earth is the loci of the intersections of the radius vector with the surface. The altitude and the track on a chart of the Earth constitute a quite useful description of the satellite motion.

The ground track of a satellite moving on a Keplerian orbit is a great circle, if the Earth is assumed spherical and non-rotating. Suppose that the satellite overflies point S of declination δ at time t. Consider the meridians passing through S and the ascending node N. Complete a spherical triangle NMP with an equatorial arc of angular length $\Delta\alpha$ between the meridians, the third vertex being the pole P. The ground track will split this triangle into two smaller ones (NMS and NSP). By applying the law of sines to both triangles, one obtains

$$\begin{aligned} \frac{\sin i}{\sin\delta} &= \frac{\sin(\pi/2)}{\sin\theta} & \Rightarrow \qquad \sin\delta &= \sin i \sin\theta \\ \frac{\sin\theta}{\sin\Delta\alpha} &= \frac{\sin(\pi/2-\delta)}{\sin(\pi/2-i)} = \frac{\sin(\pi/2)}{\sin(\pi-\psi)} & \Rightarrow \qquad \sin\Delta\alpha &= \frac{\cos i}{\cos\delta}\sin\theta \end{aligned}$$

The equations on the right side provide the satellite track in a non–rotating frame. The angular position of N with respect to the equinox line (*i.e.*, $\hat{\boldsymbol{g}}_1$) is called local sidereal time $\alpha = \Omega + \Delta\alpha$. The law of sines for the triangle NMP also provides the equation

$$\cos i = \cos\delta \sin\psi$$

an important relation that will be discussed in the following chapter.

Geographical latitude La and longitude Lo are needed to locate point S in a cartographical representation of the Earth surface. The former is simply Lat $= \delta$, while the latter is given by Lo $= \alpha - \alpha_G$ and requires the knowledge of the Greenwich sidereal time, which is given by (see the previous paragraph)

$$\alpha_G = \alpha_{G_0} + \omega_\oplus (t - t_0)$$

3.11 Ground Visibility

For the sake of simplicity, consider a satellite moving a circular orbit of altitude z above the ground. Adequate visibility from a base L on the Earth surface requires a minimum elevation ε above the horizon, dependent on radiofrequency and characteristics of the radio station. The satellite is seen from the base when its position on the ground track is inside a circle, which is drawn around L on the sphere, and whose radius has angular width Σ. The same angle describes the Earth surface around the satellite and visible from it.

The law of sines of plane trigonometry applied to the triangle OLS (S is the limit position for visibility),

$$\frac{\sin(\pi - \Sigma - \pi/2 - \varepsilon)}{r_\oplus} = \frac{\sin(\pi/2+\varepsilon)}{r_\oplus + z} = \frac{\sin\Sigma}{d}$$

provides the angle Σ:

$$\cos(\Sigma + \varepsilon) = \frac{r_\oplus}{r_\oplus + z}\cos\varepsilon$$

and the maximum distance between a visible satellite and the ground–base base L:

$$d = (r_\oplus + z)\frac{\sin\Sigma}{\cos\varepsilon}$$

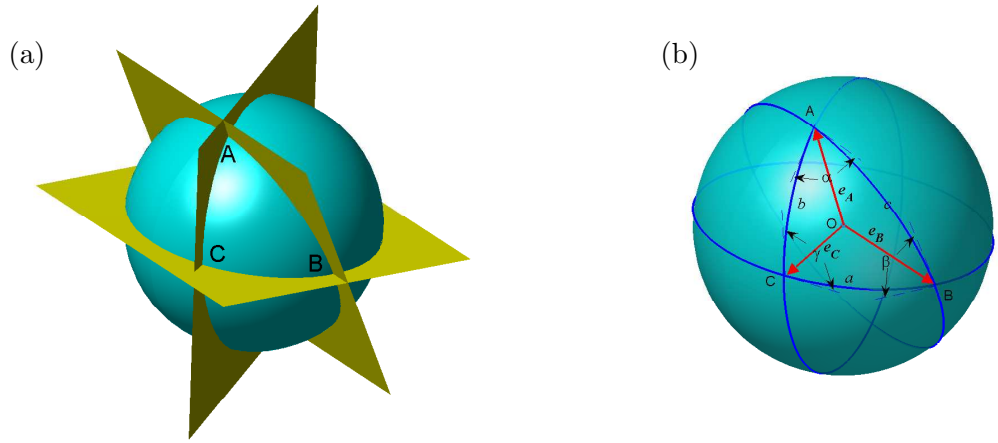

Figure 3.3: **Great circles (a) and spherical triangle ABC (b).**

The latter parameter is related to the power required for data transmission.

The time of visibility depends on

- the length of the ground track inside the circle (the closer to L the passage, the longer the track);
- the satellite altitude (the higher z the lower the spacecraft angular velocity i.e., the time to cover a unit arc of the ground track).

Note that the visibility zone is a circle on the sphere. On a cartographic representation of the Earth surface, the same zone is again a circle for an equatorial base, but is increasingly deformed for bases located at increasing latitudes.

Appendix A - Spherical Trigonometry

Consider a sphere of unit radius. The curve of intersection of the sphere and a plane passing through the center is a great circle. A spherical triangle is a figure formed on the surface of a sphere by three great circular arcs intersecting pairwise in three vertices.

Eight triangles are actually created, and we will consider only triangle ABC (Fig. 3.11(a)). The sum of the angles of a spherical triangle $\alpha + \beta + \gamma$ is greater than π, being something between π and 3π.

The study of angles and distances on a sphere is known as Spherical Trigonometry and is essentially based on two basic equations, (i) the law of cosines for the sides (a similar one for the angles exists), and (ii) the law of sines.

Let a spherical triangle be drawn on the surface of a unit sphere centered at point O, with vertices A, B, and C (Fig. 3.11(b)). The unit vectors from the center of the sphere to the vertices are $\hat{\boldsymbol{e}}_A$, $\hat{\boldsymbol{e}}_B$, and $\hat{\boldsymbol{e}}_C$, respectively. The length of the side BC, denoted by a, is seen as an angle from the center of the sphere (remember that a great circle on the unit sphere is a circle of unit radius, such that the length on an arc on the circle is equal to the angle at the center expressed in radians). The angle on the spherical surface at vertex A is the dihedral angle between planes AOB and AOC, and is denoted by α.

A.1 - Law of the cosines

The dot product $(\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_B) \cdot (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C)$ can be expressed in two ways:

$$(\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_B) \cdot (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C) \quad = \quad [(\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_B) \times \hat{\boldsymbol{e}}_A] \cdot \hat{\boldsymbol{e}}_C$$

$$\begin{aligned} &= [(\hat{\boldsymbol{e}}_A \cdot \hat{\boldsymbol{e}}_A)\hat{\boldsymbol{e}}_B - (\hat{\boldsymbol{e}}_A \cdot \hat{\boldsymbol{e}}_B)\hat{\boldsymbol{e}}_A] \cdot \hat{\boldsymbol{e}}_C \\ &= [\hat{\boldsymbol{e}}_B - \cos c\hat{\boldsymbol{e}}_A] \cdot \hat{\boldsymbol{e}}_C \\ &= \cos a - \cos b \cos c \end{aligned}$$

But it is also

$$(\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_B) \cdot (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C) = \sin b \sin c \cos \alpha$$

and equating the results, one obtains the law of cosines for the sides:

$$\cos a = \cos b \cos c + \sin b \sin c \cos \alpha$$

A.2 - Law of the sines

After computing

$$\begin{aligned} \frac{\sin \alpha}{\sin a} &= \frac{1}{\sin a} \frac{||(\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_B) \times (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C)||}{||\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_B||\ ||\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C||} \\ &= \frac{1}{\sin a} \frac{||[\hat{\boldsymbol{e}}_A \cdot (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C)]\hat{\boldsymbol{e}}_B - [\hat{\boldsymbol{e}}_B \cdot (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C)]\hat{\boldsymbol{e}}_A||}{\sin b \sin c} \\ &= \frac{||\hat{\boldsymbol{e}}_B \cdot (\hat{\boldsymbol{e}}_A \times \hat{\boldsymbol{e}}_C)\hat{\boldsymbol{e}}_A||}{\sin a \sin b \sin c} \end{aligned}$$

one recognizes that the result is independent of the selected vertex on the basis of the geometrical meaning of the numerator (see Eq. A.4 in Chapter 1). The law of sines is therefore expressed as

$$\frac{\sin a}{\sin \alpha} = \frac{\sin b}{\sin \beta} = \frac{\sin c}{\sin \gamma}$$

Appendix B - Algebraic proof of Euler's theorem

A simple algebraic demonstration of Euler's theorem can be obtained from the following considerations:

- the eigenvalues of any (real) orthogonal matrix $\boldsymbol{T}$ have unit modulus; indicating with H the Hermitian conjugate, which, for a real matrix is coincident with the transpose, one has

  $$\boldsymbol{T}\boldsymbol{a} = \lambda \boldsymbol{a} \Rightarrow \boldsymbol{a}^H \boldsymbol{T}^T \boldsymbol{T}\boldsymbol{a} = \bar{\lambda}\lambda \boldsymbol{a}^H \boldsymbol{a} \Rightarrow (1 - \bar{\lambda}\lambda)\boldsymbol{a}^H \boldsymbol{a} = 0$$

  that for any nontrivial eigenvector $\boldsymbol{a}$ implies that

  $$\bar{\lambda}\lambda = 1 \Rightarrow |\lambda| = 1$$

- (at least) one eigenvalue is $\lambda = 1$; any $n \times n$ real matrix has at least one real eigenvalue if n is an odd number, which means that a 3×3 orthogonal matrix must have at least one eigenvalue which is $\lambda_1 = \pm 1$. The other couple of eigenvalues will be, in the most general case, complex conjugate numbers of unit modulus, which can be cast in the form $\lambda_{2,3} = \exp(\pm i\phi)$. The determinant is equal to the product of the eigenvalues, which is one, for an orthogonal matrix, so that

  $$\lambda_1 \lambda_2 \lambda_3 = 1 \Rightarrow \lambda_1 = 1$$

The eigenvector relative to the first eigenvalue satisfies the relation

$$\boldsymbol{T}\boldsymbol{a} = 1 \cdot \boldsymbol{a}$$

This means that there is a direction $\boldsymbol{a}$ which is not changed under the action of transformation matrix $\boldsymbol{T}$. If $\boldsymbol{T}$ represents a coordinate change, the vector $\boldsymbol{a}$ will be represented by the same components in both the considered reference frames. For this reason, the transformation that takes the initial frame onto the final one can be considered as a single rotation α about the Euler axis $\hat{\boldsymbol{a}}$.

Chapter 4

Earth Satellites

4.1 Introduction

Several aspects of satellite launch and operation are analyzed in this chapter. The orbit cannot be reached directly from the ground and part of the propulsive effort is done at high altitude with a reduced efficiency. The orbit is perturbed by non-Keplerian forces, mainly aerodynamic drag and Earth asphericity, in the proximity of the planet, solar and lunar gravitation for satellites orbiting at higher altitudes. The choice of the orbital parameters is not free as a satellite operates in strict connection with the ground, continuously up–loading and down–loading data from and to ground–stations at each passage over them.

The mission starts with the ascent to the orbit, that is greatly dependent on the geographical location of the launch base. Orbital maneuvers may be required for the achievement of the final orbit. Maneuvering is also necessary to correct injection errors and maintain the operational orbit compensating for the effects of perturbations.

4.2 The cost of thrusting in space

The simplest way of changing the specific energy of a satellite of mass m is the use of the engine thrust T provided by exhausting an adequate mass flow $(-\dot{m})$ of propellant. A thrust impulse $T\mathrm{d}t$ is applied to a spacecraft in the absence of gravitational and aerodynamic forces. This hypothesis is not too restrictive: for instance, it is compatible with the presence of gravitation, if the displacement $\mathrm{d}r$ in the time interval $\mathrm{d}t$ is negligible,

$$\mathrm{d}\vec{\boldsymbol{v}} = \frac{\vec{\boldsymbol{T}}}{m}\mathrm{d}t \ ; \quad c = -\frac{T}{\dot{m}} \quad \Rightarrow \quad \mathrm{d}v = -c\frac{\mathrm{d}m}{m}$$

where c is the engine effective exhaust velocity. The mass expenditure $\mathrm{d}m$ (< 0) represents the cost, which is directly related to the infinitesimal velocity increment.

The burn time of a chemical engine is short when compared with the orbital period. The impulsive model (infinite thrust applied for infinitesimal time, therefore at constant radius) is undoubtedly acceptable for an interplanetary mission, where orbit periods are very long and the approximation of neglecting the thrust pulse duration is reasonable. For Earth orbiting satellites, the model still holds without excessive errors; in any case it provides the theoretical minimum propellant consumption, as the whole propulsive effort is concentrated in the optimal position for delivering the thrust pulse.

A different model (finite-thrust maneuver), which takes into account the finite magnitude of the available thrust, provides more accurate results. The use of the exact model, where pulses of finite duration are considered, is necessary when electric propulsion is exploited.

The thrust level determines the angular length of the maneuvers and is quantified by comparing the thrust and gravitational accelerations. The thrust acceleration available in present spacecraft permits low–thrust few–revolution interplanetary trajectories and very–low–thrust multi–revolution maneuvers orbit transfers around the Earth.

Assume that an impulsive maneuver is carried out to provide an assigned energy increment $\Delta\mathcal{E}$ to a spacecraft moving on a Keplerian trajectory. The thrust direction and the velocity increment $\Delta\vec{\boldsymbol{v}}$ form an angle β with the velocity $\vec{\boldsymbol{v}}_1$ just before the maneuver. The velocity after the (ideally instantaneous) thrust pulse will be

$$\vec{\boldsymbol{v}}_2 = \vec{\boldsymbol{v}}_1 + \Delta\vec{\boldsymbol{v}} \quad \Rightarrow \quad v_2^2 = v_1^2 + \Delta v^2 + 2v_1\Delta v\cos\beta$$

so that the assigned energy increment is

$$\Delta\mathcal{E} = \mathcal{E}_2 - \mathcal{E}_1 = -\frac{\mu}{2a_2} + \frac{\mu}{2a_1} = \frac{v_2^2}{2} - \frac{v_1^2}{2} = \frac{1}{2}\Delta v(\Delta v + 2v_1\cos\beta)$$

suggesting that the Δv magnitude, i.e., the cost, is minimized if

- the thrust is applied in the point of the largest velocity on the initial trajectory;
- the thrust vector is parallel to the velocity ($\beta = 0$ for $\Delta\mathcal{E} > 0$ or $\beta = \pi$ for $\Delta\mathcal{E} < 0$).

The same result can be obtained by considering the mechanical energy

$$\mathcal{E} = \frac{v^2}{2} - \frac{\mu}{r}$$

and evaluating its increment for a constant position, $r = \text{const}$, so that

$$\mathrm{d}\mathcal{E} = v\mathrm{d}v = \vec{\boldsymbol{v}}\cdot\mathrm{d}\vec{\boldsymbol{v}}$$

As a consequence, in order to obtain a given energy increment with the minimum velocity increment (*i.e.* minimum propellant consumption), one should avoid

- misalignment losses, wasting part of the thrust to rotate the spacecraft velocity);
- gravitational losses, letting the gravity reduce the velocity while the spacecraft moves away from the central body (*i.e.* fire the thrusters when r is small and v is higher).

4.3 Energetic aspects of the orbit injection

If a spacecraft flying on a circular orbit of radius r_c is considered, it is

$$v_c = \sqrt{\frac{\mu}{r_c}}\ ; \quad \mathcal{T} = \frac{2\pi r_c}{v_c} = 2\pi\sqrt{\frac{r_c^3}{\mu}}$$

As the radius increases from the minimum theoretical value $r_\oplus = 6371$ km (mean radius of the Earth), the velocity continuously decreases from the maximum value $v_I = 7.91$ km s^{-1}. The period increases starting from $84^{\mathrm{m}}\ 24^{\mathrm{s}}$, as the spacecraft becomes slower on a longer orbit, farther from the Earth centre. Figure 4.1 presents the kinetic ($\mathcal{E}_c$), gravitational ($\mathcal{E}_g$), and total ($\mathcal{E}$) nondimensional energy, scaled with respect to the absolute value of the gravitational energy on the Earth surface, as a function of the nondimensional radius $r/r_\oplus$. For higher orbits the kinetic energy diminishes, but its reduction is only half of the increment of the potential energy, that prevails, so that a greater propulsive effort $\Delta\mathcal{E}$ is required to attain higher orbit.

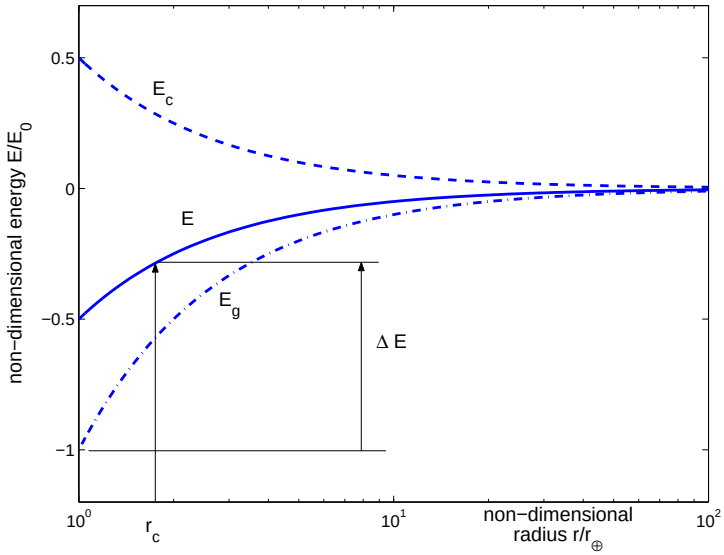

Figure 4.1: **Energy of circular orbits.**

The same Fig. 4.1 also applies to an elliptic orbit, if the semi–major axis a is considered instead of the radius r_c. An elliptic orbit is equivalent to the circular one with $r_c = a$, from the energetic point of view, but has a lower pericenter. One should note that kinetic and potential energies are not constant along the elliptic orbit; the values of Fig. 4.1 are attained when the satellite passes through the minor axis. When compared with non-equivalent ellipses, the circular orbit has the minimum energy among orbits with equal pericenter, and the maximum energy if the same apocenter is assumed.

If the ascent to an orbit of semi-major axis a_2 from the Earth surface is considered, the equivalent velocity on the ground $v_{0_{eq}}$, that corresponds to the increment of spacecraft energy sufficient to reach the energy relative to the target orbit, is obtained from the equation

$$\frac{v_{0_{eq}}^2}{2} = \Delta\mathcal{E} = \mathcal{E}_2 - \mathcal{E}_0 = -\frac{\mu}{2a_2} + \frac{\mu}{r_\oplus}$$

and lies between the first and second cosmic velocities. Unfortunately one impulsive velocity increment on the Earth surface does not permit to attain an orbit with $r_P > r_\oplus$: the spacecraft will pass again through the point where the engine has been turned off at the end of the injection maneuver (see the dotted line in Fig. 4.2). This means that in order to attain orbit with a perigee radius greater than the Earth radius, the last application of thrust must occur at $r_P > r_\oplus$.

At least two impulsive burns are therefore necessary. By means of a first velocity impulse v_0, the spacecraft leaves the ground (point 0) on a ballistic trajectory with an apogee at point 1. A second tangential burn $\Delta v = v_2 - v_1$ in 1 parallel to the velocity vector (thus avoiding misalignment losses) will raise the perigee of the orbit up to r_2. The total energy increase is obtained by means of two increments of kinetic energy:

$$\Delta\mathcal{E} = \frac{v_0^2}{2} + \frac{v_2^2}{2} - \frac{v_1^2}{2} = \frac{v_0^2}{2} + \frac{\Delta v^2}{2} + v_1\Delta v$$

The comparison between the square of the total velocity increment ($\Delta v_{tot} = v_0 + \Delta v$) and the equivalent velocity increment on the ground ($v_{0_{eq}}$) provides

$$\frac{\Delta v_{tot}^2}{2} - \frac{v_{0_{eq}}^2}{2} = \frac{(v_0 + \Delta v)^2}{2} - \Delta\mathcal{E} = (v_0 - v_1)\Delta v$$

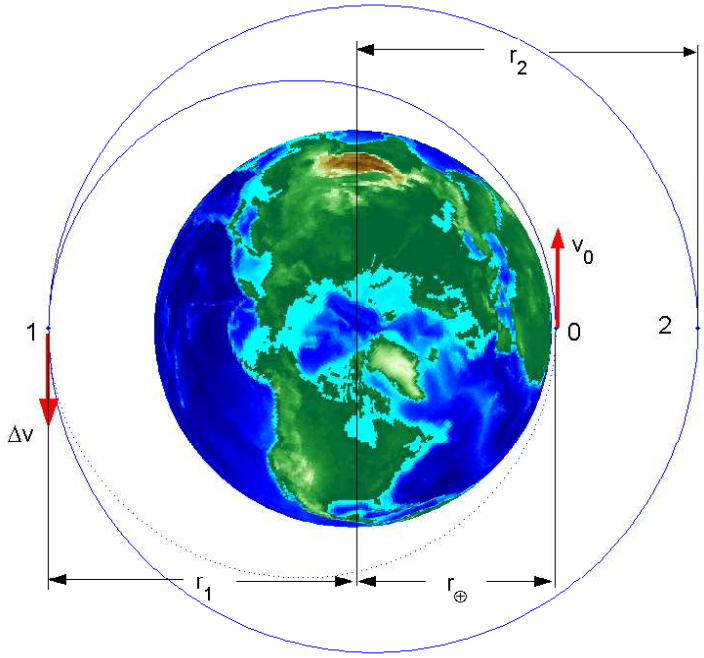

Figure 4.2: **Orbit injection with two velocity increments.**

Since the mechanical energy is constant during the ballistic ascent, for an assigned value of the radius r_1 of the injection point 1 it is

$$\frac{v_0^2}{2} - \frac{v_1^2}{2} = \frac{\mu}{r_0} - \frac{\mu}{r_1} = \text{const} = \frac{C}{2}$$

and therefore v_0 results to be an increasing function of v_1, since

$$v_0 - v_1 = \frac{C}{v_0 + v_1}$$

As a consequence it is possible to write the energy loss with respect to the equivalent Δv on the ground as

$$\frac{\Delta v_{tot}^2}{2} - \frac{v_{0_{eq}}^2}{2} = C\frac{v_2 - v_1}{v_0 + v_1} = C\frac{\Delta v}{v_0 + v_1}$$

where the loss, which is actually gravitational, is reduced if the ballistic trajectory reaches the injection point with the highest value of the velocity v_1, that is, the transfer ballistic orbit from 0 to 1 has the largest value of the semi-major axis possible to reach point 1, which in turn is the perigee of the final orbit. In order to minimize Δv losses during the launch phase, the spacecraft should depart horizontally from a ground base located at the antipodes of the injection point, as depicted in Fig. 4.2.

4.4 Injection errors

Consider an ascent to a circular orbit, carried out according to the model presented in the previous section. Suppose that the injection point has been exactly attained; an error δv_2 on the velocity magnitude or $\delta\varphi_2$ on the flight path angle after the injection burn are considered separately, and their effect on the perigee of the final orbit is evaluated.

By differentiating both sides of the energy equation considering r as fixed one gets

$$\frac{v^2}{2} - \frac{\mu}{r} = -\frac{\mu}{2a} \;\Rightarrow\; v\mathrm{d}v = \frac{\mu}{2a^2}\mathrm{d}a \;\Rightarrow\; \mathrm{d}a = \frac{2a^2}{\mu}v\mathrm{d}v$$

which corresponds to an increment $\mathrm{d}z = 2\mathrm{d}a$ of the orbit altitude at the antipodes of the injection point. Therefore

$$\delta z = 2\delta a_2 = \frac{4a^2}{\mu}v_2\delta v_2 \approx 4a\frac{\delta v_2}{v_2}$$

where, for a circular orbit, it is $v_2 = \sqrt{\mu/r_2} = \sqrt{\mu/a_2}$.

If the Earth mass parameter is considered, a reduction of only 0.1% on v_2 for a circular orbit at 200 km altitude (that is, 7.8 m s^{-1}) produces a 0.4% reduction of the perigee height (26.3 km), whereas exceeding v_2 of the same amount would increase the apogee altitude of the same amount.

An error in any direction on the flight path angle lowers the perigee. From the equation for the orbit parameter it is possible to express the eccentricity,

$$\frac{h^2}{\mu} = p = a(1-e^2) \quad\Rightarrow\quad e = \sqrt{1 - \frac{h^2}{\mu a}}$$

where the angular momentum vector in 1, after application of the Δv, is given by $h = r_1 v_2 \cos\varphi$. As a consequence, for small errors on the flight–path angle (and consequently quasi–circular orbits), the eccentricity becomes

$$e = \sqrt{1 - \frac{r_1^2 v_2^2 \cos^2\varphi}{\mu a_2}} \approx \sqrt{1-\cos^2\varphi} = |\sin\varphi| \approx |\delta\varphi_2|$$

Thus, an error of 1 mrad (3' 26") in the direction of the velocity vector at injection reduces the perigee altitude by 6.6 km.

4.5 Perturbations

Keplerian trajectories are conic sections that describe the relative motion of two spherically symmetric bodies under the action of their mutual gravitational attraction only. A real spacecraft is acted upon by several forces, together with the gravitational pull from the closest planet which, by the way, is never perfectly spherical. These forces, however, are usually relatively small, if compared to the main gravitational action, and the orbits of planets and satellites are well approximated by conic sections. Nonetheless, in the long run, the combined action of the other forces will result in a sizable variation of the current position from the ideal one, predicted by the Keplerian theory, so that the actual trajectories are referred to as perturbed Keplerian orbits.

The six orbital parameters are uniquely determined by position and velocity of the spacecraft at some instant. They are not constant for a perturbed trajectory, but the spacecraft appears continuously passing from a Keplerian orbit to another. The keplerian orbit that at specific time has the same orbital elements of the perturbed trajectory is called *osculating orbit*. The *osculating orbital elements* of a spacecraft are the parameters of its osculating conic section, which varies with time. In general they present a linear trend (*secular variation*) with superimposed long–period oscillations, due to interactions with variations of another element, and shorter-period oscillations usually related to phenomena repeated during each orbital revolution.

The equation of motion can be written in the form

$$\ddot{\vec{r}} = -\frac{\mu}{r^3}\vec{r} + \vec{a}_P$$

where $\vec{a}_P$ is the acceleration due to the action of the perturbing forces. In general, no closed–form analytical solution is available for describing the long term effects of orbital perturbation, and numerical techniques (*special perturbation methods*) or approximate analytical solutions (*general perturbation methods*) are necessary to estimate such effects.

4.5.1 Special perturbations

Special perturbation methods numerically integrate the equation of motion and provide the trajectory of a particular spacecraft, starting from its initial conditions at a specific time (see Bate, cap. 9.5, p. 412-9).

The simple numerical integration of the perturbed equation of motion is called Cowell's method. It makes no use of the fact that the actual trajectory can be approximated by a conic section. The integration step is small, even though the Cartesian coordinates of the original method are replaced by other (e.g., spherical) coordinates.

Encke's method assume the osculating orbit at some instant as a reference Keplerian orbit which is known by means of the analytical solution of the unperturbed equation of motion

$$\ddot{\vec{r}}_K = -\frac{\mu}{r_K^3}\vec{r_K}$$

The deviation of the actual trajectory from the reference orbit

$$\delta\vec{r} = \vec{r} - \vec{r}_K \quad \Rightarrow \quad \delta\dot{\vec{r}} = \dot{\vec{r}} - \dot{\vec{r}}_K \quad \Rightarrow \quad \delta\ddot{\vec{r}} = \ddot{\vec{r}} - \ddot{\vec{r}}_K = -\mu\left(\frac{\vec{r}}{r^3} - \frac{\vec{r_K}}{r_K^3}\right) + \vec{a}_P$$

is numerically integrated. The integration step can be chosen larger than in Cowell's method, at least until perturbations accumulate and Δr becomes large. Then, the reference orbit is corrected, i.e., a new reference orbit is assumed (Cornelisse 18.3, p. 412; Bate 9.3, p. 390-96).

The method of variation of orbital parameters rewrites the vector equation of perturbed motion as a system of differential equation for the orbital parameters of the osculating orbit. Only the rate of change of the osculating parameters, which derives from the perturbing forces, is integrated. Moreover, the effects of perturbations are made evident by the clear geometrical significance of the orbital elements (Cornelisse 18.4, p. 413; Bate 9.4, p. 396-410).

