Summer Workshop on

Advanced Topics in Astrodynamics

Barcelona, July 5-10, 2004

Solar Sailing Lectures

by

Macdonald and Hughes

Institut
d'Estudis Espacials
de Catalunya
IEEC

UNIVERSITY
of
GLASGOW

Contents

1. Introduction to Solar Sailing 4
1.1 Historical Perspective to Date 4
1.2 Solar Sail Configurations 6
1.3 Performance Metrics 7
1.4 Solar Sail Orbits 7
1.5 Further Reading 9
1.5.1 Historical interest 9
1.5.2 Selected introductory papers 10
1.5.2 Solar sailing books 11

2. Solar Radiation Pressure 12
2.1 Radiation Pressure – Electromagnetic Description 14
2.2 Force on a Perfectly Reflecting Solar Sail 16
2.3 Radiation Pressure due to a Finite Solar Disc 18
2.4 Solar Sail Force Models & Orbit Perturbations 21
2.5 Further Reading 23
2.5.1 Historical interest 23
2.5.2 Radiative transfer 23
2.5.3 Solar sail force model 23

3. Solar Sail Orbital Dynamics 24
3.1 Heliocentric Orbital Dynamics 24
3.2 Conic Section Orbits 24
3.3 Logarithmic Spiral Trajectory 25
3.4 Optimal Sail Force Vector 28
3.5 Lagrange Variational Equations 30
3.6 Locally Optimal Control Laws 31
3.7 True Minimum Time Trajectories 36
3.8 Non-Keplerian Orbits 38
3.9 Planet-Centred Orbits 44
3.9.1 Earth Escape – On / Off Switching 46

Cover image is The Statue of Sol at the British Library, in London.

3.9.2 Earth Escape – Orbit Rate Steering 47
3.9.3 Earth Escape – Locally Optimal Steering 50
3.9.4 Earth Escape – Polar Orbit 52
3.10 Heliocentric Optimisation Software 57
3.10 Further Reading 62
3.10.1 Sun-centred trajectories 62
3.10.2 Minimum time trajectories 62
3.10.3 Planet-centred trajectories 63
3.10.4 Miscellaneous 64

1. Introduction to Solar Sailing

Solar sailing is a unique and elegant concept, which transcends the reliance on reaction mass exhibited by conventional propulsion systems by gaining momentum from an ambient source, Solar Electromagnetic Radiation. Unlimited by a finite reaction mass, solar sails can provide a continuous acceleration, limited only by the lifetime of the sail materials.

The momentum carried by individual photons is extremely small. Thus, to provide a suitably large momentum transfer we require the sail to have a large surface, while maintaining as low a mass as possible. At best a solar sail will experience only 9 N of force per square kilometre of sail located in Earth orbit. Adding the impulse due to incident and reflected photons it is found that the thrust vector is directed normal to the surface of the sail, hence by controlling the orientation of the sail relative to the Sun we can gain or lose orbital angular momentum.

Using momentum gained only by reflecting ambient Sunlight the sail is slowly but continuously accelerated to accomplish a wide-range of potential missions. Solar sailing, with the analogies to conventional sailing, may seen a whimsical and romantic notion, however the potential real world benefits offered through solar sailing are quickly found to overshadow such thoughts.

1.1 Historical Perspective to Date

Solar sailing has been considered as a practical means of spacecraft propulsion only relatively recently, however the fundamental ideas are much older. The original concept of solar sailing dates back to the era of the Soviet pioneers of astronautics and before. The existence of light pressure was demonstrated in theory by the Scottish physicist James Clerk Maxwell in 1873, and later measured experimentally in a precision laboratory test by the Russian physicist Peter Lebedew in 1900. Around this period a number of science fiction authors wrote of spaceships propelled by mirrors, most notably the French authors Faure and Graffigny in 1889, however it was not the until early 20th century that the idea of a practical solar sail was articulated. In the early 1920s the Soviet father of astronautics Konstantin Tsiolkovsky and his colleague Fridrickh Tsander both wrote of "*using tremendous mirrors of very thin sheets*" and "*using the pressure of sunlight to attain cosmic velocities*". There is some uncertainty regarding dates but it appears that Tsander was the first to write of practical solar sailing some time late in the summer of 1924. His ideas seem to have been inspired in part by Tsiolkovsky's more general writings on propulsion using light from 1921.

Following the initial work by Tsiolkovsky and Tsander the concept of solar sailing appears to have remained unknown for over thirty years, until in the 1950s when the concept was re-invented and published once again in the popular literature. The first American author to propose solar sailing appears to have been the aeronautical engineer Carl Wiley, writing under the pseudonym Russell Sanders to protect his professional credibility. In May 1951, in

Astounding Science Fiction, Wiley discussed the design of a feasible solar sail and strategies for orbit raising in some technical detail. In particular he noted that solar sails could be 'tacked' allowing a spiral inwards towards the Sun. Even in 1951 Wiley was optimistic about the benefits of solar sailing for interplanetary travel and saw it as ultimately more practical than rocket propulsion. A similarly optimistic view was taken sometime later in 1958 through a separate proposal and evaluation by Richard Garwin, then at the IBM Watson laboratory of Columbia University. Garwin authored the first solar sail paper in a western technical publication, the journal Jet Propulsion, and coined the term "*solar sailing*". Wiley recognised that solar sails require no propellant and are continuously accelerated, therefore allowing large velocity changes over an extended period of time. Such was Garwin's enthusiasm and optimism for solar sailing he concluded that "*there are considerable difficulties connected with space travel, but those connected with the sail appear relatively small*".