4.5.2 General perturbations

General perturbations cover the analytical methods in which a perturbing acceleration is selected, expanded into series, which are opportunely truncated and integrated termwise. Analytical expressions describe the general effects of a particular perturbing force on the orbital elements as a function of time (Cornelisse 18.5, p. 425; Bate 9.4.3, p. 910-2).

4.6 Perturbed satellite orbits

The most important perturbing forces acting on an artificial Earth satellite are due to lunar and solar attraction, asphericity of the Earth, aerodynamic drag, solar radiation, and electromagnetic effects.

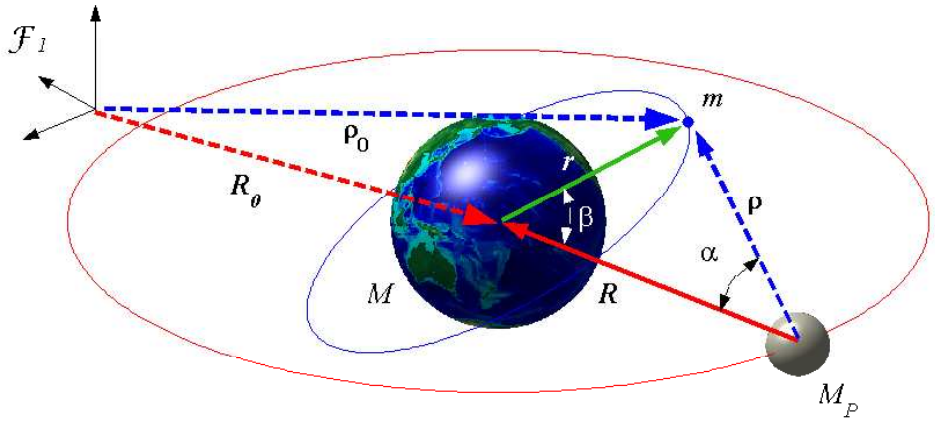

Figure 4.3: **A satellite in the Earth–Moon system.**

4.6.1 Lunar and solar attraction

Consider the three-body system, where a third body of mass M_P perturbs the motion of an Earth satellite m. Assume an inertial reference frame $\mathcal{F}_I$ where $\vec{\boldsymbol{R}}_0$ and $\vec{\boldsymbol{\rho}}_0$ are the position vectors of the Earth and of the satellite, respectively, while $\vec{\boldsymbol{R}}$ and $\vec{\boldsymbol{\rho}}$ define their position relative to the perturbing body. After writing the equations of motion for the spacecraft and the Earth,

$$\begin{aligned} m\ddot{\vec{\boldsymbol{\rho}}}_0 &= -G\frac{Mm}{r^3}\vec{\boldsymbol{r}} - G\frac{M_P m}{\rho^3}\vec{\boldsymbol{\rho}} \\ M\ddot{\vec{\boldsymbol{R}}}_0 &= G\frac{Mm}{r^3}\vec{\boldsymbol{r}} - G\frac{M_P M}{R^3}\vec{\boldsymbol{R}} \end{aligned}$$

where $\vec{\boldsymbol{r}} = \vec{\boldsymbol{\rho}}_0 - \vec{\boldsymbol{R}}_0$, it is possible to operate as in Section 1.5. Letting $\mu_P = GM_P$, one gets

$$\begin{aligned} \ddot{\vec{\boldsymbol{r}}} &= -\frac{\mu}{r^3}\vec{\boldsymbol{r}} - \frac{\mu_P}{\rho^3}\vec{\boldsymbol{\rho}} + \frac{\mu_P}{R^3}\vec{\boldsymbol{R}} \\ &= -\frac{\mu}{r^3}\vec{\boldsymbol{r}} + \vec{\boldsymbol{a}}_P \end{aligned}$$

where the disturbing acceleration,

$$\vec{\boldsymbol{a}}_P = \mu_P\left[\frac{\mu_P}{R^3}\vec{\boldsymbol{R}} - \frac{\mu_P}{\rho^3}\vec{\boldsymbol{\rho}}\right]$$

can be expressed as (see Fig. 4.3) as

$$a_P = \mu_P\sqrt{\frac{1}{R^4} + \frac{1}{\rho^4} - \frac{2\cos\alpha}{R^2\rho^2}}$$

which for $\rho \ll R$ reduces to $a_P \approx \mu_P/\rho^2$.

For an Earth orbiting satellite it is $\varepsilon = r/R \ll 1$. The angular position with respect to the Earth, β, is related to the angular position with respect to the third body, α, by the equation

$$\cos\alpha = \frac{R - r\cos\beta}{\rho} = \frac{R}{\rho}(1 - \varepsilon\cos\beta)$$

so that the perturbing acceleration can be expressed as

$$a_P = \frac{\mu_P}{\rho^2}\sqrt{1 + \frac{\rho^4}{R^4} - 2\frac{\rho}{R}(1 - \varepsilon\cos\beta)}$$

Table 4.1: **Perturbing attraction of a third body for r = 42,200 km.**

Disturbing Body	$\mu_P/\mu_\oplus$	R/r	$(a_P/a_\oplus)_{\max}$
Moon	0.0123	9.1	$3.3\ 10^{-5}$
Sun	332,946	$3.48\ 10^3$	$1.6\ 10^{-5}$

The distance from the perturbing body can be expressed using the (plane) cosine theorem, that is

$$\rho^2 = R^2 + r^2 - 2rR\cos\beta \quad \Rightarrow \quad \left(\frac{\rho}{R}\right)^2 = 1 - 2\varepsilon\cos\beta + \varepsilon^2$$

Using the binomial series expansion

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \mathcal{O}(x^3)$$

and neglecting terms of order ε^3, it is possible to write ρ/R as

$$\begin{aligned} \frac{\rho}{R} &= \left(1 - 2\varepsilon\cos\beta + \varepsilon^2\right)^{1/2} \approx 1 - \varepsilon\cos\beta + \frac{1}{2}\varepsilon^2 - \frac{1}{2}\varepsilon^2\cos\beta \\ \left(\frac{\rho}{R}\right)^4 &= \left(1 - 2\varepsilon\cos\beta + \varepsilon^2\right)^2 \approx 1 - 4\varepsilon\cos\beta + 2\varepsilon^2 + 4\varepsilon^2\cos\beta \end{aligned}$$

By substituing the above expressions into the equation for a_P, one obtains

$$a_P = \frac{\mu_P}{\rho^2}\sqrt{\varepsilon^2(1+3\cos^2\beta)} = \mu_P\frac{r}{R^3}\sqrt{1+3\cos^2\beta}$$

Taking the ration with the gravitational acceleration $a_\oplus$ due to the Earth, it is possible to express the relative importance of the perturbation,

$$\frac{a_P}{a_\oplus} = \frac{\mu_P}{\mu_\oplus}\ \frac{r^3}{R^3}\sqrt{1+3\cos^2\beta},$$

which increases with the third power of the satellite distance from the Earth. In particular the main perturbations on geostationary satellites are the disturbing attractions of Sun and Moon, the former being approximately a half of the second.

4.6.2 Asphericity of the Earth

The gravitational potential of the Earth in the two-body problem, expressed as μ/r, is valid under the assumption of a perfectly spherically symmetric body, but the Earth is oblate (*e.g.* flattened at the poles), bulged at the equator, and generally asymmetric because of density variations inside its volume.

The asphericity of the Earth is the dominant perturbation for satellites orbiting outside of the densest layers of the atmosphere, yet below an altitude of few tens thousands of kilometers. The excess of mass near the equator produces a slight torque on the satellite about the center of the Earth. The torque causes the orbit plane to precess, and the line of the nodes moves westward for direct orbits and eastward for retrograde satellites. Earth's oblateness also causes a rotation of the line–of—apsides.

The gravitational acceleration can be written as the gradient of a scalar function,

$$\mathcal{E}_g = -\frac{\mu}{r}\left\{1 - \sum_{n=2}^{\infty} J_n\left(\frac{r_\oplus}{r}\right)^n \mathcal{P}_n(\sin La) - \sum_{n=2}^{\infty}\sum_{m=2}^{\infty} J_{n,m}\mathcal{P}_n^m(\sin La)\ \cos\left[m\left(Lo - Lo_{n,m}\right)\right]\right\}$$

which describes the gravitational potential of the Earth in the rotating geocentrical equatorial frame. The function $\mathcal{P}_n(x)$ is Legendre's polynomial of degree n, that can be expressed by Rodrigues' formula as

$$\mathcal{P}_n(x) = \frac{1}{2^n\ n!}\ \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[\left(x^2-1\right)\right],$$

while $\mathcal{P}_n^m(x)$ is the associated Legendre polynomial of degree n and order m,

$$\mathcal{P}_n^m(x) = (-1)^m(1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}\left[\mathcal{P}_n(x)\right].$$

The Newtonian point–mass potential is thus joined by a series of terms in J_n, called zonal harmonics, and terms in $J_{n,m}$, called tesseral and sectorial harmonics for $n \neq m$ and $n = m$, respectively. The zonal harmonics, which describe the deviation in the North–South direction, are dependent on latitude only. The tesseral and sectorial harmonics causes the deviation in the East–West direction; the former are dependent on both latitude and longitude, the latter on longitude only. The terms with $n = 1$ are absent, as the center of the frame and the center of mass of the Earth coincide.

The even numbered zonal harmonics are symmetric about the equatorial plane and the odd numbered harmonics antisymmetric. The second zonal coefficient, J_2, describes the Earth oblateness and its value, $J_2 = 1.082637\ 10^{-3}$, is three order of magnitude greater than any other; J_3 describes the Earth as pear–shaped. The term $J_{2,1}$ is very small, as the axis of rotation of the Earth practically coincides with a principal axis of inertia; $J_{2,2}$ describes the ellipticity of the Earth equator.

The effects of tesseral and sectorial harmonics are averaged and generally cancelled by the rotation of the Earth inside the planar orbit of the satellite. An important exception is the geostationary satellites that maintain a fixed position relative to the Earth and are very sensitive to the East–West anomalies in the gravitational field.

For most practical orbits, with the above mentioned exception of geostationary ones, only the zonal harmonics are considered, *i.e.*, the gravitational field is assumed to be axi–symmetric and the reference frame does not need to rotate with the Earth.

In particular, if the method of general perturbation is applied retaining only the second zonal harmonic, J_2, three orbital elements exhibit changes after one orbital revolution:

$$\begin{aligned}
\Delta\Omega &= -3\pi J_2\left(\frac{r_\oplus}{p}\right)^2\cos i \\
\Delta\omega &= \frac{3}{2}\pi J_2\left(\frac{r_\oplus}{p}\right)^2\left(5\cos^2 i - 1\right) \\
\Delta M &= n\mathcal{T}_d + 3\pi J_2\left(\frac{r_\oplus}{p}\right)^2\cos i\frac{(1+e\cos\omega_0)^3}{1-e^2}
\end{aligned}$$

where $\mathcal{T}_d$ is the nodal or Draconian period. The regression of the line–of–nodes and the rotation of the line–of–apsides are both decreasing functions of the orbit altitude (*i.e.*, higher orbits are less affected by the non–sphericity of the Earth). The direction of the former rotation changes passing through the polar orbit ($i = \pi/2$); the semi–major axis rotates in the same direction of the satellite motion for $i < 63.4$ deg or $i > 116.6$ deg.

4.6.3 Aerodynamic drag

The motion of a satellite inside the Earth's atmosphere is accompanied by lift and drag forces. The former is negligible in most cases; the latter is in a direction opposite to the velocity vector relative to the atmosphere, with magnitude

$$a_D = \frac{1}{2}\rho w^2\frac{S}{m}C_D$$

The area S of the frontal section depends on the attitude of the satellite, and can vary between a minimum and a maximum value that can be significantly different. The drag coefficient is close to 2, as the flow is of the free–molecular type and drag is due to the molecular impact on the satellite surface. The atmospheric density rapidly decrease with altitude, but it depends also on season, local solar time, and solar activity, resulting barely predictable. As a consequence, the magnitude of aerodynamic drag can be only roughly estimated. It is the most important perturbation below 200 km altitude, whereas can be completely neglected above 1000 km.

A satellite flying an eccentric orbit with perigee altitude below 700 km is subjected to aerodynamic forces that are much stronger at the periapsis than at the apoapsis, due to the combined effects of higher velocity and density during perigee passes. As a consequence, the apogee is rapidly lowered; after the orbit has become near-circular ($r \approx$ const) it is possible to express the orbit decay from the expression of total energy:

$$-\frac{\mu}{2a} = \frac{v^2}{2} - \frac{\mu}{r} \quad \Rightarrow \quad \frac{\mu}{2a^2}\dot{a} = v\dot{v} = -v\left(\frac{1}{2}\rho w^2 \frac{S}{m} C_D\right)$$

By approximating the velocity with respect to the atmosphere with the orbital velocity, it is possible to assume that

$$w^2 \approx v^2 \approx \frac{\mu}{a} \quad \Rightarrow \quad \dot{a} = -\sqrt{\mu a}\ \rho \frac{S}{m} C_D$$

The rate of reduction of orbit altitude is ruled by atmospheric density and becomes faster as the satellite flyes lower.

4.6.4 Radiation pressure

The direct solar radiation, the solar radiation reflected by the Earth, and the radiation emitted by the Earth itself may be regarded as a "beam" of photons that hits the satellite surface, thus imparting a momentum. The corresponding radiation pressure is quite small and usually negligible, except for spacecraft with high volume/mass ratio (Bate 9.7.3, p. 424–425). The magnitude of the resulting acceleration can be expressed approximately as

$$a_{RP} = fS/m$$

where m is the satellite mass, S the frontal area exposed to the solar radiation pressure, and $f = 4.5\ 10^{-8}\ \mathrm{m^3 kg^{-1} s^{-2}}$.

4.6.5 Electromagnetic effects

A satellite may acquire electrostatic charges in the ionized high-altitude atmosphere; electric currents circulate in its subsystems. Interactions with the Earth's magnetic field are possible but generally negligible. Control of satellite formations by electrostatic fields obtained by charging the satellites is currently investigated. In general, electromagnetic effects have a deeper impact on satellite attitude dynamics, and can be used for attitude control of small satellites.

4.7 Geographical constraints

Satellite operations for communication, intelligence, navigation, and remote sensing are evidently related to the Earth's surface. Any spacecraft with no exception, however, needs to download data to ground stations, which in turn detect its position and velocity. Satellite reconfiguration and orbital maneuvers usually occur in sight of a command station.

The analysis of the ground track is therefore essential. According to the equation

$$\cos i = \cos\delta \sin\psi$$

the ground track crosses the equator at an angle equal to the orbital inclination. The maximum latitude North or South of the equator that a satellite pass over just equals the orbit plane inclination i ($\pi - i$, for retrograde orbits). A global surveillance satellite should be in polar orbit to overfly the Earth's entire surface. The Earth rotation displaces the ground track westward by the angle the Earth turns during one orbital period. If the time required for n complete rotations of the Earth on its axis (n sidereal days) is an exact multiple of the orbital period, the satellite will retrace the same path over the Earth (a correction is required to account for the additional displacement due to the regression of the line-of-nodes). This is a desirable property for a reconnaissance satellite; in general it might be necessary when the number of available ground station is limited and the orbit has a great inclination.

The geographical position of the launch site (see Appendix to the present Chapter) is even more important. Latitude and launch azimuth influence both the rocket overall performance and the orientation of the orbit plane. The first stage booster falls to the Earth several hundreds kilometers downrange and the launch azimuth is constrained to assure the passage over uninhabited and traffic–free regions. According to section 3.9, the same rocket has a greater payload for launches from a base close to the equator ($\delta \to 0$), and for an easterly launch ($\psi = \pi/2$). The orbital inclination will be the minimum achievable and equal to the latitude of the launch site. Greater inclinations, up to the limit value $\pi - \delta$, are achievable from the same site with payload reduction; retrograde orbits require a westerly launch.

The same inclination can be achieved with two different launch azimuths, if both are inside the safety constraints. In some cases the longitude of the ascending node is also assigned; the launch is possible twice a day when the launch site passes through the envisaged orbit plane, and the azimuth constraints can be fulfilled. The lift–off time is strictly fixed for maximum payload. However a fraction of the payload is traded for propellant, thus adding maneuvering capabilities and allwoign for an out–of–plane ascent. A launch window of several minutes is created around the theoretical launch time: the rocket departs as soon as all the requirements for a reliable launch are met.

4.8 Practical orbits

The satellite orbits can be classified on the basis of their altitude. A Low Earth Orbit (LEO) is confined between 200 and 600 km, the lower limit being related to the increase in atmospheric drag for lower altitudes, the upper one due to the Van Allen radiation belts, which are harmful to humans and detrimental for solar cells and some instruments. As a consequence, also orbit eccentricity must also be quite low ($e < 0.03$). A LEO assures safe manned flight, and constitutes an efficient parking orbit for an interplanetary mission. Moreover it guarantees high photographic resolution at the price of limited ground visibility.

A High Earth Orbit (HEO), above 10,000 km altitude, is drag free and safe from radiation; a large part of the Earth's surface can be seen at one time. In the middle lie Medium Earth Orbits (MEO), which represents a compromise that permits adequately high resolution sensing of Earth's surface and low transmission power with a sufficiently wide ground visibility to limit the number of satellites necessary to cover the entire surface.

Orbit circularity is an appreciate characteristic, but is rarely mandatory, as it implies the costs of a precise injection or the need of corrective maneuvers. High-ellipticity orbits are often used for space probes, as the spacecraft operates at high altitudes for a large part of its lifetime and the launch energy is lower in respect to a circular orbit with the same apogee.

A satellite on a circular direct equatorial orbit with a period equal to a sidereal day is motionless with respect to the ground. At an altitude of 42,200 km this geostationary (GEO) satellite sees almost a half of the Earth's surface. Three satellites, 120 deg apart, cover the whole Earth, except for small regions around the poles, while being able to communicate between

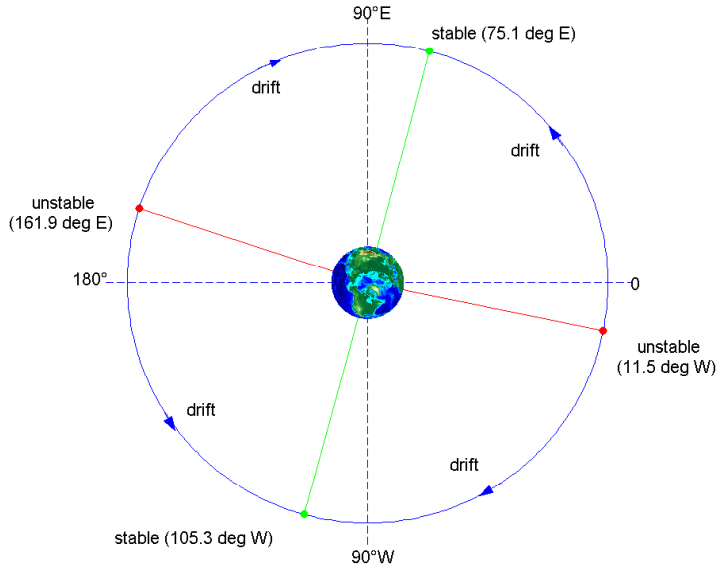

Figure 4.4: **Stable and unstable longitudes for GEO satellites.**

them. Such a satellite is extremely useful for communications, altough high power antennas are necessary. The reconnaissance role is difficult, if not impossible, as photographic resolution is poor from an altitude of 35800 km.

A geosynchronous satellite has the semi-major axis corresponding to the period of a sidereal day. If it flies on a circular inclined orbit, the satellite has a "figure–eight" ground track between latitudes equal to the inclination and approximately along the same meridian. A non-zero eccentricity would distort the eight–shape, the resulting apparent trajectory with respect to the Eart's surface depending also on the argument of perigee.

Geostationary satellites is perturbed by solar and lunar attractions, that increase the inclination at a rate close to 1 deg per year, with a minor variation related to the inclination of the lunar orbit plane to the equatorial plane, which oscillates between 18.3 deg and 28.6 deg with a period of 18.6 years. The elliptic shape of the Earth equator (described by the second sectorial coefficient $J_{2,2}$) pulls the spacecraft towards one of the two equilibrium longitudes. Maintaining the satellite nominal position requires periodic maneuvers: North-South station-keeping to compensate for the lunar and solar perturbing actions; East-West station-keeping to counteract the Earth asphericity.

Communication satellites for high–latitude regions of the northern hemisphere use high-eccentricity *Molniya orbits* with a period equal to 12 sidereal hours (even though 8-hour or 24-hour periods could also be chosen). The spacecraft spends the most part of the orbital period near the apogee, located over the highest latitude north if $\omega = -\pi/2$. A constellation of three satellites with displaced ascending nodes ($\Delta\Omega = 120$ deg) allows for the continuous coverage of the target region. On the converse, an inclination $i = 63.4$ deg is mandatory in order to avoid the rotation of the line–of–apsides induced by the J_2 effect.

A satellite is on a Sun-synchronous orbit if the line–of–nodes rotates in an easterly direction with the rate of one revolution per year (0.986 deg per day) that equals the mean angular motion of the Earth about the Sun. This rate can be naturally obtained from the Earth oblateness by carefully selecting semi–latus rectum and inclination of a retrograde Low Earth Orbit, for instance a 600 km altitude circular LEO with $i = 97.76$ deg. Sun-synchronous satellites are very attractive for Earth observation because of constant ground lighting conditions; a suitable orbital period would add a synchronism with the ground. A satellite with the line–of–nodes perpendicular to the Sun–Earth line will move approximately along two meridians at local

Table 4.2: **Launch sites in former Soviet Union.**

Launch Site	Country	Latitude	Longitude
Baikonur	Kazakhstan	45.6° N	63.4° E
Plesetsk	Russia	62.8° N	40.4° E
Kapustin Yar	Russia	48.5° N	45.8° E
Svobodniy	Russia	51.7° N	128.0° E

Websites.
http://www.aerospaceguide.net/spacebusiness/launchfacilities.html
http://www.orbireport.com/Linx/Sites.html
http://www.spacetoday.org/Rockets/Spaceports/LaunchSites.html
http://www.russianspaceweb.com/centers.html
http://www.esa.int/SPECIALS/ESA_Permanent_Mission_in_Russia/SEMHIZW4QWD_0.html

sunrise and sunset, respectively, a favorable condition for intelligence. The spacecraft could also remain continuously in sunlight, thus resorting only on solar cells for power. One should note that a short eclipse behind either pole is possible during few months every year.

Appendix – Launch Sites

Baikonur Cosmodrome is the oldest space–launching facility; it is located about 200 miles southwest of the small mining town of Baikonur, in a region of flat grasslands in the former Soviet republic of Kazakhstan, northeast of the Aral Sea at 45.6° N, 63.4° E. Baikonur Cosmodrome is the launch complex where Sputnik 1, Earth's first artificial satellite, was launched. The rocket that lifted Yuri Gagarin, the first human in orbit, was also launched from Baikonur. In fact, all Russian manned missions are launched from Baikonur, as well as all geostationary, lunar, planetary, and ocean surveillance missions. All Space Station flights using Russian launch vehicles are launched from Baikonur, since all nonmilitary launches take place there. Baikonur is also the only Russian site that has been used to inject satellites into retrograde orbits.

Plesetsk Cosmodrome, located at 62.8° N, 40.4° E, has launched the most satellites since the beginning of the Space Age. Plesetsk's location makes it ideal for launching into polar or high-inclination orbits (63 to 83° inclination), typical for military reconnaissance and weather satellites. It continues to be highly active today, especially for military launches and all Molniya-class communications satellites.

Kapustin Yar is the Russian oldest missile test site, located at 48.5° N, 45.8° E. Although used quite often for launches of smaller Cosmos satellites during the 1960s, the number of launches from this site fell dramatically during the 1970s and 1980s to about one orbital launch per year.

Svobodniy Cosmodrome, located 51°42' N, 128°00' E, (Minimum Inclination: 51.0 degrees. Maximum Inclination: 110.0 degrees) was planned during the 1990s to accommodate the new modular Angara launch vehicles in the medium and heavy categories, as the break-up of the Soviet Union left the main Russian launch site (Baikonur) on foreign territory, while Plesetsk Cosmodrome did not have facilities for large launch vehicles and was not suited for support of launches into lower-inclination orbits. Sufficient funding was not available for this massive project, and by 2000 conversion of unfinished Zenit pads at Plesetsk for use with Angara was underway. It seems that Svobodniy will be limited to launches if Rokot, Strela and Start-1 rockets, while Plesetsk would be developed as the main Russian cosmodrome. Geosynchronous launches from Plesetsk could be made economically by using Moon's gravity to change the

orbital plane of the satellite and Svobodniy would still be needed to reach the 51.6 degree orbit of the International Space Station.

Chapter 5

Orbital Manoeuvres

5.1 Introduction

Orbital manoeuvres are carried out to change some of the orbital elements. This can be necessary

- when the final orbit is achieved via a parking orbit;
- to correct injection errors;
- to compensate for orbital perturbations (stationkeeping).

The task is usually accomplished by the satellite propulsion system, and occasionally, in the first two cases, by the rocket last stage. Elementary manoeuvres, that involve a maximum of three parameters and require a maximum of three impulses, are analyzed in this chapter. When it is necessary, two or more elementary manoeuvres are combined and executed according to a more complex strategy that permits, in general, a minor propellant consumption. Nevertheless, it has been demonstrated that any optimal impulsive manoeuvre requires a maximum of four burns.

The propellant consumption for the manoeuvre (m_p) is evaluated after the total velocity change has been computed. The equation presented in Section 4.2 is easily integrated under the hypothesis of constant effective exhaust velocity c

$$\mathrm{d}v = -c\frac{\mathrm{d}m}{m} \quad \Rightarrow \quad \Delta v = c\log\frac{m_i}{m_f} \quad \Rightarrow \quad m_f = m_i \exp\left(-\frac{\Delta v}{c}\right)$$

and provides the relationship between the spacecraft final and initial mass, which is known as *rocket equation* or *Tsiolkovsky's equation*. Therefore

$$m_p = m_i - m_f = m_i\left[1 - \exp\left(-\frac{\Delta v}{c}\right)\right]$$

5.2 One-impulse manoeuvres

A single velocity impulse is sufficient to change all of the orbital elements. Unfortunately, it is not always possible to avoid changes in more than one orbit parameter at the same time. As an example, if the initial orbit is circular, it is not possible to increase its semi-major axis (that is, its radius) maintaining zero eccentricity, so that transfer between two circula orbits will require at least two pulses even in the co–planar case.

Simple cases are analyzed in this section where subscripts 1 and 2 denote characteristics of the initial and final orbit, respectively. In the figures, dashed lines indicate the orbit prior to the manoeuver, solid lines the orbit after the velocity increment.

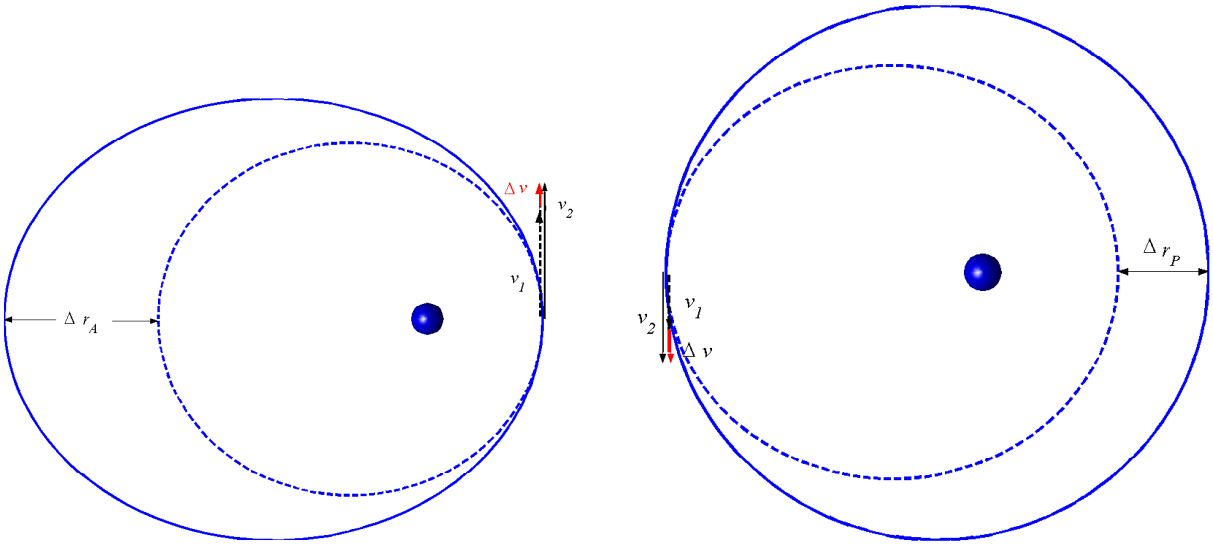

Figure 5.1: **Single–impulse manoeuvre for apogee (left) and perigee (right) increment.**

5.2.1 Adjustment of perigee and apogee height

An efficient way of changing the height of perigee and apogee relies on a velocity increment provided at the opposite (first) apsis. Misalignment losses and rotation of the semi-major axis are avoided. Once the required variation Δz of the second apsis altitude is known, one easily deduces the new length of the semi-major axis $a_2 = a_1 + \Delta z/2$, which is used in the energy equation to compute the velocity v_2 at the first apsis after the burn, which therefore provides the required $\Delta v = v_2 - v_1$.

For small variations of the second apsis radius, the differential relationship of Section 4.4 can be applied, obtaining

$$\Delta z = 2\Delta a \approx \frac{4a^2}{\mu} v_1 \Delta v$$

One should note that in some manoeuvres the apsides may interchange their role (from periapsis to apoapsis, and vice versa).

This type of manoeuvre may be required for stationkeeping of a Low Earth elliptical Orbit. The effects of atmospheric drag will be higher were the height of the orbit is lowest, that is, close to periapsis passage. The effect of atmospheric drag can be modeled as a negative Δv that slightly decreases the apogee height at every perigee pass, thus reducing both the semi–major axis and the orbit period. In order to compensate for this effect, a station–keeping manoeuvre is required where a positive Δv at perigee pass takes back the apogee radius to the nominal value. It should be noted that a positive Δv at apogee pass will rise the perigee of the orbit.

5.2.2 Simple rotation of the line-of-apsides

A simple rotation $\Delta\omega$ of the line of apsides without altering size and shape of the orbit is obtained by means of one impulsive burn at either point where the initial and final ellipses intersect on the bisector of the angle $\Delta\omega$. The polar equation of the trajectory and the constant position of the burn point in a non-rotating frame imply

$$\begin{aligned} r_2 = r_1 \quad &\Rightarrow \quad \nu_2 = 2k\pi - \nu_1 \quad , \qquad k = 0, 1 \\ \vartheta_2 = \vartheta_1 \quad &\Rightarrow \quad \nu_2 = \nu_1 + \omega_1 - \omega_2 = \nu_1 - \Delta\omega \end{aligned}$$

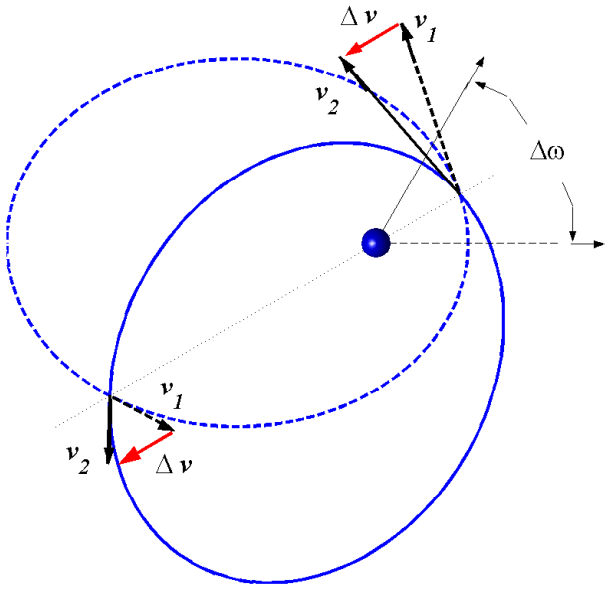

Figure 5.2: **Single–impulse manoeuvre for periapsis rotation.**

These relations imply that the point for the application of the velocity increment by the propulsive system is

$$\nu_1 = k\pi + \Delta\omega \quad , \qquad k = 0, 1$$

Energy and angular momentum are unchanged by the manoeuvre; this corresponds to conserving, respectively, magnitude and tangential component of the velocity. Therefore, the only permitted change is the sign of the radial component $\dot{r} = v_r$. The velocity change required is equal to

$$\Delta v = 2v_r = 2\frac{\mu}{h}e|\sin\nu_1| = 2\sqrt{\frac{\mu}{h}}\,\frac{e}{\sqrt{1+e}}\left|\sin\frac{\Delta\omega}{2}\right|$$

It should be noted that the Δv is the same in either point eligible for the manoeuvre. The rightmost term can been obtained bu using the relationship $h^2/p = a(1-e^2) = r_P(1+e)$.

5.2.3 Simple plane change

The change of the orientation of the orbit plane requires that the velocity increment $\Delta\vec{\boldsymbol{v}}$ has a component perpendicular to the plane of the orbit prior to the manoeuvre. A simple plane change rotates the orbital plane by means of one impulsive burn ($r = \text{const}$), without altering size and shape of the orbit. Energy and magnitude of the angular momentum are unchanged, which corresponds to conserving, respectively, magnitude and tangential component of the velocity. Therefore, the radial component v_r must also remain constant, while the tangential component v_θ is rotated of the desired angle $\Delta\psi$. From the resulting isosceles triangle in the horizontal plane one obtains (Fig. 5.3)

$$\Delta v = 2v_\theta \sin(\Delta\psi/2)$$

The analysis of the azimuth equation presented in Section 4.7,

$$\cos i = \cos\delta \sin\psi$$

indicates that a simple plane change at $\delta \neq 0$ implies $\Delta i < \Delta\psi$ (Fig. 5.4); moreover the inclination after the manoeuvre cannot be lower than the local latitude ($i_2 \geq \delta$). In a general case the manoeuvre changes both inclination and longitude of the ascending node. If the plane

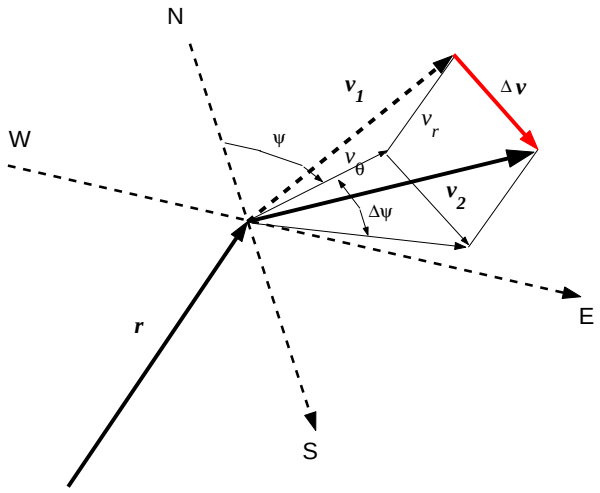

Figure 5.3: **Rotation $\Delta\psi$ of the transverse component of velocity.**

change aims to change the inclination, the most efficient manoeuvre is carried out when the satellite crosses the equatorial plane (therefore at either node) and Ω remains unchanged.

A plane rotation requires a significant velocity change (10% of the spacecraft velocity for a 5.73 deg rotation) with a considerable propellant expenditure associated with the manoeuvre, with no energy gain. Gravitational losses are not relevant and the manoeuvre is better performed where the velocity is low, inasmuch as a smaller velocity increment is sufficient to rotate the velocity vector, when the initial value of its magnitude is smaller.

In most cases, by means of a careful mission planning, a specific manoeuvre for plane change can be avoided or executed at a lower cost in the occasion of a burn aimed to change the spacecraft energy.

5.2.4 Combined change of apsis altitude and plane orientation

Consider an adjustment of the apsis altitude combined with a plane rotation $\Delta\psi$, which is therefore the angle included between the vectors $\vec{\boldsymbol{v}}_1$ and $\vec{\boldsymbol{v}}_2$. Without any loss of generality, suppose $v_2 > v_1$. If the manoeuvres are separately performed, the rotation is conveniently executed before the velocity has been increased, and the total velocity change is therefore given by

$$\Delta v_s = \Delta v_r + \Delta v_e = 2v_1 \sin(\Delta\psi/2) + (v_2 - v_1)$$

The velocity increment of the combined manoeuvre is given by

$$\Delta v_c = \sqrt{v_2^2 + v_1^2 - 2v_1 v_2 \cos\Delta\psi}$$

and the benefit achieved $\Delta v_s - \Delta v_c$ is presented in Fig. 5.4 for different values of $\Delta v_e/v_1$. One should note that, for small variations of the velocity vector azimuth angle, $\Delta\Psi$ it is $\cos\Delta\psi \approx 1$, so that $\Delta v_c \approx v_2 - v_1 = \Delta v_e$. This means that the plane rotation is actually free, in terms of fuel expenditure.

5.3 Two-impulse manoeuvres

In this case the manoeuvre starts at point 1 on the initial orbit, where the spacecraft is inserted into a transfer orbit (labeled with the subscript t) that ends at point 2 with injection on the final orbit.

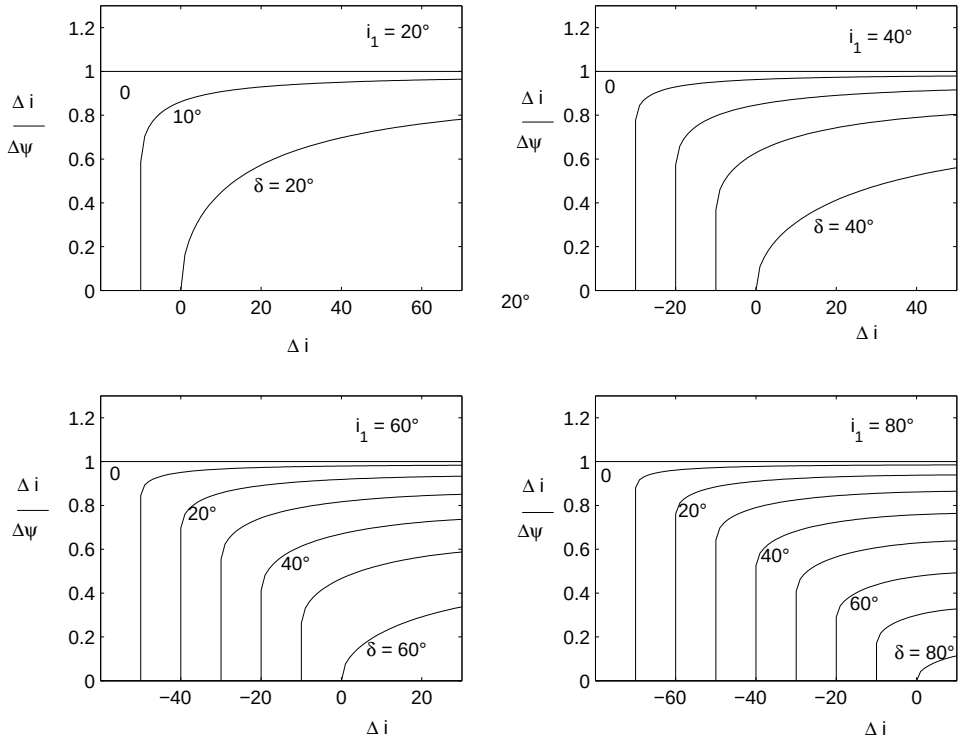

Figure 5.4: **Efficiency of an inclination change carried out at different latitudes.**

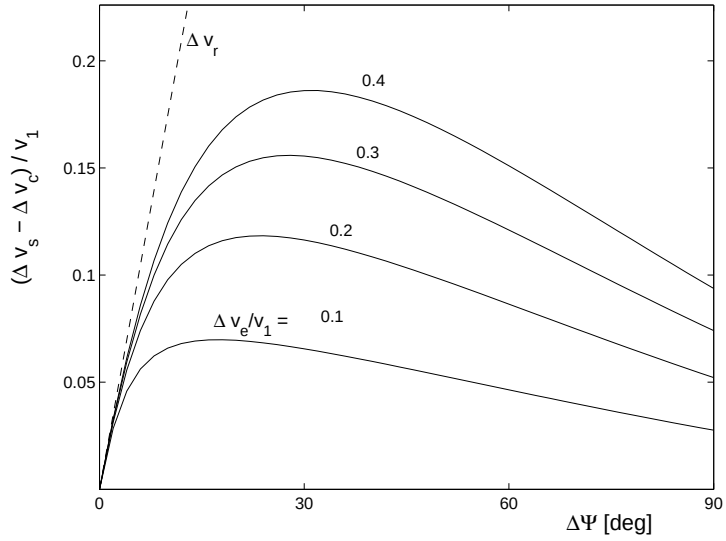

Figure 5.5: **Benefit of the combined change of apsis altitude and plane orientation.**

5.3.1 Change of the time of periapsis passage

A change ΔT of the time of periapsis passage permits to phase the spacecraft on its orbit. When the orbit is circular and the line of apsis is not defined, the manoeuvre is called orbit phasing, as the spacecraft is transferred onto a different angular position on the same circular orbit. This manoeuvre is important for geostationary satellites that need to get their design station and keep it against the East-West displacement caused by the Earth's asphericity.

Assuming $\Delta T > 0$, the manoeuvre is accomplished by moving the satellite on an outer waiting orbit where the spacecraft executes n complete revolutions. The period $\mathcal{T}_t$ of the waiting orbit is selected so that the following equation is satisfied,

$$\Delta T = n(\mathcal{T}_t - \mathcal{T}_0) = n\Delta\mathcal{T}$$

where $\mathcal{T}_0$ is the period of the nominal orbit. When the change ΔT required is higher that $\mathcal{T}_0/2$ (half of the nominal orbit period) it is more convenient to put the spacecraft on a lower, faster orbit in order to catch up the desired position along the orbit, instead of waiting for it. In this case it is

$$\mathcal{T}_0 - \Delta T = n(\mathcal{T}_0 - \mathcal{T}_t) = n\Delta\mathcal{T}$$

In this latter case, it is necessary that the perigee of the waiting orbit remains high enough above the Earth atmosphere in order not to cause energy dissipation at perigeee passes because of atmospheric drag.

Two velocity increment Δv are needed, equal in magnitude, but in opposite directions: the first puts the spacecraft on the waiting orbit; the second restores the original trajectory. According to general considerations, the engine thrust is applied at the perigee and parallel to the spacecraft velocity. The required Δv is smaller if $\Delta\mathcal{T}$ is reduced by increasing the number n of waiting orbits. A time constraint must be enforced in order to obtain a meaningful solution and avoid a manoevure wheren an infinite number of revolution with and infinitesimal $\Delta\mathcal{T}$ is necessaty to complete the transition.

The problem is equivalent to the rendezvous with another spacecraft on the same orbit. One should note that thrust apparently pushes the chasing spacecraft away from the chased one.

5.3.2 Transfer between circular orbits

Consider the transfer of a spacecraft from a circular orbit of radius r_1 to another with radius r_2, without reversing the rotation. Without any loss of generality, assume $r_2 > r_1$ (the other case only implies that the velocity-change vectors are in the opposite direction). The transfer orbit (subscript t) must intersect (or be tangent) at some point to both the circular orbits, so that the following inequalities must be satisfied:

$$r_{P_t} = \frac{p_t}{1+e_t} \le r_1 \quad ; \quad r_{A_t} = \frac{p_t}{1-e_t} \ge r_2$$

As a consequence the admissible values for eccentricity and semi–latus rectum, e_t and p_t, of the transfer orbit lie in the shadowed area of Fig. 5.6, inside which a suitable point is selected.

In order to determine the Δv necessary to complete the transfer, it is necessary to compute the energetic parameters of the transfer orbit,

$$\mathcal{E}_t = -\frac{\mu}{2a_t} = -\mu\frac{1-e_t^2}{2p_t} \quad ; \quad h_t = \sqrt{\mu p_t}$$

and then velocity and flight path angle after the first burn,

$$v_1^2 = 2\left(\frac{\mu}{r_1} + \mathcal{E}_t\right) \quad ; \quad \cos\varphi_t = \frac{h_t}{r_1 v_1}$$

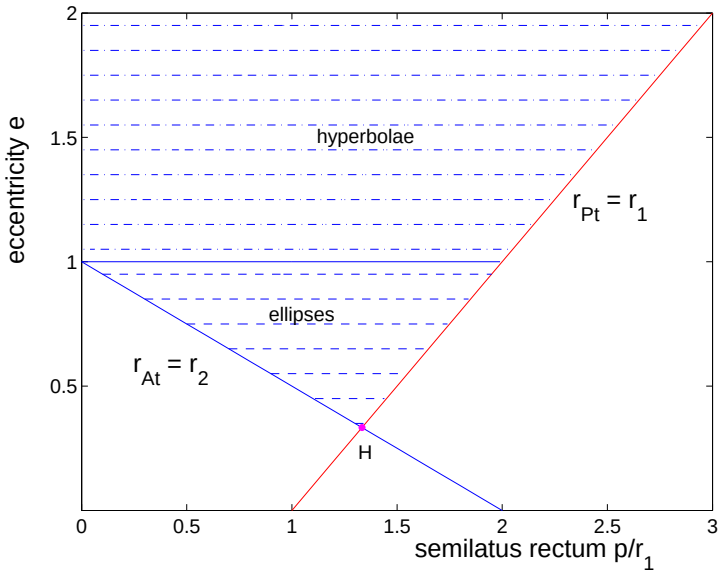

Figure 5.6: **Admissible parameters for a transfer orbit between circular orbits** ($r_2 = 2r_1$)**.**

The first velocity increment is thus

$$\Delta v_1^2 = v_1^2 + v_{c_1}^2 - 2v_1 v_{c_1} \cos\varphi_1$$

The second velocity increment at point 2 is evaluated in a similar way.

5.3.3 Hohmann transfer

The minimum velocity change required for a two-burn transfer between circular orbits corresponds to using an ellipse, which is tangent to both circles, such that $r_{P_t} = r_1$ and $r_{A_t} = r_2$. In this case the transfer orbit has a semi–major axis and energy equal to

$$a_t = \frac{r_1 + r_2}{2} \quad ; \quad \mathcal{E}_t = -\frac{\mu}{r_1 + r_2}$$

On leaving the inner circle, the velocity, parallel to the circular velocity $v_{c_1} = \sqrt{\mu/r_1}$, is

$$v_1^2 = \mu\left(\frac{2}{r_1} - \frac{1}{a_t}\right) = 2\frac{\mu}{r_1}\left(1 - \frac{r_1}{r_1 + r_2}\right) = v_{\text{esc}_1}^2 \frac{r_2}{r_1 + r_2}$$

so that the velocity increment provided by the first burn is

$$\Delta v_1 = v_1 - v_{c_1} = v_{c_1}\left(\sqrt{\frac{2r_2}{r_1 + r_2}} - 1\right)$$

The velocity on reaching the target outer orbit at $r = r_2 = r_{A_t}$, that is

$$v_2^2 = \mu\left(\frac{2}{r_2} - \frac{1}{a_t}\right) = 2\frac{\mu}{r_2}\left(1 - \frac{r_2}{r_1 + r_2}\right) = v_{\text{esc}_2}^2 \frac{r_1}{r_1 + r_2} = v_1^2 \frac{r_1^2}{r_2^2},$$

is again parallel, but smaller, than the circular velocity $v_{c_2} = \sqrt{\mu/r_2}$. Therefore a second burn is required, such that

$$\Delta v_2 = v_{c_2} - v_2 = v_{c_2}\left(1 - \sqrt{\frac{2r_1}{r_1 + r_2}}\right)$$

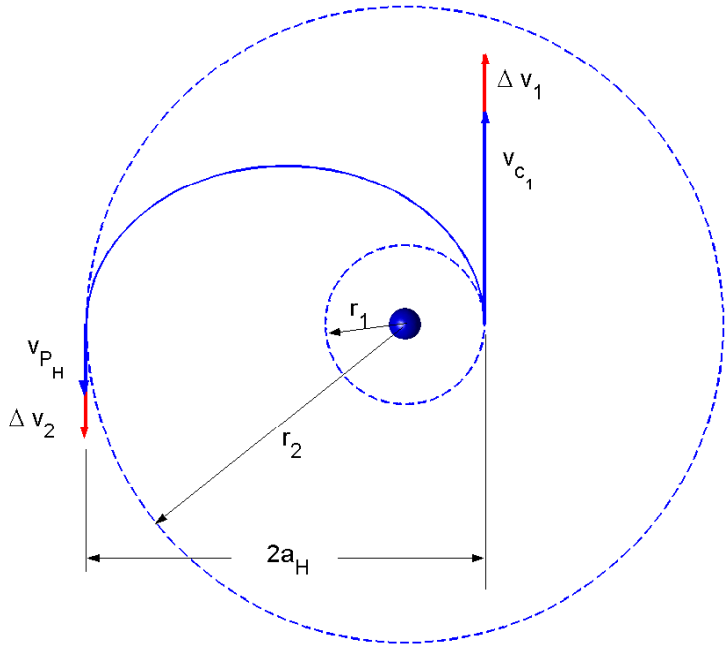

Figure 5.7: **Hohmann transfer between circular orbits.**

The time-of-flight is just half the period of the transfer ellipse,

$$\Delta T_t = \frac{\mathcal{T}_t}{2} = \pi\sqrt{\frac{a_t^3}{\mu}}$$

One should note that the Hohmann transfer requires the minimum total velocity increment $\Delta v_{\text{tot}} = \Delta v_1 + \Delta v_2$, that is, it is the cheapest two–burn transfer between circular orbits. It is also the two–burn transfer requiring the longest time. Increasing the apogee of the transfer orbit, keeping constant its perigee radius reduces significantly the time-of-flight. The transfer orbit remains tangent to the lower circular orbit but intersect the target orbit at a higher velocity (Fig. 5.8). The resulting transfer segment is shorter and the spacecraft flies along it at a higher velocity. A greater tangential Δv_1 is required upon leaving the initial orbit, in order to achieve a higher energy. Also a higher Δv_2 is necessary for circularization upon reaching $r = r_2$, as the velocity of the transfer orbit will not be tangential to the circulat target one, as it is at the apogee of the Hohmann transfer orbit. If the initial velocity increment Δv_1 is sufficiently high, the transfer may follow a parabolic or a hyperbolic trajectory.

As it is evident from Fig. 5.8, the total velocity increment Δv_{tot} needed is always higher than the fuel–optimal Hohmann transfer, but an increment of only 10% of the Δv employed for the transfer allows for a 22% reduction of the transfer time. The advantage is reduced considerably on faster orbits: if the Δv_{tot} used is 1.5 times that of the Hohmann transfer, the transfer time is reduced by only 40%, and using twice as much Δv_{tot}, the reduction is only 55%.

5.3.4 Noncoplanar Hohmann transfer

A transfer between two circular inclined orbits is analyzed ($r_2 > r_1$); a typical example of application is the *geostationary transfer orbit* (GTO) that moves a spacecraft from an inclined LEO to an equatorial GEO. The axis of the Hohmann ellipse coincides with the intersection of the initial and final orbit planes. Both impulses provide a combined change of apsis altitude and plane orientation (see Section 5.2.4). The greater part of the plane change is performed at the

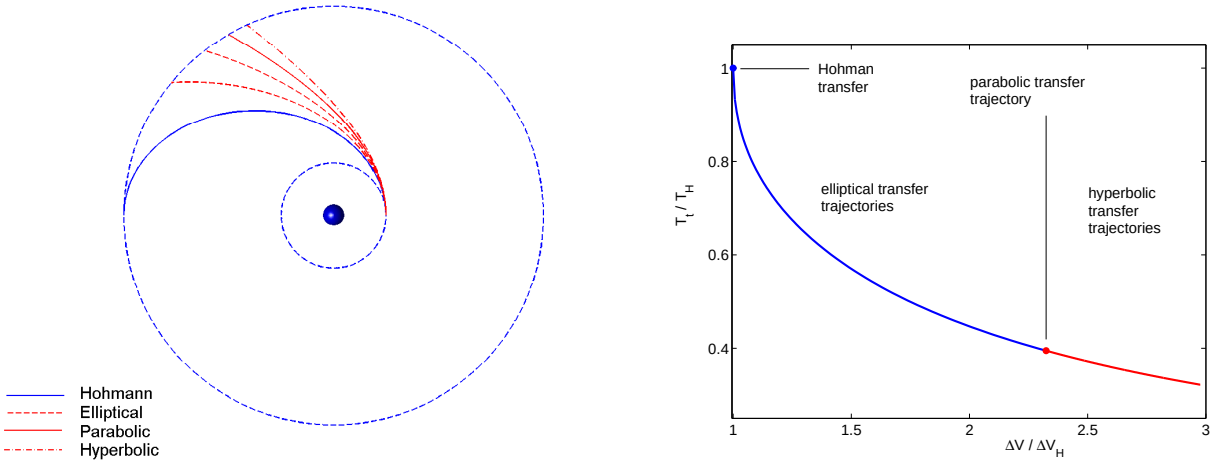

Figure 5.8: **Transfer between circular orbits ($r_2 = 4r_1$) for increasing tangential Δv_1: trajectories (left) and $\Delta v_{\rm tot}$ vs transfer time (right).**

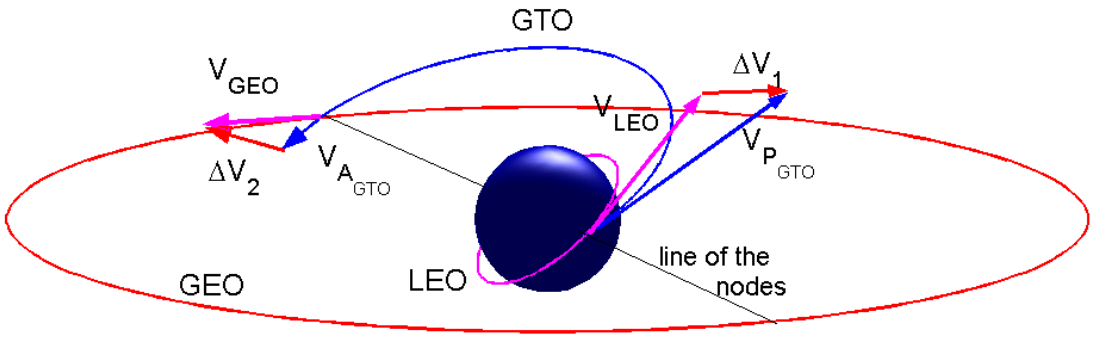

Figure 5.9: **Non–coplanar Hohmann transfer.**

apogee of the transfer orbit, where the spacecraft velocity attains the minimum value during the manoeuvre. Nevertheless, even for $r_2 \gg r_1$, a small portion of the rotation (typically 10% for LEO–to–GEO transfers) can be attained during the perigee burn almost without any additional cost.

5.4 Three–impulse manoeuvres

In special circumstances, some manoeuvres, which have been analyzed in the previous sections, are less expensive if executed according to a three–impulse scheme, which is essentially a combination of two Hohmann transfers; subscript 3 denotes the point where the intermediate impulse is applied.

5.4.1 Bielliptic transfer

The cost of the Hohmann transfer does not increase monotonically for increasing final orbit radius, but it reaches a maximum for $r_2 = r_H = 15.58 r_1$ (Fig. 5.10). Beyond this value, it is convenient to begin the mission on a transfer ellipse with apogee at $r_3 > r_2$, where the spacecraft trajectory is not circularized, but a smaller Δv_3 takes the spacecraft directly onto a Hohmann–transfer–like trajectory towards the target orbit of radius r_2, where the spacecraft is slowed down to circularize its orbit. The larger is r_3, the smaller the total cost of the *bielliptic manoeuvre*;

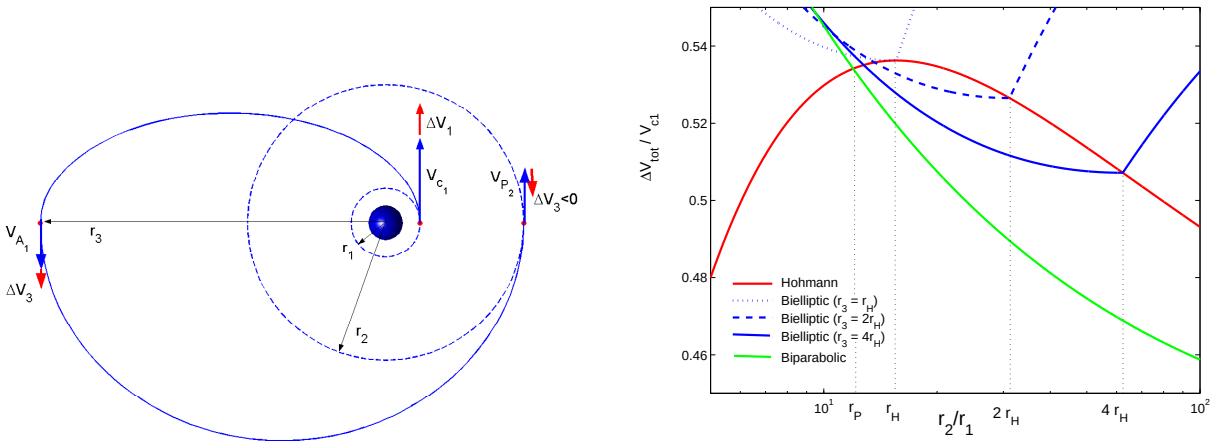

Figure 5.10: **Bielliptic orbit transfer: trajectory (left) and comparison with Hohmann transfer (right).**

one easily realizes that the minimum total $\Delta v_{\rm tot}$ is obtained with a biparabolic transfer. In this case two impulses transfer the spacecraft from a circular orbit to a minimum energy escape trajectory and vice versa; a third infinitesimal impulse is given at infinite distance from the main body to move the spacecraft between two different parabolae.