Following the discussion of solar sailing by Garwin, more detailed studies of the orbits of solar sails were undertaken during the late 1950s and early 1960s. Several authors were able to show that, for a fixed sail orientation, solar sail orbits are of the form of logarithmic spirals. Simple comparisons of solar sailing with chemical and ion propulsion systems showed that solar sails could match or out perform these systems for a range of mission applications. Such early studies explored the fundamental problems and benefits of solar sailing, but lacked a specific mission to drive detailed analyses and to act as a focus for future utilisation.

By 1973 NASA was funding low-level studies of solar sailing at the Battelle laboratories in Ohio, which gave positive recommendations for further investigation. During the continuation of this work Jerome Wright, who would later move to JPL, discovered a trajectory that could allow a solar sail to rendezvous with comet Halley at its perihelion in the mid-1980s. The flight time of four years would allow for a late 1981 or early 1982 launch. Until then a difficult rendezvous mission was thought to be near impossible in such a short time using the technology of the day. A seven to eight year mission had been envisaged using solar-electric ion propulsion, requiring a launch as early as 1977. The science community saw a rendezvous as essential for a high quality mission. These positive results prompted then NASA Jet Propulsion Laboratory director Bruce Murray to initiate an engineering assessment study of the potential readiness of solar sailing. Newly appointed from a faculty position at the California Institute of Technology, Murray adopted solar sailing as one his bold '*purple pigeon*' projects, as opposed to the more timid '*grey mice*' missions he believed JPL had been proposing for future missions. Following this internal assessment, a formal proposal was put to NASA management in September 1976. The design of a comet Halley rendezvous mission using solar sailing was initiated in November of the same year. As a result of the interest in solar sailing, proponents of solar-electric propulsion re-evaluated their performance estimates and in the end were competing directly with solar sailing for funding. The solar electric propulsion system had a larger advocacy group both within NASA and in industry. As a result of an evaluation of these two advanced propulsion concepts

NASA selected the solar electric system in September 1977, upon its merits of being a less, but still considerable risk for a comet Halley rendezvous. A short time later a rendezvous mission using solar electric propulsion was also dropped due to escalating cost estimates. The enthusiasm of the science community for a rendezvous mission had in fact lead to the lower cost fly-past option being discounted. Therefore, when the advanced propulsion required to enable a rendezvous mission was deleted, some careful back stepping was required to justify a lower cost, but less capable, fly-past mission. Ultimately though, it was too late and Comet Halley was intercepted by an armada of Soviet, Japanese and European spacecraft, but no American craft.

A true solar sail has yet to fly, however significant steps have been taken since the beginning of the 1990's. In February 1993, under the guidance of Vladimir Syromiatnikov, the Russian Space Regatta Consortium deployed a 20-m spinning reflector, Znamya, form a Progress supply vehicle. Observed from the Mir station this deployment showed that spin deployment could be controlled by passive means. In May 1996 a large deployable reflector was demonstrated during the Shuttle mission STS-77. The 14-m Inflatable Antenna Experiment was primarily designed as a radio frequency reflector, but the promise of inflatable technology towards solar sailing was clearly demonstrated despite mission anomalies. In Köln in December 1999 the German space agency, DLR, in association with ESA deployed a square, 20-m solar sail. This deployment now forms the basis for a future DLR/ESA in-orbit deployment demonstration perhaps in early 2006. Furthermore, in August 2004 the Institute of Space and Astronautical Science in Japan, ISAS, deployed two solar sails in space from a S-310 sounding rocket. NASA and several private enterprises all seeking to advance solar sailing are also conducting further work. It is thus clear that the technology is currently undergoing a revival in interest and a renewed drive towards flight status.

1.2 Solar Sail Configurations

The essential requirement of any solar sail design is to provide a large, reflective surface with minimal structural support mass. Furthermore, this surface should be easily deployable and manufactured. The first sail design concept considered is the square sail, the optimum design of which has four deployable spars cantilevered from a central load-bearing hub, as shown in **Figure 1.1**. Attitude control of the square solar sail can be achieved by inducing torques generated by articulated reflecting vanes attached to the spar tips, or through relative translation of the centre-of-mass and centre-of-pressure of the sail.