The biparabolic transfer has better performance than the Hohmann transfer for $r_2 > 11.94 r_1$. If the final radius r_2 is between r_B and r_H, also the bielliptic manoeuvre performs better than the Hohmann transfer if the intermediate radius r_3 is sufficiently great (Fig. 5.10).

The biparabolic transfer permits a maximum Δv reduction of about 8% for $r_2 \approx 50 r_1$, but the minor propellant consumption of bielliptic and biparabolic transfers is achieved at the cost of a considerably increment of the flight time. A three-impulse manoeuvre is seldom used for transfers between coplanar orbits; it becomes more interesting for noncoplanar transfer as a large part of the plane change (expensive in terms of fuel consumption) can be performed during the second impulse, far away from the central body, when the velocity is small.

5.4.2 Three-impulse plane change

A plane change can be obtained using two symmetric Hohmann transfers that move the spacecraft to and back from a far point where a cheaper rotation is executed, In particular, any arbitrary rotation is possible with an infinitesimally small velocity increment at an infinite distance from the central body. Assume that the spacecraft is in a low-altitude circular orbit. The velocity change of a simple plane rotation $\Delta v_{\rm spr}$ is compared to the $\Delta v_{\rm bp}$ required to enter and leave an escape parabola (biparabolic trajectory,

$$\Delta v_{\rm spr} = 2 v_{c_1} \sin(\Delta\psi/2) \quad ; \qquad \Delta v_{\rm bp} = 2(v_{\rm esc} - v_{c_1}) = 2(\sqrt{(}2) - 1) v_{c_1}$$

The biparabolic transfer is more convenient for deviation of the flight path larger than $\Delta\psi = 48.94$ deg. However a bielliptic transfer performs better than a single–burn rotation between 38.94 deg and 60 deg (dashed line in Fig. 5.11). Moreover, a small fraction ψ_1 of the total rotation $\Delta\psi$ can be efficiently obtained upon leaving the circular orbit and then again on reentering it (see 5.2.4). In this case the three–impulse bielliptic plane change is convenient for any amount of rotation, until the biparabolic manoeuvre takes over.

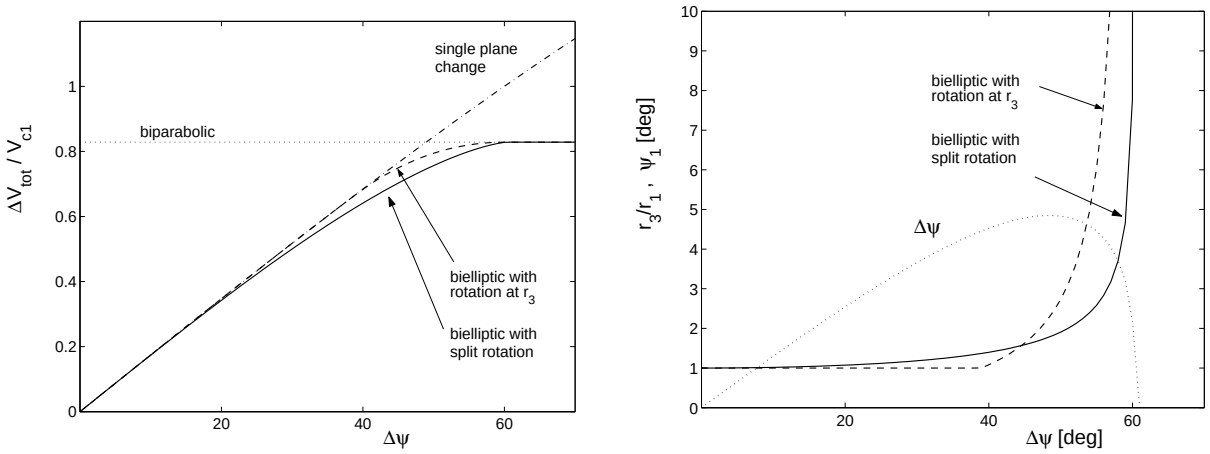

Figure 5.11: **Three–impulse plane rotation for circular orbits: Δv for different strategies (left) and optimal parameters r_3 and ψ_1 (right).**

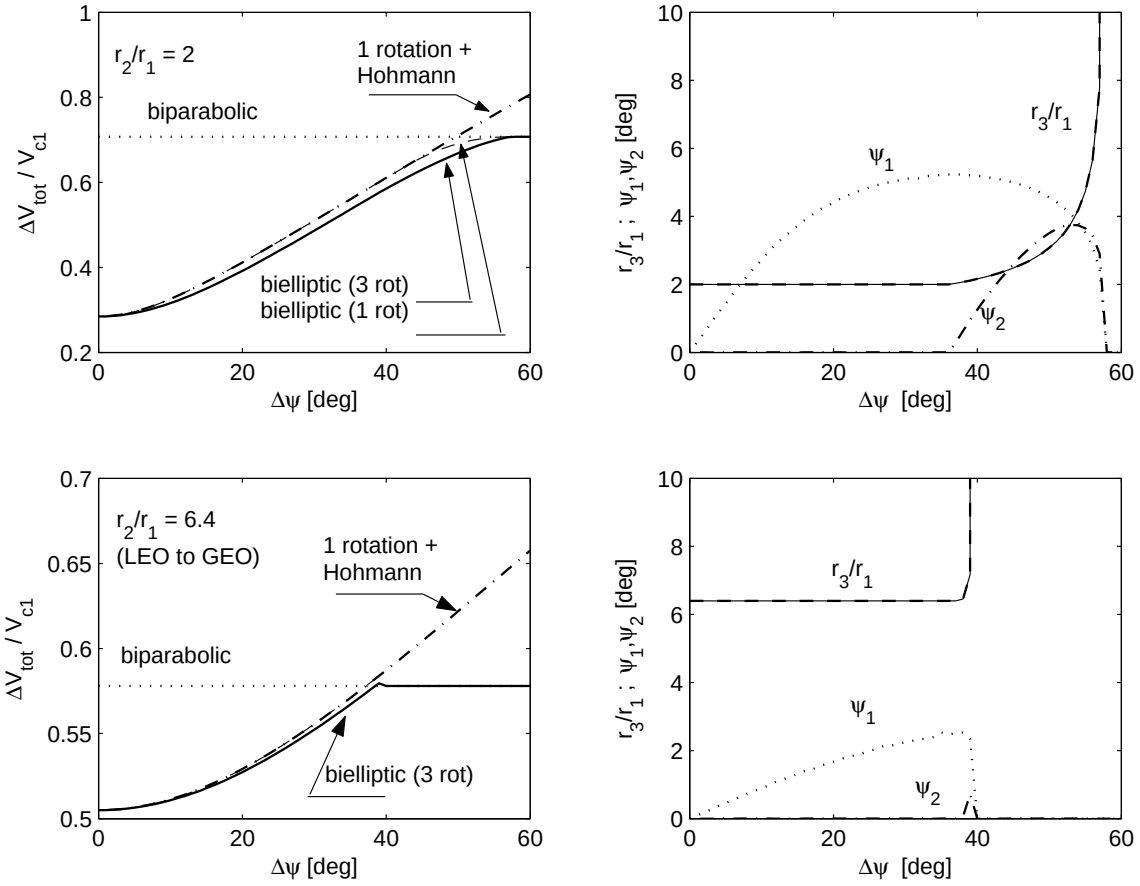

Figure 5.12: **Three–impulse transfer between noncoplanar circular orbits: Δv for different strategies (top) and optimal parameters r_3, ψ_1 and $\psi 2$ (right).**

5.4.3 Three-impulse noncoplanar transfer between circular orbits

Similar concepts also apply to the noncoplanar transfer between circular orbits of different radii. As it appears in Fig. 5.12, the range of optimality of the bielliptic transfer becomes narrower as the radius-ratio r_2/r_1 increases above unity, either if the rotation is concentrated at the maximum distance r_3, or if it is split among the three burns.

In particular, the classical noncoplanar Hohmann transfer is optimal for $\Delta\psi < 38$ deg in the case of the LEO-GEO transfer (plots on the right of Fig. 5.12, for $r_2/r_1 = 6.4$). For $\Delta\psi$ just above this limit, the optimal manoeuvre is biparabolic.

Chapter 6

Lunar trajectories

6.1 The Earth-Moon system

The peculiarity of the Earth–Moon system, a peculiarity that has considerable effects on lunar trajectories, is the relative size of Earth and Moon, whose mass ratio is 81.3, which is far larger than any other binary system in the solar system, the only exception being the Pluto and Caron pair, with a mass ratio close to 7.

The Earth-Moon average distance, that is, the semi-major axis of the geocentric lunar orbit, is 384,400 km. The two bodies actually revolve on elliptic paths about their centre of mass, which is distant 4,671 km from the centre of the planet, i.e., about 3/4 of the Earth radius. The Moon's average barycentric orbital speed is 1.010 km/s, while the Earth's is 0.012 km/s. The total of these speeds gives the geocentric lunar average orbital speed, 1.022 km/s (see Bate *et al.*, cap. 7.2; Cornelisse, p. 343).

The computation of a precision lunar trajectory requires the numerical integration of the equation of motion starting from tentative values for position and velocity at injection time, when the spacecraft leaves a LEO parking orbit to enter a ballistic trajectory aimed at the Moon. Solar perturbations (including radiation), the oblate shape of the Earth, and mainly the terminal attraction of the Moon must be taken into account. Because of the complex motion of the Moon, its position is provided by lunar ephemeris. Approximate analytical methods, which only take the predominant features of the problem into account, are required to narrow down the choice of the launch time and injection conditions. These methods will be developed in the following sections.

6.2 Simple Earth-Moon trajectories

A very simple analysis permits to assess the effect of the injection parameters, namely the radius of the parking orbit r_0, the velocity v_0, and the flight path angle φ_0, on the time–of–flight. The analysis assumes that the lunar orbit is circular with radius $R = 384{,}400$ km, and neglects the terminal attraction of the Moon. The spacecraft trajectory is in the plane of the lunar motion, a condition that actual trajectories approximately fulfill to avoid expensive plane changes.

From the (constant) values of energy and angular momentum along the ballistic trajectory,

$$\mathcal{E} = \frac{v_0^2}{2} - \frac{\mu}{r} \quad ; \qquad h = r_0 v_0 \cos\varphi,$$

the geometric parameters of the ideal Keplerian transfer orbit to the Moon can be determined:

$$a = -\frac{\mu_\oplus}{2\mathcal{E}} \; ; \quad p = \frac{h^2}{\mu_\oplus} \; ; \quad e = \sqrt{1 - \frac{p}{a}}$$

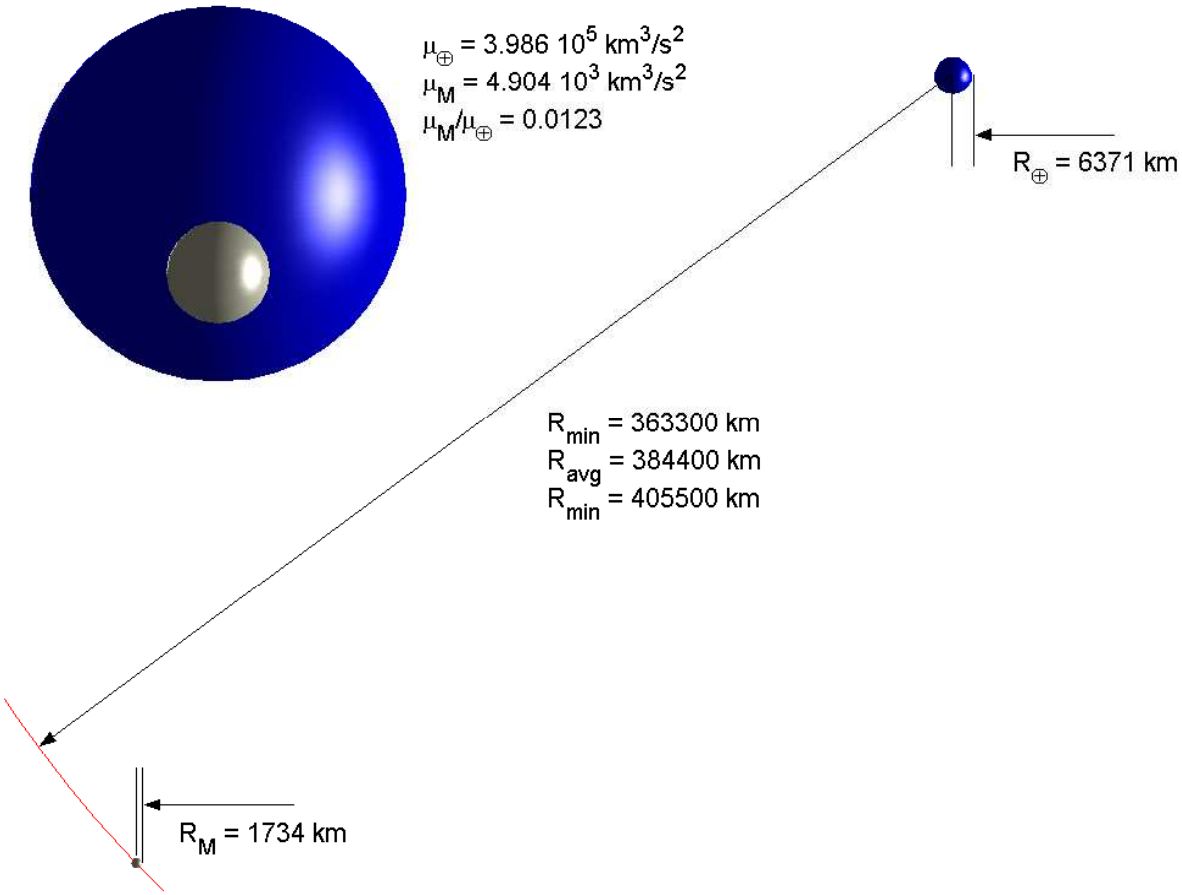

Figure 6.1: **Representation of the Earth–Moon system (radii are in the correct scale, so that the spacecraft flying between them is not visible).**

Solving the polar equation of the conic section, one finds the true anomaly at departure, ν_0, and at the intersection with the lunar orbit, ν_1, that is

$$\cos\nu_0 = \frac{p - r_0}{e r_0} \quad ; \qquad \cos\nu_1 = \frac{p - r_1}{e r_1}$$

where $r_1 = R$ is the Moon orbit average radius.

The time–of–flight is computed using the time equations presented in Section 2.7.

Let ω_M be the orbit angular velocity of the Moon along its orbit and *gamma* the phase angle, *i.e.* the angle between the position vector of the probe and the position vector of the Moon with respect to the Earth. Its value at departure and at arrival, indicated as γ_0 and γ_1, respectively, satisfy the equation

$$\gamma_0 + \omega_M(t_1 - t_0) = (\nu_1 - \nu_0) + \gamma_1.$$

where $\gamma_1 = 0$ for a direct hit (neglecting the final attraction of the Moon). The phase angle at departure,

$$\gamma_0 = (\nu_1 - \nu_0) + \gamma_1 - \omega_M(t_1 - t_0),$$

under to the assumption of circularity for both the lunar and parking orbits, actually determines the times of launch opportunities (the launch windows).

The total propulsive effort is evaluated by adding the theoretical velocity required to attain the parking orbit energy (Section 4.3), $v_{0_{\rm eq}}$, and the magnitude of the velocity increment

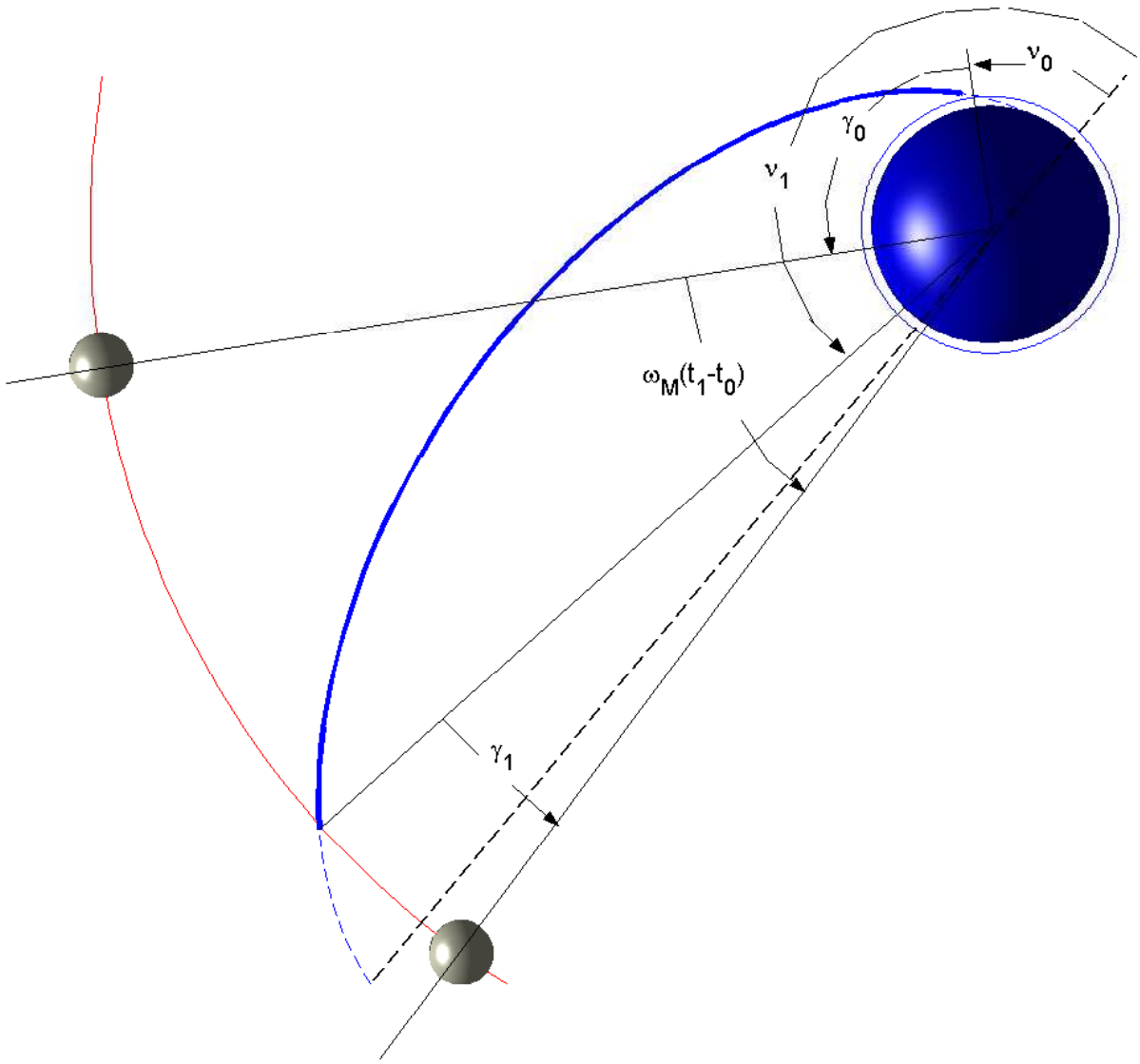

Figure 6.2: **Earth–centered transfer orbit to the Moon: definition of the angles.**

necessary for leaving the circular LEO

$$\Delta\vec{v}_0 = \vec{v}_0 - \vec{v}_{c_0} = v_0 \sin\varphi_0 \hat{\boldsymbol{i}} + (v_0 \cos\varphi_0 - v_{c_0})\hat{\boldsymbol{j}}.$$

The results presented in Fig. 6.3 suggest to depart with an impulse parallel to the circular velocity ($\varphi = 0$) from a parking orbit at the minimum altitude that would permit a sufficient stay, taking into account the decay due to the atmospheric drag.

Other features of the trajectories based on a 320 km altitude LEO are presented in Fig. 6.4 as a function of the injection velocity v_0. The minimum injection velocity of 10.82 km/s originates a Hohmann transfer that requires the maximum flight time of about 120 hour (5 days). The apogee velocity is 0.188 km/s, and the velocity relative to the Moon has the opposite direction, resulting in an impact on the leading edge of our satellite.

A modest increment of the injection velocity significantly reduces the trip time. For the manned Apollo missions, the life-support requirements led to a flight time of about 72 hour, that also avoided the unacceptable non–return risk of hyperbolic trajectories. Further increments of the injection velocity reduce the flight time and the angle $\nu_1 - \nu_0$ swept by the lunar probe

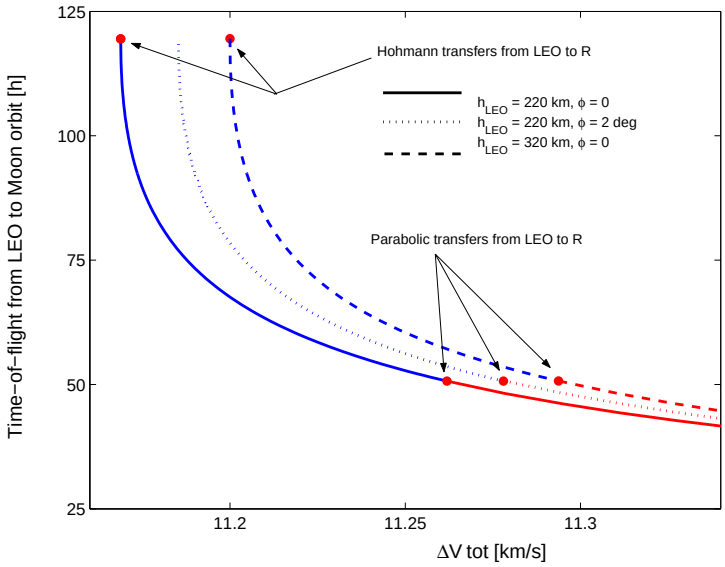

Figure 6.3: **Approximate time-of-flight of lunar trajectories.**

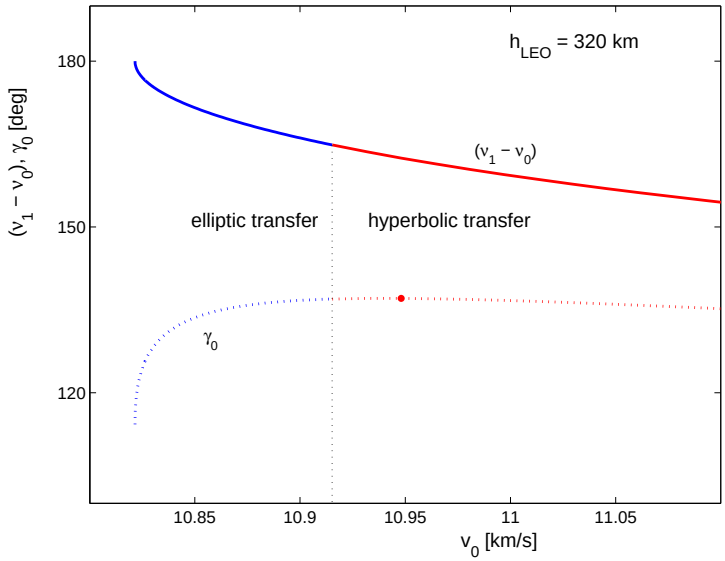

Figure 6.4: **Lunar trajectories departing from 320 km circular LEO with $\varphi_0 = 0$.**

from the injection point to the lunar intercept. In the limiting case of infinite injection speed, the trajectory is a straight line with a zero trip–time, $\nu_1 - \nu_0 = \pi/2$, and impact in the centre of the side facing the Earth.

It is also interesting to note that the phase angle at departure γ_0 presents a stationary point (that is, a maximum, for $v_0 = 10.94$ km/s, in the range of hyperbolic transfer trajectories. In this condition, given a prescribed (*i.e.* fixed) phase angle at arrival on the Moon orbit, γ_1, and assuming a tangential burn to leave LEO, such that $\ni_0 = 0$, it is

$$\frac{\mathrm{d}\gamma_0}{\mathrm{d}v_0} = \frac{\mathrm{d}\gamma_0}{\mathrm{d}t_1}\frac{\mathrm{d}t_1}{\mathrm{d}v_0} = \left(\frac{\mathrm{d}\nu_1}{\mathrm{d}t_1} - \omega_M\right)\frac{\mathrm{d}t_1}{\mathrm{d}v_0} \quad \Longrightarrow \quad \mathrm{d}\nu_1 = \omega_M \mathrm{d}t_1$$

This means that, by choosing this strategy, an error on the initial speed, e.g., a higher speed upon leaving LEO, reduces the swept angle, but the error is almost exactly compensated ($\Delta\nu_1 \approx \omega_M \Delta t_1$) by the reduced time required by the Moon to reach the new intersection point.

The practical significance of this condition is scarce, because the model is oversimplified and hyperbolic trajectories are not an option for the increased fuel necessary and the absence, for

manned missions, of free–return trajectories. However, this kind of consideration are useful to remind the reader that space missions are designed looking not only for reduced Δv, but also for safety from errors.

6.3 The patched-conic approximation

Any prediction of the lunar arrival conditions requires accounting for the terminal attraction of the Moon. A classical approach is bases on the *patched-conic approximation*, which is still based on the analytical solution of the two-body problem. The spacecraft is considered under the only action of the Earth until it enters the Moon's sphere of influence: inside it, Earth's attraction is neglected.

The concept of sphere of influence, introduced by Laplace, is conventional, as the transition from geocentric to selenocentric motion is a gradual process that takes place on a finite arc of the trajectory where both Earth and Moon affect the spacecraft dynamics equally. Nevertheless this approach is an acceptable approximation for a preliminary analysis and mainly for evaluating the injection.[1] Solar perturbation is the main reason that makes the description of the trajectory after the lunar encounter only qualitative.

The previous assumption of circular motion of the Moon in the same plane of the spacecraft trajectory is retained. The probe enters, before the apogee, the lunar sphere of influence whose radius

$$\rho = R\left(\frac{\mu_M}{\mu_\oplus}\right)^{2/5} = 66,200km$$

was assumed according to Laplace's definition.

6.3.1 Geocentric leg

The geocentric phase of the trajectory is specified by four initial conditions:

- r_0, radius of the LEO parking orbit;
- v_0, velocity at injection onto the ballistic transfer orbit;
- φ_0, flightpath angle at injection time t_0;
- γ_0, phase angle with the Moon at injection time t_0 (see Fig. 6.2).

An iterative process allows for the determination of point 1, where the spacecraft enters the lunar sphere of influence. The phase angle γ_1 is conveniently replaced by another independent variable, that is, the angle λ_1 which specifies the position of point 1, where the trajectory crosses the boundary of the lunar sphere of influence, with respect to the position vector of the Moon in geocentric coordinates.

The computation of the geocentric leg is carried out using the equations presented in Section 4.2, with the only remarkable exception of radius and phase angle at end of the geocentric leg, which are obtained by means of elementary geometry (Fig. 6.5):

$$r_1 = \sqrt{\rho^2 + R^2 - 2\rho R\cos\lambda_1} \quad ; \qquad \sin\gamma_1 = \frac{\rho}{r_1}\sin\lambda_1$$

The last equation implies that the angle λ_1 has the same sign as the phase angle, γ_1, which is positive if the Moon is ahead of the spacecraft. In Fig. 6.5 a case with $\lambda_1 < 0$ is represented.

[1] The method of patched–conics works better for interplanetary missions, where the approximations introduced by the concept of sphere of influence have a minor impact on long interplanetary transfers, where the transition to and from the sphere of influence of one planet takes place over a much smaller angular segment of the Heliocentric interplanetary transfer orbit.

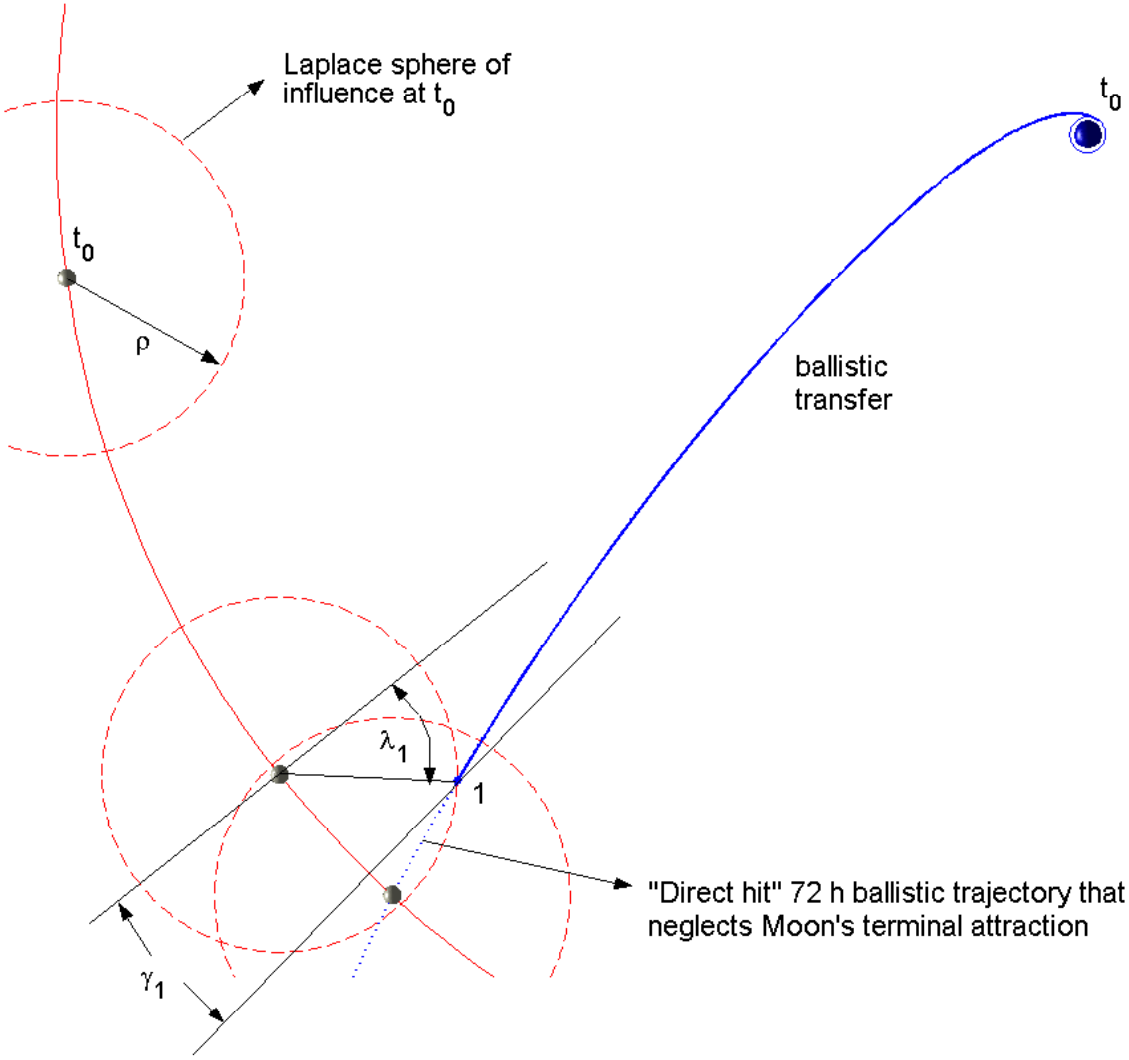

Figure 6.5: **Geocentric leg (transfer to the Moon's sphere of influence).**

The conservation of energy and angular momentum provide the geocentric velocity and flight path angle at point 1,

$$v_1 = \sqrt{2\left(\mathcal{E} + \frac{\mu_\oplus}{r_1}\right)} \quad ; \qquad \cos\varphi_1 = \frac{h}{r_1 v_1}$$

where $\varphi > 0$, since the arrival occurs prior to apogee.

It is worthwhile to note that the energy at injection $\mathcal{E} = v_0^2/2 - \mu_\oplus/(r_\oplus + h_{\text{LEO}})$ must be sufficient to reach point 1. In mathematical terms this meand that the argument of the square root, $2(\mathcal{E} + \mu_\oplus/r_1)$, must be positive. Note also that, in this respect, the energy necessary to enter the sphere of influence of the Moon is less than that required by the Hohmann transfer that takes the space probe to a distance equal to R, since r_1 is strictly less than R.

6.3.2 Selenocentric leg

Spacecraft position and velocity on entering the sphere of influence (time t_1) must be expressed in a non-rotating selenocentric reference frame, in order to compute the trajectory around the Moon. Both vector quantities that describe the initial condition of the selenocentrc leg will be indicated by the subscript 2. Position can be described in polar coordinates by the radius $r_2 = \rho$

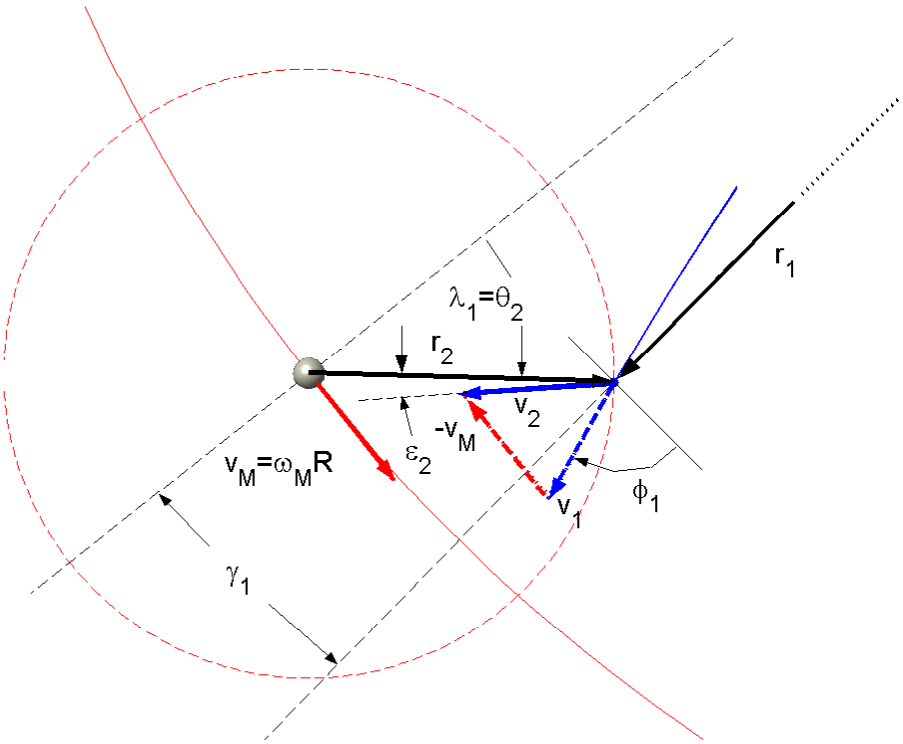

Figure 6.6: **Entry in the Moon's sphere of influence and start of the selenocentric leg.**

and the anomaly $\vartheta_2 = \lambda_1$; the velocity relative to the Moon is

$$\vec{\boldsymbol{v}}_2 = \vec{\boldsymbol{v}}_1 - \vec{\boldsymbol{v}}_M$$

The angle between the vectors $\vec{\boldsymbol{v}}_1$ and $-\vec{\boldsymbol{v}}_M$ in Fig. 6.6 is $\varphi_1 - \gamma_1$, where φ_1 is the flight path angle of the probe in point 1, so that the magnitude of $\vec{\boldsymbol{v}}_2$ is

$$v_2 = \sqrt{v_1^2 + v_M^2 - 2v_1 v_M \cos(\varphi_1 - \gamma_1)}$$

The selenocentric frame can be chosen by taking

$$\hat{\boldsymbol{\imath}}_M = \frac{\vec{\boldsymbol{r}}_2}{r_2} \quad \text{and} \quad \hat{\boldsymbol{k}}_M = \frac{\vec{\boldsymbol{h}}_M}{h_M}$$

where $\vec{\boldsymbol{h}}_M = \vec{\boldsymbol{R}} \times \vec{\boldsymbol{v}}_M$ is the angular momentum of the Moon in its motion around the Earth, and the transverse unit vector $\hat{\boldsymbol{\jmath}}_M = \hat{\boldsymbol{k}}_M \times \hat{\boldsymbol{\imath}}_M$ completes a right–handed triad (Fig. 6.7).

The direction of $\vec{\boldsymbol{v}}_2$ in the selenocentric frame is identified by the angle between $\vec{\boldsymbol{v}}_2$ and the radius vector $\vec{\boldsymbol{r}}_2$, namely ε_2, where a positive angle means a counterclockwise lunar trajectory. In Fig. 6.6 a case with $\varepsilon_2 < 0$ is represented. By taking the dot product $\vec{\boldsymbol{v}}_2 \cdot \hat{\boldsymbol{\jmath}}$, where $\hat{\boldsymbol{\jmath}}$ is the transverse unit vector of the selenocentric frame, one obtains the following relation:

$$\begin{aligned} v_2 \sin\varepsilon_2 &= v_M \cos\lambda_1 - v_1 \sin[\pi/2 + \gamma_1 - \varphi_1 - \lambda_1] \\ &= v_M \cos\lambda_1 - v_1 \cos(\varphi_1 - \gamma_1 + \lambda_1) \end{aligned}$$

The geocentric velocity v_1 is usually quite low (some hundred of m/s) and a considerable portion of the selenocentric velocity v_2 is due to the Moon velocity. In most cases v_2 is greater than the lunar escape velocity at the boundary of the sphere of influence and the spacecraft will approach the Moon along a hyperbola.

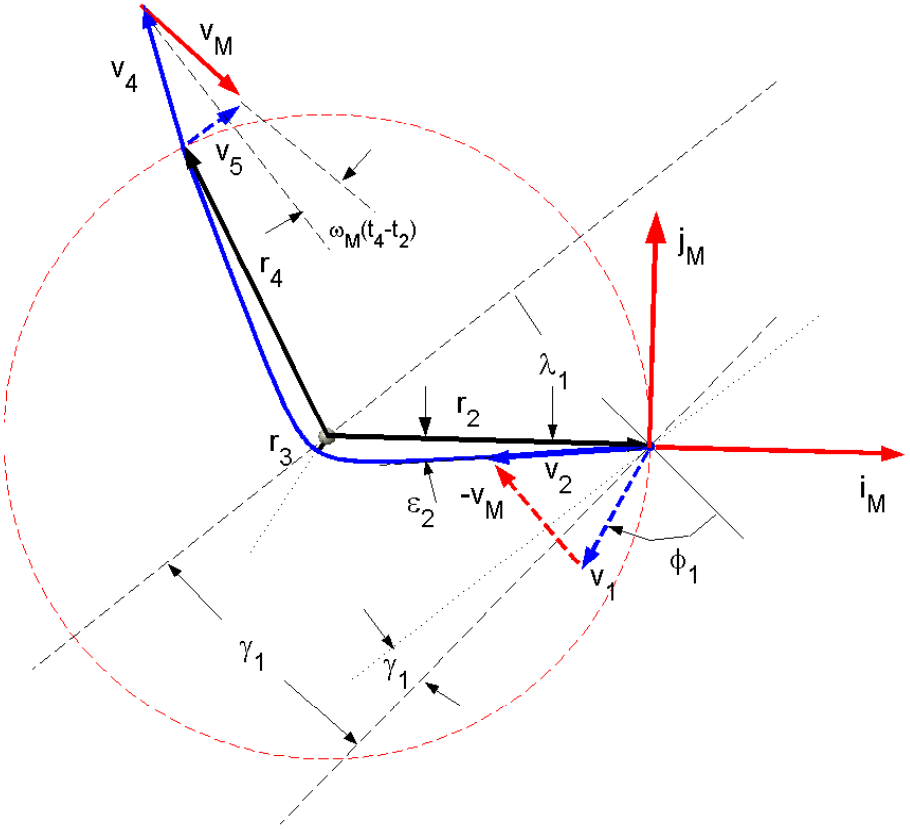

Figure 6.7: **Selenocentric leg.**

The energy and angular momentum of the selenocentric motion can now be computed as

$$\mathcal{E}_M = \frac{v_2^2}{2} - \frac{\mu_M}{r_2} \quad ; \quad h_M = r_2 v_2 \cos\varphi_2 = r_2 v_2 \sin\varphi_2$$

The orbital elements that describe the shape of the lunar trajectory are thus given by

$$a_M = \frac{\mu_M}{2\mathcal{E}_M} \quad ; \quad p_M = \frac{h^2}{\mu_M} \quad ; \quad e_M = \sqrt{\left(1 - \frac{p_M}{a_M}\right)}$$

The true anomaly on entering the sphere of influence

$$\cos\nu_2 = \frac{p_M - r_2}{e_M r_2}$$

is always negative. The radius of the periselenium,

$$r_{P_S} = r_3 = a_M(1 - e_M)$$

plays a crucial role, since there are three possibilities depending mainly on its value:

1. if $r_3 < r_M$ (= 1738 km) the spacecraft hits the Moon;
2. if $r_3 > r_M$ the spacecraft
   (a) can be slowed down by the engines and inserted into a lunar orbit;
   (b) flies–by the Moon and crosses again the sphere of influence (fly–by and free–return).

The first case comprises, together with a destructive impact, a soft landing on the lunar surface without passing through an intermediate parking orbit. A retrorocket system or inflatable cushions are required for landing.

Entering in a lunar orbit require a braking manoeuvre at the periselenium, where energy variations are more efficient. If a semimajor axis a_o is selected for the lunar orbit, it is

$$\Delta v_3 = v_3 - v_{o_3}$$

where the velocities before and after the manoeuvre are given by

$$v_3 = \sqrt{2\left(\mathcal{E}_M + \frac{\mu_M}{r_3}\right)} \quad ; \quad v_{o_3} = \sqrt{\mu_M\left(\frac{2}{r_3} - \frac{1}{a_o}\right)} < \sqrt{2\frac{\mu_M}{r_3}} = v_{\mathrm{esc}_3}$$

where v_{o_3} must be less than the escape velocity at r_3.

The velocity change necessary to perform an insertion onto a lunar orbit is reduced by a low periselenium (minimum gravitational losses) and a high eccentricity orbit with a far apocentre, which however should permit the permanent capture by the Moon. Circular orbits are preferred when a rendezvous is programmed with a vehicle ascending from the lunar surface.

If no action is taken at the periselenium, the spacecraft crosses again the sphere of influence at point 4 with relative velocity $v_4 = v_2$ in the outward direction and $\varepsilon_4 = \pi - \varepsilon_2$. The knowledge of the geocentric velocity

$$\vec{\boldsymbol{v}}_5 = \vec{\boldsymbol{v}}_4 + \vec{\boldsymbol{v}}_M$$

is necessary for the analysis of the successive geocentric path. In what follows, subscript 5 denotes the same exit point on the boundary of the sphere of influence as subscript 4, but refers to quantities measured in a non-rotating, geocentric reference frame.

One should note that the Moon has traveled the angle $\omega_M(t_4 - t_2)$ around the Earth in the time elapsed during the selenocentric phase; its velocity v_M has rotated counterclockwise of the same angle. Alternately, the exit point can be rotated clockwise of the angle $\omega_M(t_4 - t_2)$, if one is interested in keeping fixed the direction of the Earth-Moon line. In this latter case the azimuth of the probe in selenocentric, polar coordinates based on the Earth–Moon line is

$$\vartheta_4 = \lambda_1 \pm 2\nu_2 - \omega_M(t_4 - t_2)$$

where the upper sign holds for $\varepsilon_2 > 0$, *i.e.* counterclockwise lunar trajectories.

A passage in front of the leading edge of the Moon rotates clockwise the relative velocity so that $v_5 < v_1$ and, by assuming $r_5 \approx r_2$, it is also $\mathcal{E}_5 < \mathcal{E}_2$. On the converse, a passage near the trailing edge of the Moon rotates counterclockwise the relative velocity. This means that $v_5 > v_1$ and the geocentric energy $\mathcal{E}_5 > \mathcal{E}_2$ may be sufficient to escape from the Earth gravitation.[2] Only the former trajectory suits a manned mission aimed to enter a lunar orbit, as, in the case of failure of the braking maneuver, it should result into a low-perigee return trajectory.

6.3.3 Noncoplanar lunar trajectories

The preceding analysis was based on a two–dimensional approach, where the probe is assumed to fly always in the plane of the Moon orbit. This is seldom true, as the launch usually puts the spacecraft on an inclined parking orbit in the geocentric–equatorial reference frame and the orbit plane of the Moon is inclined with respect to the base–plane of this reference.

[2]One should remember the a transfer orbit to the Moon has a very high eccentricity, higher than 0.95. This means that the velocity at injection is close to the escape velocity and also a marginal increment can be sufficient to leave the Earth's sphere of influence.

In particular, the orbit of the Moon lies on a plane at an angle of $5°8' \pm 9'$ with respect to the Ecliptic, with the line of the nodes making one complete rotation westward every 18.6 years. The inclination of the equatorial plane with respect to the ecliptic, on the converse, is fairly constant, the precession of the Earth spin axis taking as much as 26,000 years. This means that when the ascending node of the Moon orbit coincides with the direction of the vernal equinox the inclination of the Moon orbit with respect to the equatorial plane is maximum, and it is equal to $5°8' + 23°27' = 28°35'$. When the descending node lies in the direction of the vernal equinox, this inclination is minimum, equal to $23°27' - 5°8' = 18°19'$.

In general, a plane change is always necessary during a Moon mission, which strongly depends on the launch epoch and Lunar ephemeris at epoch. More details can be found in Bate *et al.*, Section 7.5 (p. 344 and following).

6.4 Restricted Circular Three–Body Problem

The circular restricted three-body problem represents one of the long standing problems in orbital dynamics. Originally formulated by Euler in 1772, it represents a particular case of the more general three–body problem, that is, the problem of three point–masses that move under the action of their mutual gravitational action only (Section 1.4).

The restricted circular three body problem concerns the motion of a small mass (→ *restricted*) in the vicinity of two larger primary masses (→ a total of *three–bodies*) which move about each other on circular orbits (→ *circular*), unaffected by the presence of the small one, since that is assumed not to perturb the motion of the two primaries.

In spite of the simplifying assumptions, the three-body problem is very complex and difficult to analyze; the resulting motion of the small mass under the gravitational action of the primary bodies can be chaotic. The restricted problem (both circular and elliptical) was worked on extensively by Lagrange in the 18th century and Poincar at the end of the 19th century.[3] Although it has been shown that the problem cannot be solved analytically (*i.e.* in terms of a closed-form solution of known constants and elementary functions), particular solutions do exist and approximate ones can be calculated by means of series expansion, numerical methods or perturbation methods.

Among the particular solutions, there are five equilibria *Lagrangian points* (or *libration points*), where the small mass remains at rest relative to the two primary bodies, and periodic solutions around some of them.

The circular restricted three-body problem is of importance in space system engineering as it may be used as a dynamical model of the Earth–Moon system, an acceptable approximation as the eccentricity of the Moon orbit is relatively small. Lunar transfer trajectories can be obtained along with free–return trajectories. The Lagrange points also provide ideal locations for space based telescopes and other space science missions.

6.4.1 System Dynamics

First of all it is necessary to define the circular restricted three–body problem, formulating the equations of motion in a properly chosen set of Cartesian coordinates. Let us consider two primary masses m_1 and m_2, in circular orbit around their centre of mass $\mathcal{O}$. We can put a third body of negligible mass m and study its motion with respect to the two primary masses. The main assumption is that, being

$$m \ll m_2 < m_1$$

[3] Poincar's work on the restricted three-body problem was the foundation of deterministic chaos theory.

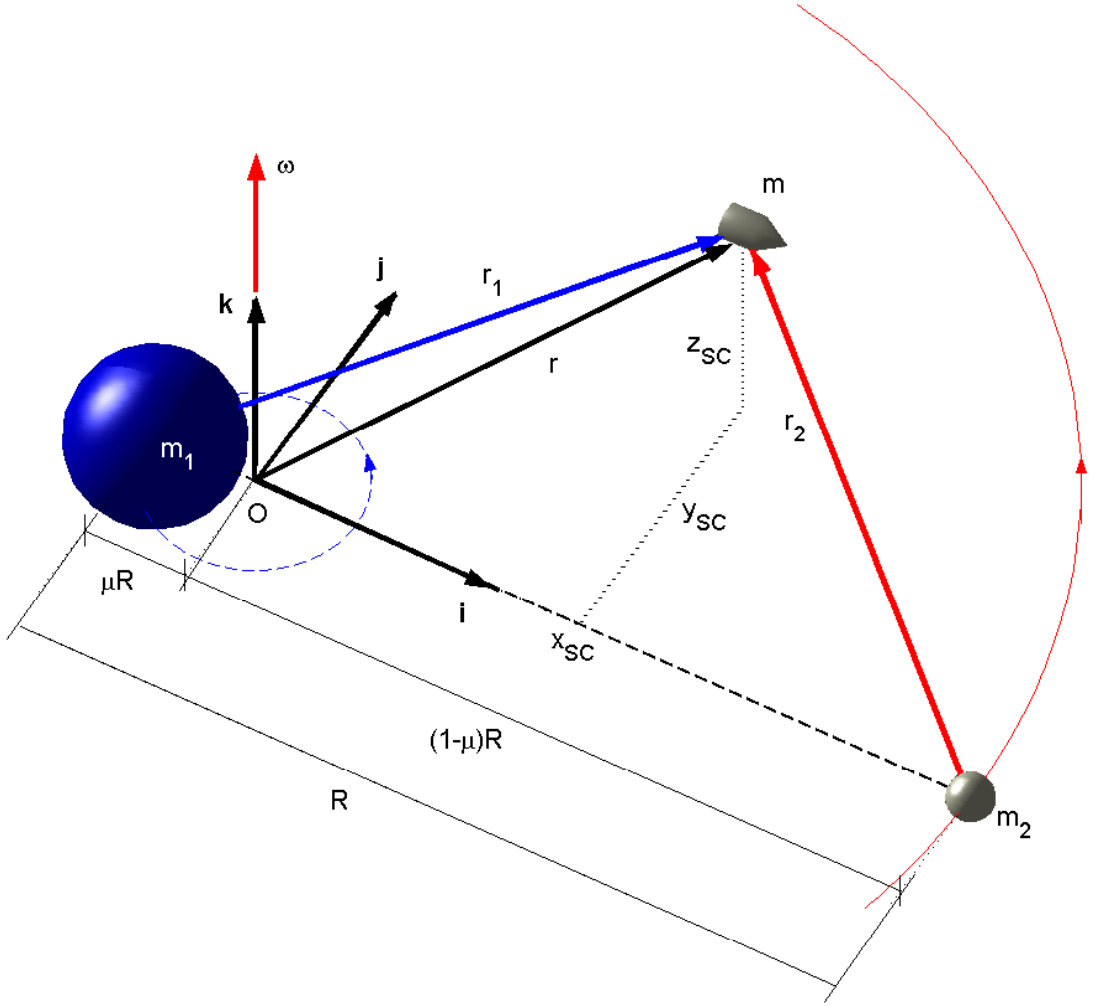

Figure 6.8: **The resticted circular 3–body problem.**

the third body does not influence the motion of the primary masses. This assumption can be used for the motion of any spacecraft in the Earth–Moon system or in the Earth–Sun system. The total mass of the system is

$$M = m_1 + m_2$$

(being m negligible with respect to both the other two). The mass ratio is defined as

$$\mu = \frac{m_2}{M}$$

(it is $\mu = 0.01215$ for the Earth–Moon system, and only $3.0359\ 10^{-6}$ for the Earth–Sun system). Then, it is

$$\begin{aligned} m_2 &= \mu M \\ m_1 &= (1-\mu)M \end{aligned}$$

The position of m_2 with respect to m_1 is identified by the vector $\vec{\boldsymbol{R}}$. It is possible to define a set of Cartesian coordinates, centered in the centre of mass $\mathcal{O}$, where the x–axis lies along $\vec{\boldsymbol{R}}$, the z–axis is perpendicular to the orbit plane, parallel to the angular velocity vector, while the y–axis is perpendicular to $\vec{\boldsymbol{R}}$ and lies in the orbit plane (Fig. 6.8). For the circular restricted three–body problem, the positions of the two primary bodies on the x–axis are fixed, m_1 being

at $x = -\mu R$ and m_2 at $x = (1-\mu)R$, and the reference frame rotates with a constant angular velocity

$$\omega = \sqrt{\frac{GM}{R^3}}$$

The equation of motion of a small mass m in this rotating frame is writtem in vector form as

$$\underbrace{\ddot{\vec{r}}}_{\text{relative}} + \underbrace{\vec{\omega} \times (\vec{\omega} \times \vec{r})}_{\text{centripetal}} + \underbrace{2\vec{\omega} \times \dot{\vec{r}}}_{\text{Coriolis}} = \frac{1}{m}\left(\vec{F}_1 + \vec{F}_1\right)$$

where

$$\begin{aligned} \vec{\omega} &= \omega \hat{\boldsymbol{k}} \\ \vec{r} &= x\hat{\boldsymbol{i}} + y\hat{\boldsymbol{j}} + z\hat{\boldsymbol{k}} \\ \dot{\vec{r}} &= \dot{x}\hat{\boldsymbol{i}} + \dot{y}\hat{\boldsymbol{j}} + \dot{z}\hat{\boldsymbol{k}} \end{aligned}$$

Then

$$\begin{aligned} 2\vec{\omega} \times \dot{\vec{r}} &= -2\omega\dot{y}\hat{\boldsymbol{i}} + 2\omega\dot{x}\hat{\boldsymbol{j}} \\ \vec{\omega} \times (\vec{\omega} \times \vec{r}) &= -\omega^2 x\hat{\boldsymbol{i}} - \omega^2 y\hat{\boldsymbol{j}} \end{aligned}$$

The gravitational acceleration on m due to the action of the two primary bodies is

$$\frac{1}{m}\left(\vec{F}_1 + \vec{F}_1\right) = -GM\left(\frac{(1-\mu)\vec{r}_1}{r_1^3} + \frac{\mu\vec{r}_2}{r_2^3}\right)$$

where

$$\begin{aligned} \vec{r}_1 &= (x+\mu R)\hat{\boldsymbol{i}} + y\hat{\boldsymbol{j}} + z\hat{\boldsymbol{k}} \\ \vec{r}_2 &= [x-(1-\mu)R]\hat{\boldsymbol{i}} + y\hat{\boldsymbol{j}} + z\hat{\boldsymbol{k}} \end{aligned}$$

are the position vectors of m relative to m_1 and m_2. Collecting all the terms, one obtains

$$\begin{aligned} \ddot{x} - 2\omega\dot{y} - \omega^2 x &= -G(1-\mu)M\frac{x+\mu R}{r_1^3} - G\mu M\frac{x+(1-\mu)R}{r_2^3} \\ \ddot{y} + 2\omega\dot{x} - \omega^2 y &= -G(1-\mu)M\frac{y}{r_1^3} - G\mu M\frac{y}{r_2^3} \\ \ddot{z} &= -G(1-\mu)M\frac{z}{r_1^3} - G\mu M\frac{z}{r_2^3} \end{aligned} \tag{6.1}$$

These are the equations of motion for m in the rotating frame of reference.

It is possible to recast Eqs. (6.1) in nondimansional form, by defining the dimensionless position vector

$$\vec{\rho} = (\xi, \eta, \zeta)^T = \vec{r}/R \;\text{ with }\; \xi = x/R\,; \quad \eta = y/R\,; \quad \zeta = z/R$$

and the nondimensional time,

$$\tau = \omega t$$

such that

$$(\cdot)^{\cdot} = \frac{\mathrm{d}(\cdot)}{\mathrm{d}t} = \omega\frac{\mathrm{d}(\cdot)}{\mathrm{d}\tau} = \omega(\cdot)'$$

By substituting these definitions in Eqs. (6.1) and simplfying, one gets

$$\begin{aligned} \xi'' - 2\eta' - \xi &= -(1-\mu)\frac{\xi+\mu}{\rho_1^3} - \mu\frac{\xi(1-\mu)}{\rho_2^3} \\ \eta'' + 2\xi' - \eta &= -(1-\mu)\frac{\eta}{\rho_1^3} - \mu\frac{\eta}{\rho_2^3} \\ \zeta'' &= -(1-\mu)\frac{\zeta}{\rho_1^3} - \mu\frac{\zeta}{\rho_2^3} \end{aligned} \tag{6.2}$$

where

$$\rho_1 = \sqrt{(\xi+\mu)^2 + \eta^2 + \zeta^2}\ ; \quad \rho_2 = \sqrt{[\xi-(1-\mu)]^2 + \eta^2 + \zeta^2}$$

There are no closed form solutions of these equations, in the form

$$\xi = f_1(t)\ ; \quad \eta = f_3(t)\ ; \quad \zeta = f_3(t)$$

although it is possible to demonstrate that there are 5 equilibrium solutions, the *Lagrangian points*,the stability of which can be easily determined. Moreover, the system of ordinary differential equations (6.2) can be numerically integrated, in order to get some insight on the dynamic behaviour of the third body. Numerical investigation shows that there exist also periodic solutions, but the majority of the evolutions are chaotic, in the sense that there is a very strong dependence on the initial conditions.

6.4.2 Jacobi Integral

It is possible to define a three–body potential, from which acceleration terms may be derived,

$$\mathcal{U} = \frac{1}{2}(\xi^2 + \eta^2) + \frac{1-\mu}{\rho_1} + \frac{\mu}{\rho_2}$$

where the first term is relative to the centripetal acceleration, while the second and the third determine the gravity acceleration induced by the two primary masses.

The partial derivative of $\mathcal{U}$ w.r.t. ξ is given by

$$\frac{\partial \mathcal{U}}{\partial \xi} = \xi - \frac{1-\mu}{\rho_1^2}\frac{\partial \rho_1}{\partial \xi} - \frac{\mu}{\rho_2^2}\frac{\partial \rho_2}{\partial \xi}$$

From the definition of ρ_1 it is

$$\rho_1^2 = (\xi+\mu)^2 + \eta^2 + \zeta^2$$

so that

$$2\rho_1\frac{\partial \rho_1}{\partial \xi} = 2(\xi+\mu)$$

and

$$\frac{\partial \rho_1}{\partial \xi} = \frac{\xi+\mu}{\rho_1}$$

Analogously, for ρ_2, one obtains

$$2\rho_2\frac{\partial \rho_2}{\partial \xi} = 2[\xi-(1-\mu)]$$

that is

$$\frac{\partial \rho_2}{\partial \xi} = \frac{\xi-(1-\mu)}{\rho_2}$$

Collecting all the terms,

$$\frac{\partial \mathcal{U}}{\partial \xi} = \xi - \frac{1-\mu}{\rho_1^2}\frac{\xi+\mu}{\rho_1} - \frac{\mu}{\rho_2^2}\frac{\xi-(1-\mu)}{\rho_2}$$

Using the same procedure for $\partial\mathcal{U}/\partial\eta$ and $\partial\mathcal{U}/\partial\zeta$, it is

$$\frac{\partial \mathcal{U}}{\partial \eta} = \eta - \frac{1-\mu}{\rho_1^2}\frac{\partial \rho_1}{\partial \eta} - \frac{\mu}{\rho_2^2}\frac{\partial \rho_2}{\partial \eta}$$

and

$$\frac{\partial \mathcal{U}}{\partial \zeta} = -\frac{1-\mu}{\rho_1^2}\frac{\partial \rho_1}{\partial \zeta} - \frac{\mu}{\rho_2^2}\frac{\partial \rho_2}{\partial \zeta}$$

with

$$\begin{aligned} \frac{\partial \rho_1}{\partial \eta} &= \frac{\eta}{\rho_1} \quad ; \quad \frac{\partial \rho_2}{\partial \eta} = \frac{\eta}{\rho_2} \; ; \\ \frac{\partial \rho_1}{\partial \zeta} &= \frac{\zeta}{\rho_1} \quad ; \quad \frac{\partial \rho_2}{\partial \zeta} = \frac{\zeta}{\rho_2} \end{aligned}$$

Finally, it is

$$\begin{aligned} \frac{\partial \mathcal{U}}{\partial \eta} &= \eta - \frac{1-\mu}{\rho_1^2}\frac{\eta}{\rho_1} - \frac{\mu}{\rho_2^2}\frac{\eta}{\rho_2} \\ \frac{\partial \mathcal{U}}{\partial \zeta} &= -\frac{1-\mu}{\rho_1^2}\frac{\zeta}{\rho_1} - \frac{\mu}{\rho_2^2}\frac{\zeta}{\rho_2} \end{aligned}$$

Thus Eq. (6.2) can be rewritten in terms of gradient of the potential $\mathcal{U}$ as

$$\begin{aligned} \xi'' - 2\eta' &= \frac{\partial \mathcal{U}}{\partial \xi} \\ \eta'' + 2\xi' &= \frac{\partial \mathcal{U}}{\partial \eta} \\ \zeta'' &= \frac{\partial \mathcal{U}}{\partial \zeta} \end{aligned} \tag{6.3}$$

Multiplying the first equation by ξ', the second by η' and the third by ζ', and summing up, one gets

$$\xi'\xi'' + \eta'\eta'' + \zeta'\zeta'' = \xi'\partial\mathcal{U}\xi + \eta'\partial\mathcal{U}\eta + \zeta'\partial\mathcal{U}\zeta$$

which can be rewritten as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\tau}(\xi'^2 + \eta'^2 + \zeta'^2) = \frac{\mathrm{d}\mathcal{U}}{\mathrm{d}\tau}$$

Letting $V = \sqrt{\xi'^2 + \eta'^2 + \zeta'^2}$ be the magnitude of the (non dimensional) relative velocity of the third body, the latter equation can be easily integrated, giving

$$\xi'^2 + \eta'^2 + \zeta'^2 = V^2 = 2\mathcal{U} - C$$

where the constant of integration C is the *Jacobi's constant*. The quantity

$$V^2 = 2\mathcal{U} - C$$

is the *Jacobi's integral*, that can be used to bound the motion of the mass.

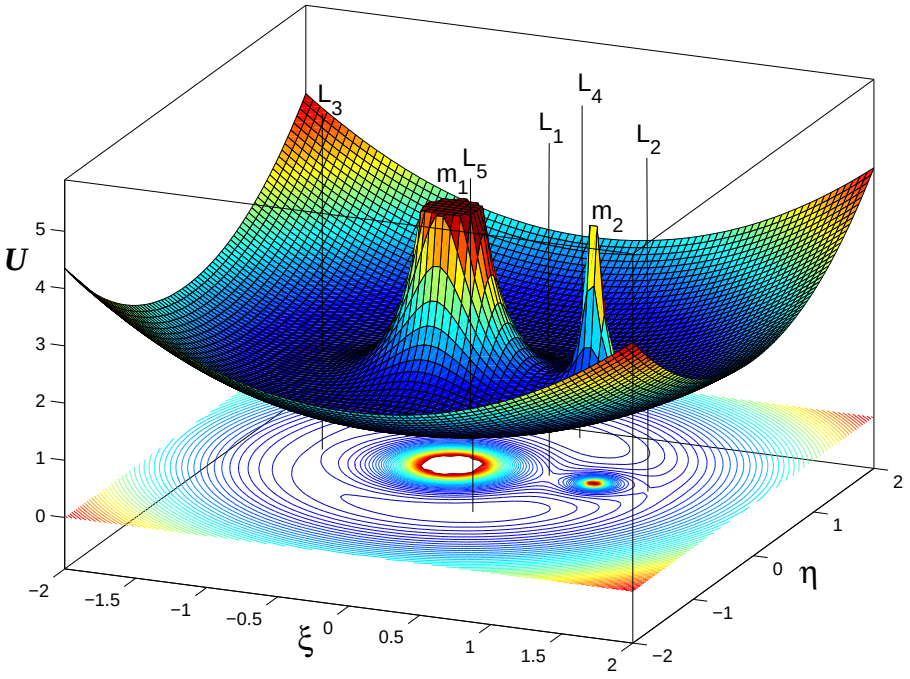

Figure 6.9: **Three–body potential function in the $\xi - \eta$ plane.**

6.4.3 Jacobi Integral in terms of absolute velocity

The absolute velocity $\vec{\boldsymbol{v}}_a$ in a non-rotating frame (with the same origin, such that $\vec{\boldsymbol{r}} = \vec{\boldsymbol{r}}_a$) is expressed as a function of the relative velocity $\vec{\boldsymbol{v}}$ in the rotating one $\{\mathcal{O}; \hat{\boldsymbol{i}}, \hat{\boldsymbol{j}}, \hat{\boldsymbol{k}}\}$ as

$$\vec{\boldsymbol{v}}_a = \vec{\boldsymbol{v}} + \vec{\boldsymbol{\omega}} \times \vec{\boldsymbol{r}}$$

In non dimensional terms, by defining $\vec{\boldsymbol{w}} = \vec{\boldsymbol{v}}_a/(\omega R)$, it is

$$\begin{aligned} \dot{\vec{\boldsymbol{\rho}}} = \vec{\boldsymbol{w}} - \hat{\boldsymbol{k}} \times \vec{\boldsymbol{\rho}} \implies V^2 &= W^2 - 2\vec{\boldsymbol{w}} \cdot (\hat{\boldsymbol{k}} \times \vec{\boldsymbol{\rho}}) + ||\boldsymbol{k} \times \vec{\boldsymbol{\rho}}||^2 \\ &= W^2 - 2\hat{\boldsymbol{k}} \cdot (\vec{\boldsymbol{\rho}} \times \vec{\boldsymbol{w}}) + (\xi + \eta)^2 \\ &= W^2 - 2\hat{\boldsymbol{k}} \cdot \vec{\boldsymbol{h}} + (\xi + \eta)^2 \;=\; W^2 - 2h_3 + (\xi + \eta)^2 \end{aligned}$$

where $\vec{\boldsymbol{h}} = \vec{\boldsymbol{\rho}} \times \vec{\boldsymbol{w}}$ is the (nondimensional) angular momentum vector of the third body. It must be remembered that the total angular momentum of the system is constant. Since the primary bodies have an absolute motion around the origin $\mathcal{O}$ of the system on a fixed plane with constant angular velocity, the component of $\vec{\boldsymbol{h}}$ along the vertical, h_3 must also remain constant. Substituting the definitions of the Jacobi integral, $V^2 = 2\mathcal{U} - C$ and of the Jacobi potential, $\mathcal{U} = (1/2)(\xi^2 + \eta^2) + (1-\mu)/\rho_1 + \mu/\rho_2$, in the last equation, one gets

$$\frac{1-\mu}{\rho_1} + \frac{\mu}{\rho_2} - \frac{C}{2} = \frac{W^2}{2} - h_3$$

that is

$$\frac{W^2}{2} - \frac{1-\mu}{\rho_1} - \frac{\mu}{\rho_2} - h_3 = \mathcal{E} - h_3 = -\frac{C}{2}$$

As a consequence, Jacobi's integral can be also expressed in a non–rotating frame as a combination of the total specific energy and the vertical component of the angular momentum of the third body, both constant in an inertial frame of reference.

6.4.4 Lagrangian libration points

Given two massive bodies in circular orbits around their common centre of mass, it can be shown that there are five positions in space, the *Lagrangian libration points*, where a third body of comparatively negligible mass would maintain its position relative to the two massive bodies. As seen in the frame which rotates with the same period as the two co-orbiting bodies, the gravitational fields of two massive bodies combined with the centrifugal force are in balance at the Lagrangian points.

For a static equilibrium, it is necessary that the following equations hold,

$$\xi' = \xi'' = \eta' = \eta'' = \zeta' = \zeta'' = 0$$

Substituting these conditions in Eq. (6.3), one obtains for the equilibrium condition the following expression:

$$\frac{\partial \mathcal{U}}{\partial \xi} = \frac{\partial \mathcal{U}}{\partial \eta} = \frac{\partial \mathcal{U}}{\partial \zeta} = 0$$

After recalling the expressions of the components of the gradient of the potential $\mathcal{U}$,

$$\begin{aligned}
\frac{\partial \mathcal{U}}{\partial \xi} &= \xi - (1-\mu)\frac{\xi+\mu}{\rho_1^3} - \mu\frac{\xi-(1-\mu)}{\rho_2^3} \\
\frac{\partial \mathcal{U}}{\partial \eta} &= \eta - (1-\mu)\frac{\eta}{\rho_1^3} - \mu\frac{\eta}{\rho_2^3} \\
\frac{\partial \mathcal{U}}{\partial \zeta} &= -(1-\mu)\frac{\zeta}{\rho_1^3} - \mu\frac{\zeta}{\rho_2^3}
\end{aligned}$$

a first observation is that, from the third equation, in order to have $\partial\mathcal{U}/\partial\zeta = 0$ one must have $\zeta = 0$, that is, **all the equilibrium points lie in the $\xi - \eta$ plane**.

collinear points

The third equilibrium condition require that $\zeta = 0$. If it is also $\eta = 0$ one gets that $\partial\mathcal{U}/\partial\eta = 0$ as well, so that only the first equation, $\partial\mathcal{U}/\partial\xi = 0$, is necessary for determining equilibrium points along the ξ axis. With $\eta = \zeta = 0$ one gets

$$\xi - (1-\mu)\frac{\xi+\mu}{\rho_1^3} - \mu\frac{\xi-(1-\mu)}{\rho_2^3} = 0$$

with $\rho_1 = |\xi + \mu|$ and $\rho_2 = |1 - \mu - \xi|$.

Plotting the graphs of $\mathcal{U}$ and $d\mathcal{U}/d\xi$ it is possible to determine graphically the position of the collinear lagrangian points (Fig. 6.10). It is also possible to determine them numerically, with greater accuracy. From the previous equation, three cases are obtained as a function of the sign of $\xi + \mu$ and $1 - \mu - \xi$:

$$\begin{aligned}
-\mu < \xi < 1-\mu \;\Rightarrow\; \rho_1 = \xi+\mu \;;\;\; \rho_2 = 1-\mu-\xi \;\Rightarrow\; \rho_1+\rho_2 = 1 \\
\xi > 1-\mu \;\Rightarrow\; \rho_1 = \xi+\mu \;;\;\; \rho_2 = \mu+\xi-1 \;\Rightarrow\; \rho_1-\rho_2 = 1 \\
\xi < -\mu \;\Rightarrow\; \rho_1 = -\xi-\mu \;;\;\; \rho_2 = 1-\mu-\xi \;\Rightarrow\; \rho_2-\rho_1 = 1
\end{aligned}$$

Case 1: $-\mu < \xi < 1 - \mu$

In this case, the equilibrium condition can be rewritten as

$$1 - \mu - \rho_2 - \frac{1-\mu}{\rho_1^2} + \frac{\mu}{\rho_2^2} = 0$$

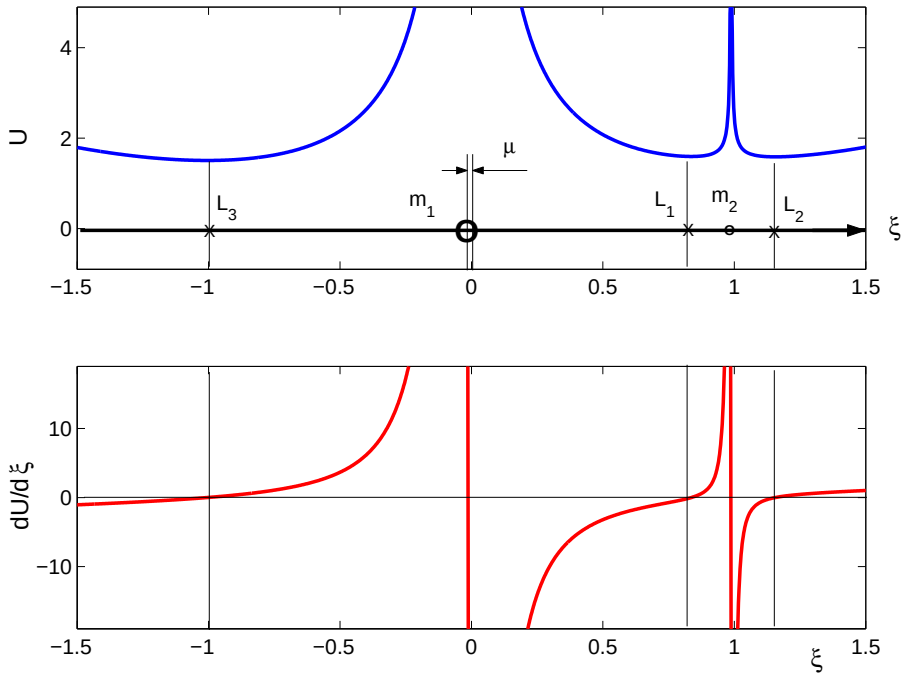

Figure 6.10: **Position of the collinear points along the ξ axis.**

Multiplying by $\rho_1^2\rho_2^2$ and setting the numerator of the resulting fraction to 0, one gets

$$(1-\mu-\rho_2)\rho_1^2\rho_2^2-(1-\mu)\rho_2^2+\mu\rho_1^2=0$$

Substituting $\rho_1 = 1-\rho_2$ yields the fifth order polynomial

$$p_1(\rho_2)=\rho_2^5-(3-\mu)\rho_2^4+(3-2\mu)\rho_2^3-\mu\rho_2^2+2\mu\rho_2-\mu=0$$

When m_2 is significantly less than m_1, as in the Earth–Moon case, the only real solution for this polynomial is approximately given by $\rho_2 \approx \sqrt[3]{\mu/3}$. In the Earth–Moon case this point is known as the *cislunar point*.

Case 2: $\xi > 1-\mu$

Similarly, in Case 2, the equilibrium condition becomes

$$1-\mu-\rho_2-\frac{1-\mu}{\rho_1^2}-\frac{\mu}{\rho_2^2}=0$$

that, multiplied by $\rho_1^2\rho_2^2$ and setting the numerator of the resulting fraction to 0, becomes

$$(1-\mu-\rho_2)\rho_1^2\rho_2^2-(1-\mu)\rho_2^2+\mu\rho_1^2=0$$

Substituting $\rho_1 = \rho_2+1$ in the above expression and rearranging, one gets the fifth order polynomial

$$p_2(\rho_2)=\rho_2^5+(3-\mu)\rho_2^4+(3-2\mu)\rho_2^3-\mu\rho_2^2-2\mu\rho_2-\mu=0$$

When m_2 is significantly less than m_1, the only real solution for this polynomial is approximately given again by $\rho_2 \approx \sqrt[3]{\mu/3}$, but this point is on the other side, w.r.t. to the greater of the two primary masses, where $\xi = 1-\mu+\rho_2$. In the Earth–Moon case this point is called the *translunar point*.

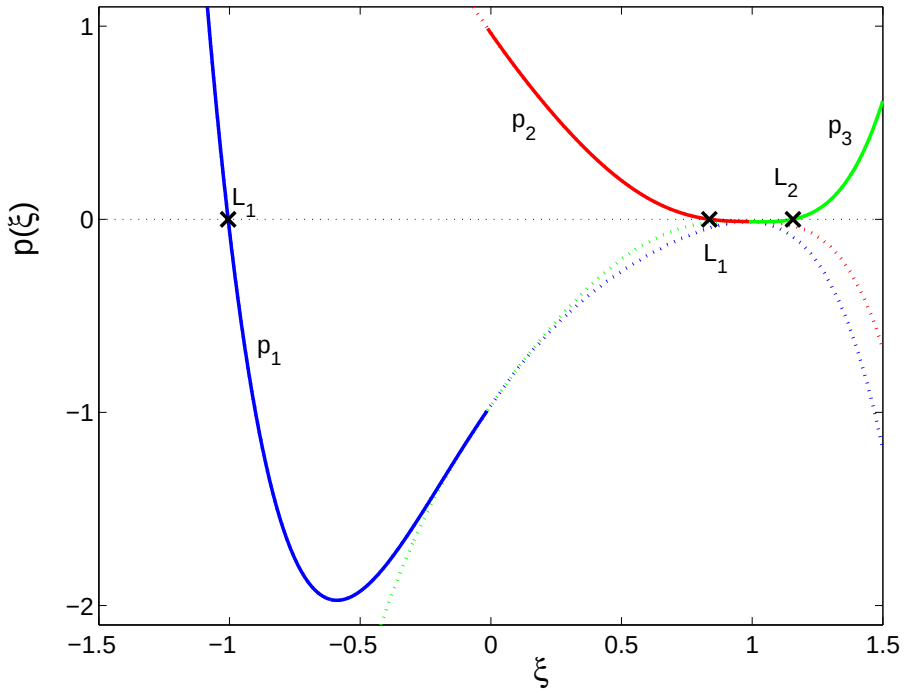

Figure 6.11: **Position of the collinear points along the ξ axis.**

Case 3: $\xi < -\mu$

Finally, in the third case, the equilibrium condition can be rewritten as

$$1 - \mu - \rho_2 + \frac{1-\mu}{\rho_1^2} + \frac{\mu}{\rho_2^2} = 0$$

Multiplying by $\rho_1^2\rho_2^2$ and setting the numerator of the resulting fraction to 0, one gets

$$(1 - \mu - \rho_2)\rho_1^2\rho_2^2 + (1-\mu)\rho_2^2 + \mu\rho_1^2 = 0$$

In this latter case it is $\rho_1 = 1 + \rho_2$ and the fifth order equation is

$$p_3(\rho_2) = \rho_2^5 - (3-\mu)\rho_2^4 + (3-2\mu)\rho_2^3 - (2-\mu)\rho_2^2 + 2\mu\rho_2 - \mu = 0$$

This equation has one real solution for $\rho_2 \approx 2$, if $m_2 \ll m_1$, that is $\rho_1 \approx 1$. This point lies along the ξ axis, approximately in the symmetric position of m_1 w.r.t. the system origin. In the Earth–Moon case the third lagrangian point is called the *trans–Earth point*.

Figure 6.11 shows the plots of the three polynomials $p_i(\xi)$, for $i = 1, 2, 3$ and the position of their real roots in the three intervals of the ξ axis, that is the position of the collinear lagrangian points.

Equilateral points

From the first two equations, we see that if $\rho_1 = \rho_2 = 1$, it is

$$\begin{aligned} \frac{\partial \mathcal{U}}{\partial \xi} &= \xi - (1-\mu)(\xi+\mu) - \mu\xi - \mu(1-\mu) = 0 \\ \frac{\partial \mathcal{U}}{\partial \eta} &= \eta - (1-\mu)\eta - \mu\eta = 0 \end{aligned}$$

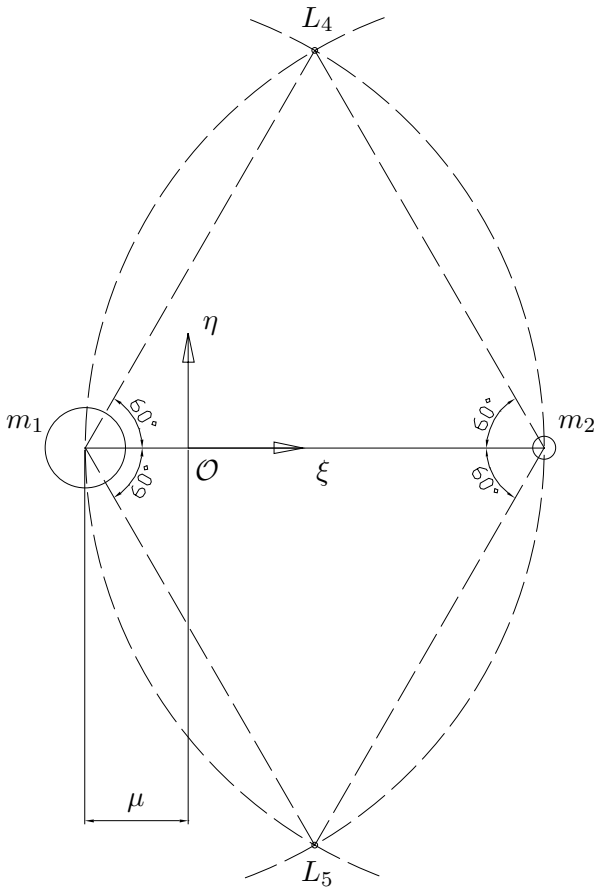

Figure 6.12: **Position of equilateral lagrangian points in the $\xi - \eta$ plane.**

This means that at the intersections of the circumferences of radius 1 (i.e. radius R, in dimensional terms) we have two equilibria, L_4 and L_5, which are the vertices, with m_1 and m_2 of two equilateral triangles. In nondimensional coordinates, the position of L_4 and L_5 is given by

$$L_4 = (1/2 - \mu; \sqrt{3}/2; 0) \; ; \quad L_5 = (1/2 - \mu; -\sqrt{3}/2; 0)$$

In the Lagrangian points L_4 and L_5 of the Sun–Jupiter system it is possible to observe the so called *Trojan Asteroids*. More details are reported at the end of this section.

6.4.5 Lagrange Point stability

By linear perturbation analysis it is possible to evaluate the stability of the five lagrangian points. Letting $L_i = (\xi_i; \eta_i)$ denote the position of the i–th libration points, it is possible to determine the linearized expression of the equation of motion written in the form

$$\begin{aligned} \xi'' - 2\eta' &= \frac{\partial \mathcal{U}}{\partial \xi} \\ \eta'' + 2\xi' &= \frac{\partial \mathcal{U}}{\partial \eta} \\ \zeta'' &= \frac{\partial \mathcal{U}}{\partial \zeta} \end{aligned}$$

where we recall that

$$\mathcal{U} = \frac{1}{2}(\xi^2 + \eta^2) + \frac{1-\mu}{\rho_1} + \frac{\mu}{\rho_2}$$

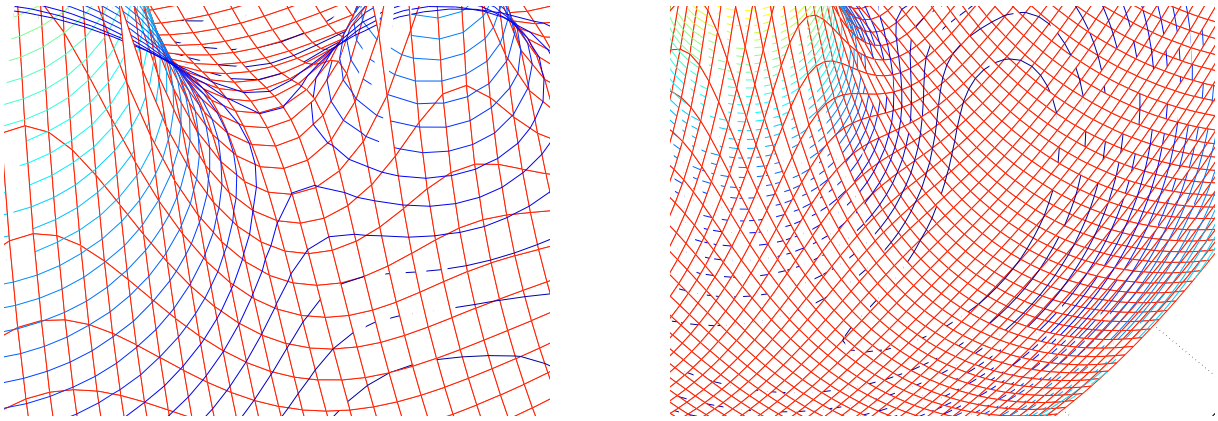

Figure 6.13: **Saddle point and local minimum.**

Substituting in the equation of motion the perturbed values of the position vector components,

$$\xi = \xi_i + \varepsilon_\xi \; ; \quad \eta = \eta_i + \varepsilon_\eta \; ; \quad \zeta = 0 + \varepsilon_\zeta$$

and evaluating the Jacobian matrix of the potential

$$\begin{aligned} \varepsilon''_\xi - 2\varepsilon'_\eta &= \frac{\partial^2 \mathcal{U}}{\partial \xi^2}\varepsilon_\xi + \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \eta}\varepsilon_\eta + \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \zeta}\varepsilon_\zeta \\ \varepsilon''_\eta + 2\varepsilon'_\xi &= \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \eta}\varepsilon_\xi + \frac{\partial^2 \mathcal{U}}{\partial \eta^2}\varepsilon_\eta + \frac{\partial^2 \mathcal{U}}{\partial \eta \partial \zeta}\varepsilon_\zeta \\ \varepsilon''_\zeta &= \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \zeta}\varepsilon_\xi + \frac{\partial^2 \mathcal{U}}{\partial \eta \partial \zeta}\varepsilon_\eta + \frac{\partial^2 \mathcal{U}}{\partial \zeta^2}\varepsilon_\zeta \end{aligned}$$

The eigenvalues of the symmetric jacobian matrix

$$\nabla^2_\rho \mathcal{U} = \begin{bmatrix} \frac{\partial^2 \mathcal{U}}{\partial \xi^2}\varepsilon_\xi & \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \eta}\varepsilon_\eta & \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \zeta}\varepsilon_\zeta \\ \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \eta}\varepsilon_\xi & \frac{\partial^2 \mathcal{U}}{\partial \eta^2}\varepsilon_\eta & \frac{\partial^2 \mathcal{U}}{\partial \eta \partial \zeta}\varepsilon_\zeta \\ \frac{\partial^2 \mathcal{U}}{\partial \xi \partial \zeta}\varepsilon_\xi & \frac{\partial^2 \mathcal{U}}{\partial \eta \partial \zeta}\varepsilon_\eta & \frac{\partial^2 \mathcal{U}}{\partial \zeta^2}\varepsilon_\zeta \end{bmatrix}$$

provide the required information on the stability of the lagrangian points. It can be demonstrated that equilateral points can be local minima for the potential $\mathcal{U}$, i.e. be stable equilibria, if

$$\frac{\mu}{1-\mu} + \frac{1-\mu}{\mu} \geq 25 \quad \Longrightarrow \quad \mu < 0.03852$$

a condition that is satisfied by both the Sun–Jupiter and Earth–Moon systems.

On the other side, for every value of the mass parameter μ the quadratic form $\nabla^2_\rho \mathcal{U}$ for $i = 1$, 2, and 3 (collinear points) is that of a saddle point, that means that collinear points are always unstable. This can be easily seen by considering the L_1 point. If an object located in L_1 drifts closer to one of the masses, the gravitational attraction it feels from the closer mass would be greater, and it would be pulled closer. However, a test mass displaced perpendicularly from the central line would feel a force pulling it back towards the equilibrium point. In other words, the first three Lagrangian points are stable only if constrained to the plane perpendicular to the line between the two bodies. This is because in this latter case the lateral components of the two

masses' gravity would add to produce a restoring force, whereas the components along the axis between them would balance out.

Although L_1, L_2, and L_3 points are nominally unstable, it turns out that, at least in the restricted three-body problem, it is possible to find stable periodic orbits around these points in the plane perpendicular to the line joining the primaries.

6.4.6 Surfaces of zero velocity

The magnitude of the velocity vector must be a real number, so that:

- if $2\mathcal{U} > C$ then $V^2 > 0$ and the motion is possible;
- if $2\mathcal{U} < C$ then $V^2 < 0$ and no motion is possible.

Thus, the equation

$$2\mathcal{U} - C = 0$$

defines the boundary between the allowed and the forbidden region in the rotating frame, that is, the *surface of zero velocity* or *Hill surfaces*, where $C' = 2\mathcal{U}$.

Recalling the definition of the potential $\mathcal{U}$, it is

$$C' = 2\mathcal{U} = \xi^2 + \eta^2 + 2\frac{1-\mu}{\rho_1} + 2\frac{\mu}{\rho_2}$$

which is a surface in the three-dimensional space $\xi - \eta - \zeta$. The surfaces of zero velocity are symmetrical with respect to the ξ-η plane; only their intersection with this plane will be analyzed in the following. Letting $\zeta = 0$, it is possible to observe the intersection of the surfaces in the $\xi - \eta$ plane (Fig. 6.14).

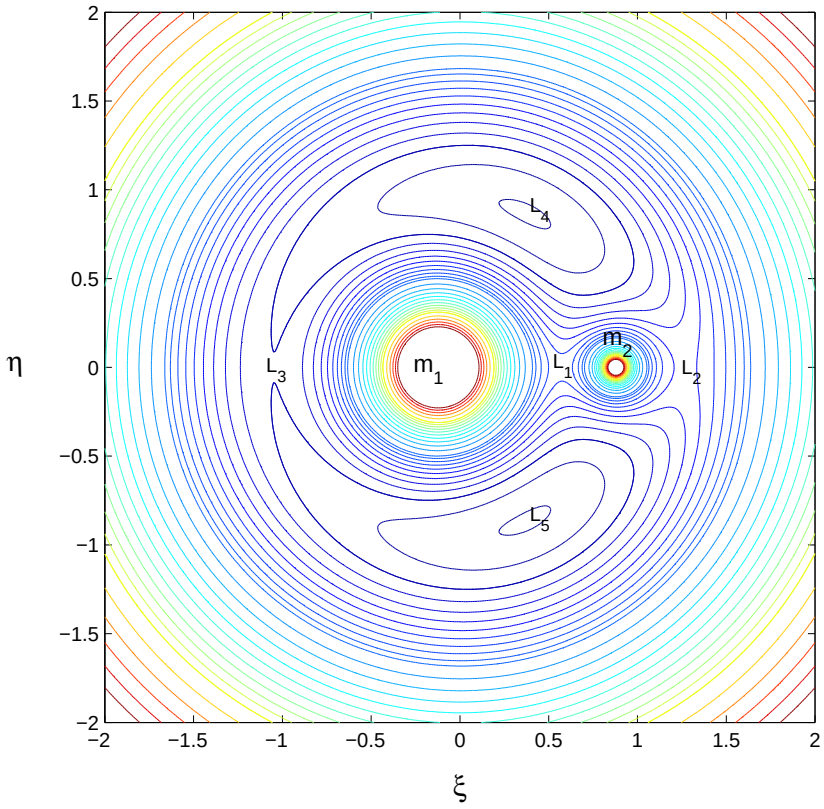

Figure 6.14: **Contour lines of the three–body potential function $\mathcal{U}$ in the $\xi - \eta$ plane.**

Note that if ξ and η are large, the terms $1/\rho_1$ and $1/\rho_2$ are both small and

$$C \approx \xi^2 + \eta^2$$

that is a circle in the $\xi - \eta$ plane. If ρ_1 or ρ_2 is small (i.e. the third body is near one of the two primary masses), the correspondent term $1/\rho$ becomes dominant, and the contour lines are circles around the nearest primary mass,

$$C \approx 2\frac{1-\mu}{\rho_1} \qquad \text{or} \qquad C \approx 2\frac{\mu}{\rho_2}$$

The position dependent function,

$$C' = \xi^2 + \eta^2 + 2\frac{1-\mu}{\rho_1} + 2\frac{\mu}{\rho_2}$$

can be compared with the Jacobi integral C, which is a function of initial position and velocity of the mass m,

$$C = \xi_0^2 + \eta_0^2 + 2\frac{1-\mu}{\rho_{1_0}} + 2\frac{\mu}{\rho_{2_0}} - V_0^2$$

A particle (a spacecraft, in our case) may have relative motion only in regions of space corresponding to values of C' higher then (or at most equal to) its own C.

The constant $-C/2$ corresponds to the total energy of the third body in the non-rotating frame, and in fact represents the maximum value of potential energy that the spacecraft can attain by zeroing its velocity. The spacecraft can only access the region of space with $C' \geq C$, where the potential energy is less than its total energy. As V_0 increases, the third mass can access larger regions of space. Eventually, it can cross from m_1 to m_2 and further escape from the system.

Figure 6.15 presents some zero velocity curves and, in particular, those for C_1, C_2, C_3, corresponding to the values of potential in the collinear Lagrangian points. If C' is very large, the zero velocity curves consist of three circles: the largest one has approximately radius $\sqrt{C}$

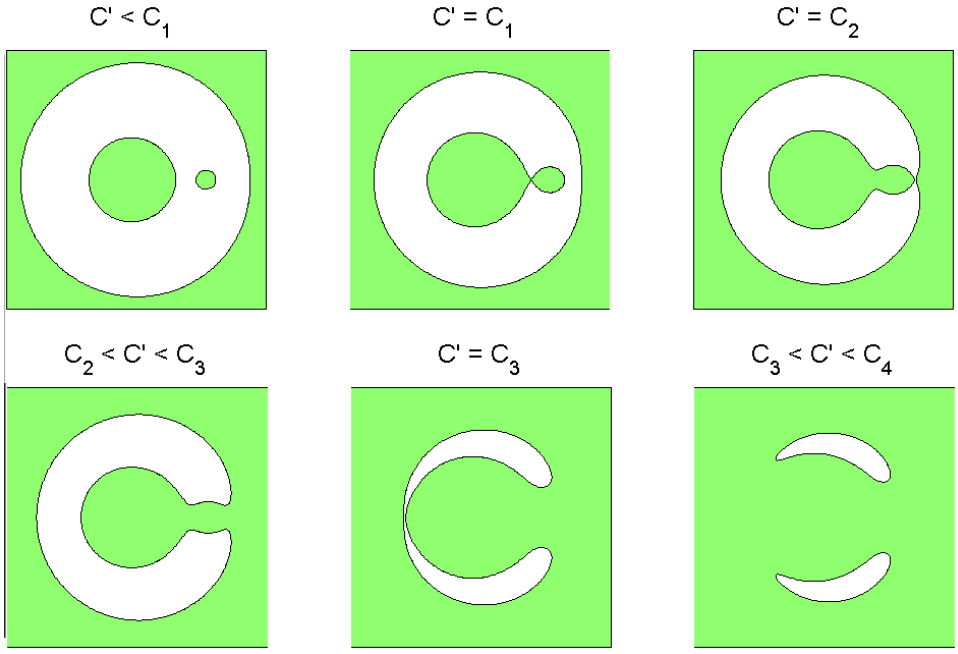

Figure 6.15: **Intersections of zero-velocity surfaces with the $\xi - \eta$ plane.**

and is centered on the origin of the frame; two smaller circles enclose the primaries; a spacecraft orbiting the main body with $C = C'$ is confined to move around it. A spacecraft with $C > C'$ can leave the region around m_1 and orbit around m_2. For $C' > C_2$ the spacecraft can escape from the system, but only leaving the region around the primaries from the side of the second body. This limit disappears for $C' > C_3$. Any place can be reached when $C > C_4 = C_5$.

6.5 Summary and practical considerations

Summing up, in the circular problem there exist five equilibrium points. Three are collinear with the masses in the rotating frame and are unstable. The remaining two are located 60 degrees ahead of and behind the smaller mass in its orbit about the larger mass. For sufficiently small mass ratio of the primaries, these triangular equilibrium points are stable, such that (nearly) massless particles will orbit about these points that in turn orbit around the larger primary.

The L_1 point lies on the line defined by the two main bodies, and between them. An object which orbits the Earth more closely than the Moon would normally have a shorter orbital period than the Moon, but if the object is directly between the bodies, then the effect of the lunar gravity is to weaken the force pulling the object towards the Earth, and therefore increase the orbital period of the object. As the object gets closer to the Moon, this effect become greater until, at L_1 point (the *cislunar lagrangian point*), the orbital period of the object becomes exactly equal to the lunar orbital period.

The L_2 point lies on the line defined by the two primary masses, beyond the smaller of the two. On the side of the Moon away from the Earth, the orbital period of an object would normally be greater than that of the Moon. The extra pull of the lunar gravity decreases the orbital period of the object, which at the L_2 point (the *translunar lagrangian point*) has the same period as the Moon.

If the second body has mass m_2 much smaller than the mass m_1 of the largest one, then L_1 and L_2 are at approximately equal distances r_2 from the second body (L_1 is actually a little closer), given by the radius of the Hill sphere

$$r_2 = R\sqrt[3]{\frac{m_2}{3m_1}}$$

where R is the distance between the two bodies. The Sun-Earth L_1 and L_2 points are distant 1,500,000 km from the Earth, and the Earth-Moon points 61,500 km from the Moon.

The L_3 point lies on the line defined by the two large bodies, beyond the larger of the two. In the Earth-Moon system L_3 is on the opposite side of the Moon, a little further away from the Earth than the Moon is, where the combined pull of the Moon and Earth again causes the object to orbit with the same period as the Moon.

The L_4 and L_5 points lie at the third point of an equilateral triangle with the base of the line defined by the two masses, such that the point is respectively ahead of, or behind, the smaller primary mass in its orbit around the larger one. L_4 and L_5 are sometimes called equilateral Lagrange points or Trojan points.

Perturbations may change this scenario, and an object could not remain permanently stable at any one of these five points. In any case a spacecraft can orbit around them with modest fuel expenditure to maintain such a position, as the sum of the external actions is close to zero.

In the more general case of elliptical orbits, there are no longer stationary points in the same sense: it becomes more of a Lagrangian "area" where the third body makes small odd-shaped orbits about the invisible Lagrangian point; these orbits are commonly referred to as Halo orbits. The Lagrangian points constructed at each point in time as in the circular case form stationary elliptical orbits which are similar to the orbits of the massive bodies. This is due to the fact that Newton's second law, $\dot{\vec{\boldsymbol{p}}} = m\dot{\vec{\boldsymbol{v}}} = \vec{\boldsymbol{f}}$ (where $\vec{\boldsymbol{p}}$ is the momentum, m the mass, $\vec{\boldsymbol{v}}$ the velocity and

$\vec{f}$ the force), remains invariant if force and position are scaled by the same factor. A body at a Lagrangian point orbits with the same period as the two massive bodies in the circular case, implying that it has the same ratio of gravitational force to radial distance as they do. This fact is independent of the circularity of the orbits, and it implies that the elliptical orbits traced by the Lagrangian points are solutions of the equation of motion of the third body.

Stable periodic orbits around collinear points in the plane perpendicular to the line joining the primaries, also known as *halo orbits*, can be exploited for some space missions even if, such perfectly periodic orbits do not exist in a full n-body dynamical system such as the solar system. Quasi-periodic (i.e. bounded but not precisely repeating) Lissajous orbits do exist even in the more realistic n-body scenario. These quasi-periodic orbits are what all libration point missions to date have used. Although they are not perfectly stable, a relatively modest propulsive effort can allow a spacecraft to stay in a desired Lissajous orbit for an extended period of time. It also turns out that, at least in the case of Sun–Earth L1 missions, it is actually preferable to place the spacecraft in a large amplitude (100,000–200,000 km) Lissajous orbit instead of having it sit at the libration point, since this keeps the spacecraft off of the direct Sun–Earth line and thereby reduces the impacts of solar interference on the Earth–spacecraft communications links. Another interesting and useful property of the collinear libration points and their associated Lissajous orbits is that they serve as "gateways" to control the chaotic trajectories of the Interplanetary Superhighway.

The Sun–Earth L_1 is ideal for making observations of the Sun. Objects here are never shadowed by the Earth or the Moon. The sample return capsule Genesis returned from L_1 to Earth in 2004 after collecting solar wind particles there for three years. The Solar and Heliospheric Observatory (SOHO) is stationed in a Halo orbit at the L_1 and the Advanced Composition Explorer (ACE) is in a Lissajous orbit, also at the L_1 point. The Earth–Moon L_1 allows easy access to lunar and Earth orbits with minimal Δv, and would be ideal for a half-way manned space station intended to help transport cargo and personnel to the Moon and back.

The Sun–Earth L_2 offers an exceptionally favorable environment for a space–based observatory since its instruments can always point away from the Sun, Earth and Moon while maintaining an unobstructed view to deep space. The Wilkinson Microwave Anisotropy Probe (WMAP) is already in orbit around the Sun–Earth L_2 and observes the full sky every six months, as the L_2 point follows the Earth around the Sun. The future Herschel Space Observatory as well as the proposed James Webb Space Telescope will be placed at the Sun–Earth L_2. Earth–Moon L_2 would be a good location for a communications satellite covering the Moon's far side.

The Sun–Earth L_3 is only a place where science fiction stories put a Counter-Earth planet sharing the same orbit with the Earth but on the opposite side of the Sun.

By contrast, L_4 and L_5 are stable equilibrium points, provided the ratio of the primary masses is larger than 24.96. This is the case for the Sun–Earth and Earth–Moon systems, though by a small margin in the latter case. When a body at these points is perturbed, it moves away from the point, but the Coriolis force then acts, and bends the object's path into a stable, kidney bean–shaped orbit around the point, as seen in the rotating frame of reference (Fig. 6.16).

In the Sun–Jupiter system several thousand asteroids, collectively referred to as Trojan asteroids, are in orbits around the Sun–Jupiter L_4 and L_5 points (Greek and Trojan camp, respectively). In 1904 Edward Emerson Barnard observed what is now believed the first Trojan Asteroid ever observed from the Earth, but the importance of his observation was not understood at the time. Only three years later the German astronomer Max Wolf discovered an asteroid in the Lagrangian point L_4 of the Sun–Jupiter system, and named it 588 Achilles. The peculiarity of its orbit was soon clear, and many other asteroids were later discovered in both the triangular Lagrange point of the Sun-Jupiter system. There are currently, as of July 2004, 1679 known

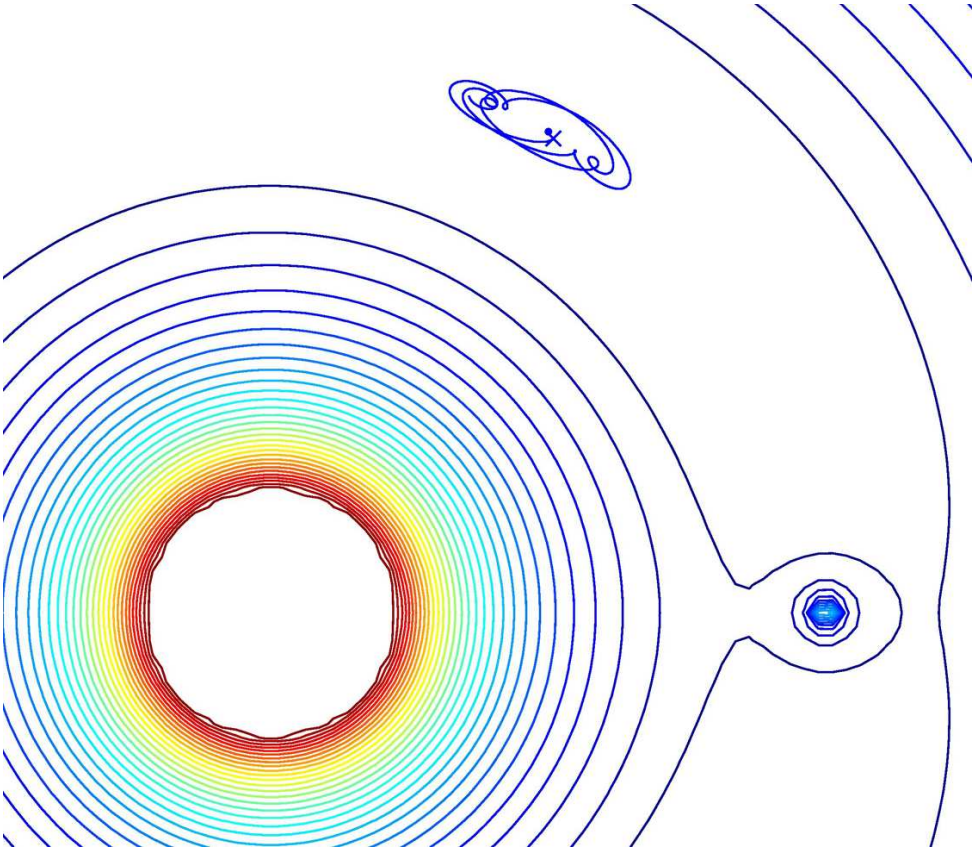

Figure 6.16: **Trajectory in the neighbourhood of an equilateral point.**

Trojan asteroids, 1051 at L_4 and 628 at L_5. There are undoubtedly many others too small to be seen with current instruments.

Other bodies can be found in the Sun–Neptune (four bodies around L_4) and Sun–Mars (5261 Eureka in L_5) systems. There are no known large bodies in the Sun–Earth system's Trojan points, but clouds of dust surrounding the L_4 and L_5 were discovered in the 1950s.

Clouds of dust, called Kordylewski clouds, may also be present in the L_4 and L_5 of the Earth–Moon system. There is still controversy as to whether they actually exist, due to their extreme faintness; they might also be a transient phenomenon as the L_4 and L_5 points of the Earth–Moon system are actually unstable, because of the perturbation induced by the Sun. Instead, the Saturnian moon Tethys has two smaller Moons in its L_4 and L_5 points, Telesto and Calypso, respectively. The Saturnian Moon Dione also has two Lagrangian co–orbitals, Helene at its L_4 point and Polydeuces at L_5. The Moons wander azimuthally about the Lagrangian points, with Polydeuces describing the largest deviations, moving up to 32 degrees away from the Saturn–Dione L_5 point. Tethys and Dione are hundreds of times more massive than their "escorts" and Saturn is far more massive still, which makes the overall system stable.

Appendix - Sphere of influence

Consider a unit–mass spacecraft moving in the proximity of a body of mass m_2, which orbits around a primary body of much greater mass m_1. The distance between the large bodies is R; no forces other than gravitation are considered. One is interested in defining a region of space where the action of the second body is dominant. A precise boundary does not exist, as the transition from the dominance of the primary body to the prevailing action of the second

Table 6.1: **Planetary spheres of influence (Laplace definition)**

Planet	**Mass ratio $\times 10^4$ (Planet/Sun)**	**Radius of the sphere of influence (10^5 km)**
Mercury	0.00164	1.12
Venus	0.0245	6.16
Earth	0.0304	9.29
Mars	0.00324	5.78
Jupiter	9.55	482
Saturn	2.86	545
Uranus	0.436	519
Neptune	0.518	868
Pluto	0.000066	341

body is quite smooth. Practical reasons suggest that it is possible to approximate such a region with a sphere, whose radial extension is conventional. The most credited definitions of sphere of influence, which are due to Laplace and Hill, do not coincide: the former deals with the transition from a computational model to another; the latter definition is based on the limit altitude of stable orbits.

The concept of sphere of influence is therefore a mere convention and to neglect the action of the other primary mass is an approximation. Other forces can be significant, such as radiation pressure or the attraction of other bodies (*e.g.* the gravitation pull of the Sun on an object moving in the Earth–Moon system). Also the definition of the sphere of influence requires that the third object is of negligible mass, in order not to introduce significant contribution to the gravitation field through its own gravity.

Orbits just within the sphere are not stable in the long term; numerical methods demonstrate that stable satellite orbits are inside 1/2 to 1/3 of the Hill sphere radius, with retrograde orbits being more stable than prograde orbits.

The spheres of influence of the planets in the solar system are given in Table 6.1. Data refers to the mean distance from the Sun. The planet with the largest sphere is Neptune; its great distance from the Sun amply compensates for its small mass relative to Jupiter. Neptune has also the largest Hill sphere (0.775 AU of radius), greater than Jupiter's one (whose Hill sphere measures 0.355 AU).

Laplace sphere of influence

The motion of a spacecraft can be studied in a non-rotating reference frame with its origin in the centre of m_2. The only actions on the spacecraft are the attractions of the central body Φ and of the primary body m_1, which is seen as a perturbation $\Delta\Phi$ of the gravitational pull of m_2. Remembering the results presented in Section 4.6.1, it is

$$\Phi = \frac{\mu_2}{\rho^2} \qquad \text{and} \qquad \Delta\Phi = \mu_1 \frac{\rho}{R^3}\sqrt{1+3\cos^2\beta}$$

where ρ is the distance of the spacecraft from m_2. If on the converse, the motion of the spacecraft is analyzed in a non-rotating frame centered in the centre of mass of the largest primary body, the main force F exerted by m_1 and the perturbation force ΔF produced by m_2 in its proximity are instead

$$F = \frac{\mu_1}{r^2} \qquad \text{and} \qquad \Delta F = \frac{\mu_2}{\rho^2}$$

where $r \approx R$ is the distance of the spacecraft from the largest primary mass. According to Laplace, if one desires to neglect the perturbing action, thus retaining only the major effects from the smaller yet closer body, m_2, the two approaches can be considered equivalent when

$$\frac{\Delta F}{F} = \frac{\Delta \Phi}{\Phi} \qquad \Longrightarrow \qquad \frac{\rho^5}{R^5} = \frac{m_2^2}{m_1^2} \frac{1}{\sqrt{1+3\cos^2\beta}}$$

It is convenient to underline that this equivalence is just conventional, and it does not provide any knowledge of the eventual errors.

The radius provided by the above equation defines a surface that is rotationally symmetric about an axis joining the massive bodies. Its shape differs little from a sphere; the ratio of the largest ($\beta = \pi/2$) and smallest (along line connecting the two primary bodies) values of ρ is about 1.15. For convenience the surface is made spherical by replacing the square root with unity (this is equivalent to select the largest sphere tangent to the surface). The sphere of radius

$$\rho_L = R\left(\frac{m_2}{m_1}\right)^{2/5}$$

is known as *Laplace sphere of influence* or simply *sphere of influence*, as it is the most widely used definition.

The Earth's sphere of influence has a radius of about 924,000 km, and comfortably contains the orbit of the Moon, whose sphere of influence extends out to 66,200 km.

Hill sphere

The definition of the *gravitational sphere of influence*, also known as *Hill sphere*, is due to the American astronomer G. W. Hill. It is also called the Roche sphere because the French astronomer E. Roche independently derived its definition.

The determination of the Hill sphere of influence assumes a rotating reference frame with an angular frequency equal to that of the two primary masses rotating about their centre of mass (see Section 6.4.1), and considering the three vector fields due to the centrifugal force and the attractions of the massive bodies. The Hill sphere is the largest sphere within which the sum of the three fields is directed towards the second body. A small third body will orbit the second one in a stable way (no escape allowed) only if it lies always within the Hill sphere during its motion, where the resultant force is centripetal.

The Hill sphere extends between the Lagrangian points L1 and L2, which lie along the line of centres of the two bodies. The region of influence of the second body is shortest in that direction, and so it acts as the limiting factor for the size of the Hill sphere. Beyond that distance, a third object in orbit around the second one would spend at least part of its orbit outside the Hill sphere, and would be progressively perturbed by the main body until it is pulled out from the neighbourhood of m_2 and attracted towards m_1, ending up orbiting the largest primary body.

The distance ρ of L_1 and L_2 from the smaller body is obtained by equating the attractive accelerations of the two primaries to the centrifugal acceleration, that is

$$\frac{\mu_1}{(R\pm\rho)^2} \pm \frac{\mu_2}{\rho^2} = \omega^2(R\pm\rho) \qquad \Longrightarrow \qquad \frac{\mu_1}{(R^2(1\pm\varepsilon)^2} \pm \frac{\mu_2}{\rho^2} = \omega^2 R(1\pm\varepsilon)$$

where the upper sign applies to L_2 and the lower one to L_1, while $\varepsilon = \rho/R$. Remembering that, for $m_2 \ll m_1$ it is $\omega^2 = \mu_1/R^3$, and that

$$(1\pm\varepsilon)^{-2} \approx 1 \mp 2\varepsilon$$

one obtains

$$\frac{\mu_1}{R^2}(1 \mp 2\varepsilon) \pm \frac{\mu_2}{\rho^2} = \omega^2 R(1 \pm \varepsilon) \qquad \Longrightarrow \qquad \pm\frac{\mu_2}{\rho^2} = \pm 3\varepsilon\frac{\mu_1}{R^2} = \pm 3\rho\frac{\mu_1}{R^3}$$

and therefore the radius ρ of the Hill sphere of the smaller primary mass is given by

$$\rho_H = R\sqrt[3]{\frac{\mu_2}{3\mu_1}} = R\sqrt[3]{\frac{m_2}{3m_1}}$$

The Hill sphere for the Earth extends to about 1.5 10^6 km from its centre (approximately 0.01 AU); it should be noted that this is approximately 4 times the Moon–Earth distance, which explains the long–term stability of the motion of the Moon even in presence of the sizable solar perturbation. The radius of the Moon's sphere of influence is close to 61,500 km.

Chapter 7

(Introduction to) Interplanetary trajectories

7.1 The Solar system

7.1.1 Sun and planets

The heliocentric nature of the solar system was proposed by Copernicus in 1530 as a possible explanation for the observed motion of the planets. Although his revolutionary treatise was the first to formulate such a hypothesis in modern times, Aristarchus of Samos had already proposed in the third centurt b.C. a heliocentric model in which the Earth revolves around the Sun spinning about its axis. He based his theory on geometrical grounds, and he was also the first to evaluate the size of the Sun and of the Moon and their distance from the Earth. Unfortunately his work was strongly criticized and gradually forgot.

Nine planets (most of which accompanied by several satellites) encircle the Sun, which has more than 99% of the total mass of the system. A great number of lesser bodies, asteroids or comets, moves in the system, and additional mass is present under the form of meteors and dust clouds (see Cornelisse, 1979, p. 481, for more details).

The mean distance of the planets from the Sun approximately follows Bode's law formulated it in 1772. Given the series of integers 0, 3, 6, 12,..., that is

$$n_1 = 0,\ n_2 = 3,\ n_i = 2n_{i-1} \text{ for } i \geq 3$$

the distance of the planets in Astronomical Units[1] is approximately given by

$$R_i = (n_i + 4)/10 \text{ for } i\ = 1, 2, ..., 9$$

According to some, the ninth position of the series is shared by Neptune and Pluto, thus suggesting that Pluto might be an escaped satellite of Neptune.

The planets' orbits are described by five almost constant orbital elements, whereas the sixth (true longitude) defines the position along the orbit and rapidly changes. Different sources provide Ephemeris, that are tables of the orbit elements at different dates for a great number of major and minor bodies (Tab. 7.2). The orbits of the planets are nearly circular and located in the ecliptic plane, with the exception of the extreme planets, Mercury ($e = 0.206$, $i = 7.00$ deg) and Pluto ($e = 0.249$, $i = 17.14$ deg). In particular Pluto's perihelion lies inside the orbit of Neptune.

[1]One Astronomical Unit is the semimajor axis of the Earth's orbit around the Sun and is often used as a reference distance when dealing with interplanetary trajectories or in describing the features of the Solar Systems.

Table 7.1: **Bode's Law**

Planet	**Bode's Law distance**	**Orbit semi-major axis**
Mercury	0.4	0.39
Venus	0.7	0.72
Earth	1.0	1.00
Mars	1.6	1.52
Asteroids (average)	2.8	2.65
Jupiter	5.2	5.20
Saturn	10.0	9.52
Uranus	19.6	19.28
Neptune	38.8	30.17
Pluto	77.2	39.76

Table 7.2: **Keplerian elements and their rates, with respect to the mean ecliptic and equinox of J2000, valid for the time–interval 1800 AD - 2050 AD.**

Planet	**semimaj.axis** $\boldsymbol{a}$ [AU] ($\dot{a}$) [AU/Cy]	**eccentricity** $\boldsymbol{e}$ ($\dot{e}$) [1/Cy]	**inclination** i [deg] ($\dot{i}$) [deg/Cy]	**true anomaly** $\boldsymbol{\nu}$ [deg] ($\dot{\nu}$) [avg.deg/y]	**long.perihelion** $\boldsymbol{\Pi}$ [deg] ($\dot{\Pi}$) [deg/Cy]	**long.asc.node** $\boldsymbol{\Omega}$ [deg] ($\dot{\Omega}$) [deg/Cy]
Mercury	**0.38709927** (0.00000037)	**0.20563593** (0.00001906)	**7.00497902** (-0.00594749)	**252.25032350** (1494.72674111)	**77.45779628** (0.16047689)	**48.33076593** (-0.12534081)
Venus	**0.72333566** (0.00000390)	**0.00677672** (-0.00004107)	**3.39467605** (-0.00078890)	**181.97909950** (585.17815387)	**131.60246718** (0.00268329	**76.67984255** (-0.27769418)
Earth	**1.00000261** (0.00000562)	**0.01671123** (-0.00004392)	**-0.00001531** (-0.01294668)	**100.46457166** (359.99372450)	**102.93768193** (0.32327364	**0.0** (0.0)
Mars	**1.52371034** (0.00001847)	**0.09339410** (0.00007882)	**1.84969142** (-0.00813131)	**-4.55343205** (191.40302685)	**-23.94362959** (0.44441088)	**49.55953891** (-0.29257343)
Jupiter	**5.20288700** (-0.00011607)	**0.04838624** (-0.00013253)	**1.30439695** (-0.00183714)	**34.39644051** (30.34746128)	**14.72847983** (0.21252668)	**100.47390909** (0.20469106)
Saturn	**(9.53667594** (-0.00125060)	**0.05386179** (-0.00050991)	**2.48599187** (0.00193609)	**49.95424423** (12.22493622)	**92.59887831** (-0.41897216)	**113.66242448** (-0.28867794)
Uranus	**19.18916464** (-0.00196176)	**0.04725744** (-0.00004397)	**0.77263783** (-0.00242939)	**313.23810451** (4.28482028)	**170.95427630** (0.40805281)	**74.01692503** (0.04240589)
Neptune	**30.06992276** (0.00026291)	**0.00859048** (0.00005105)	**1.77004347** (0.00035372)	**-55.12002969** (2.18459453)	**44.96476227** (-0.32241464)	**131.78422574** (-0.00508664)
Pluto	**39.48211675** (-0.00031596)	**0.24882730** (0.00005170)	**17.14001206** (0.00004818)	**238.92903833** (1.45207805)	**224.06891629** (-0.04062942)	**110.30393684** (-0.01183482)

Table 7.3: **Physical properties of Sun and planets.**

Celestial Body	**Diameter [km]**	**(Earth = 1)**	**Rotational period***	**Oblateness**	**Axial tilt [deg]**	**Mass Earth = 1**	**Mass par. [km^{3}s^{-2}]**
Sun	1,392,000	(109)	$\approx$ 25.4 days	$\approx 10^{-5}$	7.25	333,432	$1.327\ 10^{11}$
Mercury	4,879	(0.38)	58.65 days	0.0	2.0	0.055	$2.232\ 10^{4}$
Venus	12,104	(0.95)	-243.02 days	0.0	177.4	0.815	$3.257\ 10^{5}$
Earth	12,742	(1.0)	23 hrs 56 min	0.0034	23.45	1.000	$3.986\ 10^{5}$
Mars	6,780	(0.53)	24 hrs 37 min	0.005	25.19	0.107	$4.305\ 10^{4}$
Jupiter	139,822	(10.97)	9 hrs 55 min	0.065	3.12	317.830	$1.268\ 10^{8}$
Saturn	116,464	(9.14)	10 hrs 40 min	0.108	26.73	95.159	$3.795\ 10^{7}$
Uranus	50,724	(3.98)	-17.24 days	0.03	97.86	14.536	$5.820\ 10^{6}$
Neptune	49,248	(3.87)	16 hours 7 min	0.02	29.56	17.147	$6.896\ 10^{6}$
Pluto	2,390	(0.19)	-6.38 days	0.0	119.6	0.002	$0.797\ 10^{3}$

* Negative numbers indicate retrograde rotation.

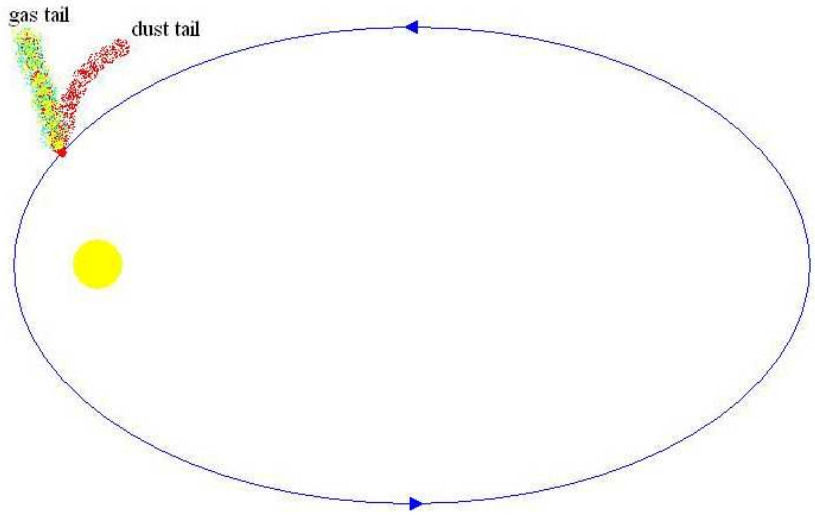

Figure 7.1: **A comet high–ellipticity orbit (note the two distinct tails).**

7.1.2 Other bodies

The fifth position in Bode's sequence is occupied by the asteroids, which might be the remnants of a destroyed body or the matter destined for a never–born planet (see Tab. 7.1). About 220 asteroids are larger than 100 km. The biggest asteroid is Ceres, which is about 1000 km across. The total mass of the Asteroid belt is estimated to be 1/35th that of the Moon, and of that total mass, one–third is accounted for by Ceres alone.

Comets are small bodies in the solar system characterized by a nucleus, generally less than 50km across, which is composed of rock, dust, and ice. When a comet approaches the inner solar system, radiation from the Sun causes its outer layers of ice to evaporate. The streams of released dust and gas form a huge but extremely tenuous coma, which can extend over 150 million km (1 AU). The coma become visible from the Earth when a comet passes closes to the Sun, the dust reflecting sunlight directly and the gases glowing due to ionization. The force exerted on the coma by the solar wind and radiation pressure cause two distinct tails to form pointing away from the sun in slightly different directions. The dust is left behind in the comet's orbit so that it often forms a curved tail; the ionized–gas tail always points directly away from the Sun, since the gas is more affected by the solar wind than dust is, and follows magnetic field lines rather than an orbital trajectory.

Comets are classified according to their orbital periods. Short period comets have orbits of less than 200 years (comet Encke never is farther from the Sun than Jupiter). Long period comets have larger orbits but remain gravitationally bound to the Sun. Single–apparition comets have parabolic or hyperbolic orbits which will cause them to permanently exit the solar system after one pass by the Sun. Short–period comets are thought to originate in the Kuiper belt, which is an area of the solar system extending from within the orbit of Neptune (at 30 AU) to 50 AU from the Sun, at inclinations consistent with the ecliptic. Long–period comets are believed to originate in the 50,000 AU distant Oort cloud, which is a spherical cloud of debris left over from the condensation of the solar nebula and containing a large amount of water in a solid state. Some of these objects were perturbed by gravitational interactions and fell from their circular orbits into extremely elliptical orbits that bring them very close to the Sun. Because of their low masses, and their elliptical orbits which frequently take them close to the giant planets, cometary orbits are often perturbed and constantly evolving. Some are moved into sungrazing orbits that destroy the comets when they get near to the sun, while others are thrown out of the solar system forever.

7.1.3 Orbital resonance

Some bodies in the solar system are in orbital resonance, a phenomenon that occurs when two orbiting bodies have periods of revolution that are in a simple integer ratio, so that they exert a regular gravitational influence onto each other. This can stabilize the orbits and protect them from gravitational perturbation. For instance, Pluto and some smaller bodies called Plutinos were saved from ejection by a 3:2 resonance with Neptune. The Trojan asteroids may be regarded as being protected by a 1:1 resonance with Jupiter. Orbital resonance can also destabilize one of the orbits. For instance, there is a series of almost empty lanes in the asteroid belt called Kirkwood gaps where the asteroids would be in resonance with Jupiter. A Laplace resonance occurs when three or more orbiting bodies have a simple integer ratio between their orbital periods. For example, Jupiter's moons Ganymede, Europa, and Io are in a 1:2:4 orbital resonance.

A companion object of the Earth has been discovered in 1988; the asteroid 3753 Cruithne is on an elliptical orbit which has almost the same period of the Earth. Due to close encounters with the planet, this asteroid periodically alternates between two regular solar orbits. When the asteroid approaches the Earth (the minimum distance is 12 10^6 km), it takes orbital energy from the planet and moves into a larger and slower orbit. Some time later, the Earth catches up with the asteroid and takes the energy back; as a consequence the asteroid falls back into a smaller, faster orbit and a new cycle begins.

Epimetheus and Janus, satellites of Saturn, have a similar relationship, though they are of similar masses and so actually exchange orbits with each other periodically (Janus is roughly 4 times more massive, but still light enough for its orbit to be altered).

7.2 Interplanetary missions

7.2.1 Generalities

The computation of a precision trajectory for an interplanetary mission requires the numerical integration of the complete equations of motion where all perturbation effects are taken into account. Initial values of state variables, namely position and velocity, and velocity increments along the trajectory are modified until a satisfying mission profile is found.

The spacecraft spends most of the trip–time under the dominant gravitational attraction of the Sun and the perturbations caused by the planets are negligible; only for brief periods the trajectory is shaped by the departure and arrival planets. In these circumstances the patched–conic model is absolutely adequate, and this simple approach is generally adopted for preliminary analyses. The Sun attraction is neglected during the planetocentric legs, and, for the sake of simplicity, the velocity relative to the planet, on exiting or entering the sphere of influence, is assumed equal to the hyperbolic excess velocity. The dimension of the sphere can be neglected in comparison with the distance from the Sun, and the extremes of the heliocentric leg are assumed to coincide with the positions of the planets at the time of planet encounter.

In spite of the apparent model simplicity, the mission design task is far from trivial. Numerically, it is necessary to evaluate spacecraft and planets orbit parameters with extremely tight accuracies. Apparently minor errors on orbit parameters on the heliocentric legs may induce considerable errors on orbit parameters upon enetring the sphere of influence of the target planet. The same consideration is valid for orbit injection errors, and the same is true also for actual errors during the mission. A certain amount of fuel for correction manoeuvres must be included in the fuel budget, nonetheless, the more accurate the trajectory, the smaller the penalty on fuel. Although the planets' orbits lie almost on the same plane, the cost of orbit plane change is made very high by the high velocity involved in interplanetary travel. Even more difficult is the case of complex interplanetary missions, where the probe is required to fly–by a

given sequence of planets. This may be useful in order to exploit the swing–by effect, where the probe gets energy by entering the planet's sphere of influence "from behind". A minor error on the first planet intercept may result in an unacceptable error on the approach to the last planet of the series.

7.2.2 Elongation and sinodic period

The design of an interplanetary mission requires also the careful evaluation of the so–called launch window, that is, the time interval during which a succesfull launch can be performed that takes the space probe in the best conditions for intercepting the planet with the minimum amount of fuel and the maximum payload.

The launch window depends on the relative motion of departure and target planet at launch time with respect to the planet. The planets configuration at departure time depends on the selected transfer trajectory. The *elongation* is the angle under which the target planet is seen from the departure one. If the target planet orbit is outside that of the departure planet, than there are four particular values for the elongation:

1) $\epsilon = 0$ (*conjunction*, where the target planet is on the same line with the departure one and the Sun, on the same side with the Sun);

2) $\epsilon = \pi$ (*opposition*, where the target planet and the sun lie on the same line with the departure one, but on opposite directions);

3 and 4) $\epsilon = \pm\pi/2$ (*eastern* and *western quadrature*).

On the converse, when the target planet orbit is inside that of the departure planet, the elongation will lie on an interval $\epsilon < \epsilon_{\max}$. In this case the four particular values are:

1 and 2) $\epsilon = 0$ (*inferior* and *superior conjunction*, where the target planet is on the same line with the departure one and the Sun, on the same or on the other side of the Sun, respectively);

3 and 4) $\epsilon = \pm\epsilon_{\max}$ (*eastern* and *western maximum elongation*).

If a launch opportunity is given for a certain value of the elongation and is lost, the same trajectory will be again available after one *sinodic period* τ, that is, the period after which the two target planet and the Sun are again in the same relative configuration with respect to the departure planet. This period is obtained from the equation $\omega_2\tau = \omega_1\tau + 2\pi$, so that it is

$$\tau = 2\pi/(\omega_2 - \omega_1)$$

7.3 Heliocentric transfer

When the patched–conic approximation is used, the study of an interplanetary mission begins with the heliocentric transfer orbit.

The heliocentric velocity at departure from the Earth

$$\vec{\boldsymbol{v}}_1 = \vec{\boldsymbol{v}}_\oplus + \vec{\boldsymbol{v}}_\infty \quad (\text{where } v_\oplus = 29.77 \text{ km s}^{-1})$$

is the sum of the Earth velocity and the hyperbolic excess velocity. The latter is assumed to be equal to the spacecraft velocity relative to the Earth. In general it is $\vec{\boldsymbol{v}}_\oplus \gg \vec{\boldsymbol{v}}_\infty$, due to the modest capabilities of present space propulsion, so that the maximum angle between $\vec{\boldsymbol{v}}_\oplus$ and $\vec{\boldsymbol{v}}_1$ is quite small. In particular the heliocentric leg will lie in a plane that can assume only a modest inclination away from the ecliptic plane. In any case the Earth velocity is better exploited if the spacecraft departs from our planet either in the same or in the opposite direction.

Table 7.4: **Hohmann transer to planets of the Solar system from the earth.**

Planet	**Radius** [AU]	τ [y]	a_H [AU]	$t_2 - t_1$ [y]	γ_1 [deg]	**Helioc.vel.** v_1 [km/s]	**Vel.incr.** Δv_1 [km/s]
Mercury	0.387	0.317	0.69355	0.289	-251.6734	22.258	7.535
Venus	0.723	1.599	0.86167	0.400	-54.0305	27.297	2.496
Mars	1.524	2.135	1.26186	0.709	44.3458	32.739	2.946
Jupiter	5.203	1.092	3.10144	2.731	97.1578	38.588	8.795
Saturn	9.537	1.035	5.26834	6.046	106.0927	40.085	10.291
Uranus	19.189	1.012	10.09458	16.036	111.3215	41.077	11.284
Neptune	30.070	1.006	15.53496	30.615	113.1596	41.450	11.657
Pluto	39.482	1.004	20.24106	45.532	113.9274	41.610	11.817

7.3.1 Ideal planets' orbits

The orbit of most of the planets can be considered circular and coplanar, and the most efficient transfer between them uses the Hohmann ellipse. One easily computes the heliocentric velocity and then the hyperbolic excess velocity on leaving the Earth sphere of influence

$$\mathcal{E}_\odot = \frac{\mu_\oplus}{R_1 + R_2} \; ; \quad v_1 = \sqrt{2\left(\mathcal{E}_\odot + \frac{\mu_\oplus}{R_1}\right)} \; ; \quad v_\infty = |v_1 - v_\oplus|$$

where R_1 and R_2 are the orbit radii of the departure and arrival planet, respectively. A transfer to an inner planets requires that the spacecraft is launched in the direction opposite to the Earth's orbital motion.

The mission can be carried out using any of the ∞^2 trajectories that intersect or are tangent to the planets' orbits. A transfer other than the Hohmann ellipse may be preferred if presents minor sensitivity to injection errors, permits an Earth return trajectory, or simply reduces the trip time.

Power requirements and solar interference on communications between ground stations and spacecraft depend on the distance and phase angle $\alpha_2 = \gamma 1 + (\omega_2 - \omega_1)(t_2 - t_1)$ between the planets at the arrival time

$$\gamma_2 = \gamma_1 + \omega_2(t_2 - t_1) - (\nu_2 - \nu_1) \; ; \quad d = \sqrt{R_1^2 + R_2^2 - R_1 R_2 \cos\alpha_2}$$

and influence the selection of the heliocentric transfer orbit. Assuming $\gamma_2 = 0$ (that is, a direct planet hit), and neglecting "third bodies" effects (that is, the presence of other celestial bodies during the transfer orbit is neglegted), the characteristics of direct Hohmann transfer to Solar system's planet can be easily evaluated (Tab. 7.4).

By increasing the velocity increment at injection onto the transfer orbit, a minor trip time to the target planet is achieved due to greater energy, at the expenses of a greater fuel consumption. Of course, the phase angle at departure must be changed accordingly, which results in a variation of the launch window with respect to the Hohmann transfer case. Some payload can be traded off for opening a launch window. It must be remembered that the same mission can be flown only after a synodic period.

7.3.2 Accounting for real planets' orbits

When a more accurate mission design is required, it is necessary to perform a parameter analysis that uses the time of departure from the Earth and the arrival time (or the trip time) as independent variables. The conic orbit is a solution of the Lambert problem, and the parameters

are varied until an optimal solution is found (in the sense of the best trade–off between minimum–time, minimum–fuel and maximum payload).

Among many other perturbations of the ideal, coplanar transfer, it must be remembered that the target planet will usually be outside of the ecliptic plane. This meand that at some point it is necessary to perform an orbit inclination change, which is expensive in terms of fuel consumption. A good solution is to apply a midcourse impulse in quadrature with the arrival point, that minimises the magnitude of the plane change required. In this situation, the required plane change variation is exactly equal to the latitude of the target planet at the time of intercept, β, so that the required velocity increment is given by

$$\Delta v = 2v_\vartheta \sin(\beta/2)$$

7.4 Earth departure trajectory

After the selection of the transfer orbit to the target planet, it is necessary to design the escape trajectory from the Earth, which takes place within the Earth's sphere of influence. It must be recalled that upon leaving the Earth's sphere of influence, tha radius of which is approximately 10^6 km, the spacecraft velocity in the heliocentric frame must be $v_1 = v_\oplus + \Delta v_1$, where $\Delta v_1 \approx v_\infty$, that is, the velocity increment to enter the transfer orbit is approximately equal to the hyperbolic excess speed upon leaving the Earth's sphere of influence.

The spacecraft velocity at departure from the Earth surface or from a point with radius r_0 on a parking LEO is thus obtained by considering the conservation of energy during the geocentric, hyperbolic leg

$$\frac{v_0^2}{2} - \frac{\mu_\oplus}{r_0} = \frac{v_\infty^2}{2} - \frac{\mu_\oplus}{\rho_\oplus} \approx \frac{v_\infty^2}{2} \quad \Longrightarrow \quad v_0 = \sqrt{v_\infty^2 + \frac{2\mu_\oplus}{r_0}}$$

Departure is usually from a parking orbit, the altitude of which derives from a compromise between gravitational losses (smaller, for lower orbits) and atmospheric drag (that are higher if r_0 is low).

The departure impulse is $\Delta v_0 = v_0 - v_{c_0}$. In order to avoid misalignment losses ($\varphi_1 = \varphi_0 = 0$), $\Delta\vec{v}_0$ and the escape asymptote must be parallel to the spacecraft orbital velocity on LEO and the Earth's orbital velocity, respectively. In order to achieve such a result, the periaxis of the hyperbolic escape trajectory (that is, the injection point) must form an angle Φ with the direction of Earth's orbital velocity which satisfies the equation

$$\cos\Phi = -1/e$$

where the eccentricity is obtained from the expressions of $a = -\mu_\oplus/v_\infty^2$ and $r_0 = r_P = a(1-e)$.

Any plane containing the nominal $\vec{v}_\infty$ is permitted, so that virtually any place on the Earth surface can be the launch site. Inasmuch as the displacement from the ecliptic remains limited, if compared to interplanetary distances, orbit inclination in the heliocentric frame will be extremely small, if $\vec{v}_\infty$ is parallel to $\vec{v}_\oplus$.

Injection errors must be kept as low as possible, as the hyperbolic excess speed depends strongly on the value of injection velocity, v_0. Assuming a given LEO of radius r_0, it is

$$v_0 \mathrm{d}v_0 = v_\infty \mathrm{d}v_\infty \quad \Longrightarrow \quad \frac{\mathrm{d}v_\infty}{v_\infty} = \left(\frac{v_0}{v_\infty}\right)^2 \frac{\mathrm{d}v_0}{v_0}$$

For a Hohmann transfer to Mars, where $v_\infty = 2.95$ km s^{-1} and $v_0 = 11.6$ km s^{-1}, this meand that a 1% error in v_0 will result in a 15% error on v_∞.

7.5 Arrival to the target planet

The heliocentric leg is always tangent to the departure planet velocity vector, in order to exploit as much as possible its momentum upon injection on the transfer orbit. On the converse, the transfer orbit usually intersects that of the target planet with an angle φ_2, unless a Hohmann transfer is selected.

The velocity v_2 at intercept is given by

$$v_2 = \sqrt{2\left(\frac{\mu_\odot}{r_2} + \mathcal{E}_t\right)}$$

while the angle φ_2 satisfies the equation

$$\cos\varphi_2 = h_t/(r_2 v_2)$$

The velocity relative to the target planet is

$$\vec{\boldsymbol{v}}_3 = \vec{\boldsymbol{v}}_2 - \vec{\boldsymbol{v}}_T$$

where $\vec{\boldsymbol{v}}_T$ is the planet orbital speed.[2] From the law of cosines it is

$$v_3^2 = v_2^2 + v_T^2 - v_2 v_T \cos\varphi_2$$

The angle ϑ between the planet's velocity and $\vec{\boldsymbol{v}}_3$ is obtained from the law of sines

$$\sin\vartheta = \frac{v_2}{v_3}\sin\varphi_2$$

If a direct hit to the target planet is sought, the phase angle at departure must be chose in such a way that $\gamma_2 = 0$. In this case, the intercept point on the planet's orbit will be achieved when also the planet is there. This means that the spacecraft enters the planet's sphere of influence with a velocity vector parallel to the local vertical to the planet centre, so that the approach trajectory to the planet is a straight line along which the probe flies at hyperbolic speed.

If a fly–by trajectory is sought, the phase angle at departure must be changed accordingly, in order to miss the planet by a certain amount distance x (measured along the planet's orbit). In this case it is

$$\gamma_1 = (\nu_2 - \nu_1) - \omega_2(t_2 - t_1) \pm x/r_2$$

where the + sign means a passage behind the planet, while the - sign is for passaged ahead of it. It should be noted that the former is typical of missions that requires injection into a planetary orbit, while the latter allows the exploitation of the swin–by effect, where the spacecraft gains energy from the planet.

Letting the offset distance of $\vec{v}_3$ be defined as $y = x\sin\vartheta = x(v_2/v_3)\sin\varphi_2$ and remembering that the velocity $\vec{\boldsymbol{v}}_3$ represents the hyperbolic excess speed upon entering the planet's sphere of influence, the (constant) magnitude of the angular momentum is $h = yv_3$, while the energy of the approach orbit is $\mathcal{E}_e \approx v_3^2/2$. As a consequence, the other orbital elements can easily be found as $p = h^2/\mu_T$ (parameter) and $e = \sqrt{(1 + 2\mathcal{E}_e h^2/\mu_T^2}$, where μ_T is the gravitational parameter of the target planet.

The periapsis radius is thus given by $r_p = p/(1+e)$ and from conservation of angular momentum, the velocity at periapsis passage is $v_p = yv_3/r_p$.

[2] In this section, the subscript T will indicate quantities related or referred to the target planet.

In general, the design variable is a given periapsis radius (e.g. for injection onto a planetary orbit), which allows for the evaluation of a certain offset distange, $y = v_p r_p / v_3$. But from conservation of energy it is also

$$\mathcal{E}_e = \frac{v_3^2}{2} = \frac{v_p^2}{2} - \frac{\mu_T}{r_p} \quad \Longrightarrow \quad v_p = \sqrt{v_3^2 + \frac{2\mu_T}{r_p}}$$

so that the offset distance is given by

$$y = \frac{r_p}{v_3}\sqrt{v_3^2 + \frac{2\mu_T}{r_p}}$$

From the practical point of view, it is important to determine the offset distance which results into a periapsis radius equal to the radius of the target panet. This value is called *impact parameter*, b, inasmuch as any offset minor than b will result with an impact on the planet surface. By letting $r_p = r_T$, it is

$$b = \frac{r_T}{v_3}\sqrt{v_3^2 + \frac{2\mu_T}{r_T}}$$

It should be noted that, being $2\mu_T/r_T > 0$, it is always $b > r_T$.

The circle of radius b placed at the boundaries of the sphere of influence and perpendicular to the direction passing through the planet in the direction of the asymptote of the entry hyperbola is the *effective collision cross section*. If the probe must enter the planet atmosphere, which is always a thin layer of width h_a several order of magnitude smaller than the planet radius, $h_a \ll r_T$, the *entry corridor* is an anulus of inner radius b and width db, where $db = (\mathrm{d}b/\mathrm{d}r_T)h_a$ is extremelemy small.

Bibliography

1. Richard Battin, *An Introduction to the Mathematics and Methods of Astrodynamics*, AIAA Education Series, New York, 1987.

2. Roger R. Bate, Donald D. Mueller, Jerry E. White, *Fundamentals of astrodynamics*, Dover, New York, 1971

3. Fred P.J. Rimrott, *Introductory orbit dynamics*, Vieweg, Braunschweig, 1989

4. J.W. Cornelisse, H.F.R. Schyer, K.F. Wakker, *Rocket propulsion and spaceflight dynamics*, Pitman, London, 1979

5. B. Wie, *Space vehicle dynamics and control*, AIAA Education Series, Reston, 1998.

6. M.H. Kaplan, *Modern Spacecraft Dynamics and Control*, J. Wiley & Sons, New York, 1976.

7. G. Mengali, *Meccanica del Volo Spaziale*, Edizioni Plus, Pisa, 2001