An alternative concept to the boom supported film is to use spin-induced tension, as in the heliogyro or spinning disc sail. At first the heliogyro appears a more attractive design as the mechanical cantilever spars are removed, however the blades may require edge stiffeners to transmit radial loads and provide torsional stiffness to allow blade rotation for attitude control. The primary advantage of the heliogyro is the ease of packing and deployment. The third concept is the disc solar sail, in which a film is held in tension through

spin, the film may be one continuous sheet or several segments, which together generate a near-continuous film, as in **Figure 1.1**. The disc sail offers the same benefits of the heliogyro but without the high aspect ratio blades.

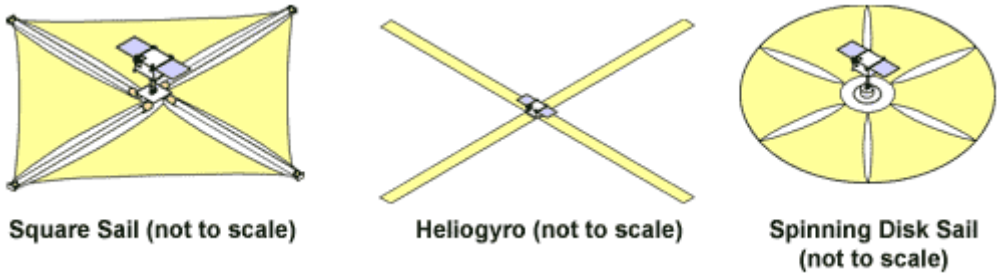

Figure 1.1 Solar sail design concepts (Image by NASA)

1.3 Performance Metrics

To enable sail design, configuration and/or requirements to be compared a standard performance metric is required. The most common metric for astrodynamics is the sail characteristic acceleration, defined as the solar radiation pressure acceleration experienced by a sail facing the Sun at the distance of one Astronomical Unit (AU). At this radius the solar radiation pressure, P, is 4.56 x 10^{-6} N m^{-2}. Therefore, multiplying the pressure by sail area, A, we get the solar radiation pressure force on the sail. Dividing by the sail mass, m, gives the sail acceleration. A factor of two is added to account for reflectivity, since reflected photons impart an equal and opposite force to incident photons. Finally, a finite sail efficiency factor is incorporated to account for non-ideal behaviour of the sail. The sail characteristic acceleration is then defined as

$$a_o = \frac{2\eta P}{\sigma} \ , \ \sigma = \frac{m}{A} \qquad [1.1]$$

where σ is the solar sail mass per unit area, termed the sail loading. We note that the actual sail acceleration is a function of heliocentric radius and sail orientation, however this metric allows an accurate and fair comparison to be made between concepts.

1.4 Solar Sail Orbits

The physics of solar radiation pressure can be described in several ways, as will be discussed in **Section 2**. The simplest is to think about the transfer of momentum to the solar sail by photons, the quantum packets of energy of which light is composed. As a photon hits the surface of the sail it imparts a momentum onto the sail film, thus applying an impulse to the entire solar sail. Then, when the photon is reflected a reaction impulse will also be exerted on the solar sail. The combination of these two impulses summed across the entire flux of photons incident on the sail film then leads to a force directed normal to the surface of the sail, as shown in **Figure 1.2**.

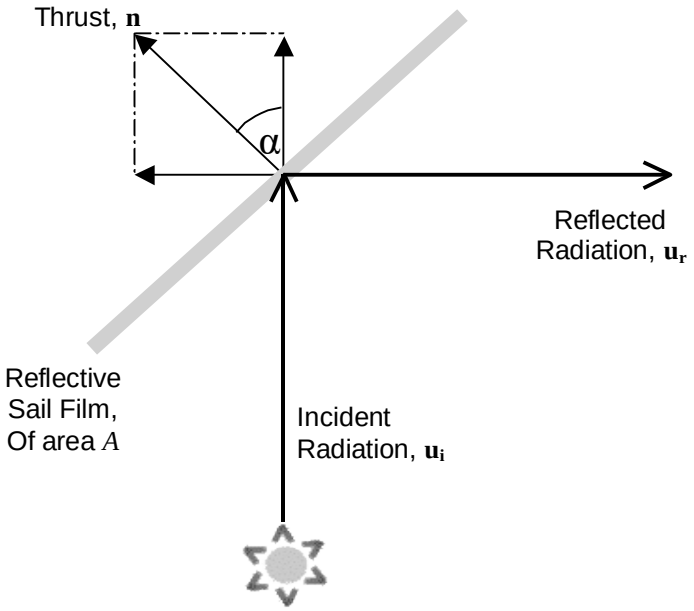

Figure 1.2 Incident and reflected forces on ideal solar sail

The orientation of the solar sail, and so the force vector, is described relative to the Sun-line by the sail pitch angle, α, and clock angle, δ. The orientation of the sail pitch and clock angles, the sail control angles, is defined in **Figure 1.3**. We note that the sail control angles can be defined as either;

$$0^o \leq \alpha \geq 90^o \quad \& \quad 0^o \leq \delta \geq 360^o$$

$$\text{or,} \quad -90^o \leq \alpha \geq 90^o \quad \& \quad 0^o \leq \delta \geq 180^o$$

however, each is equally correct and equally widely used within literature.

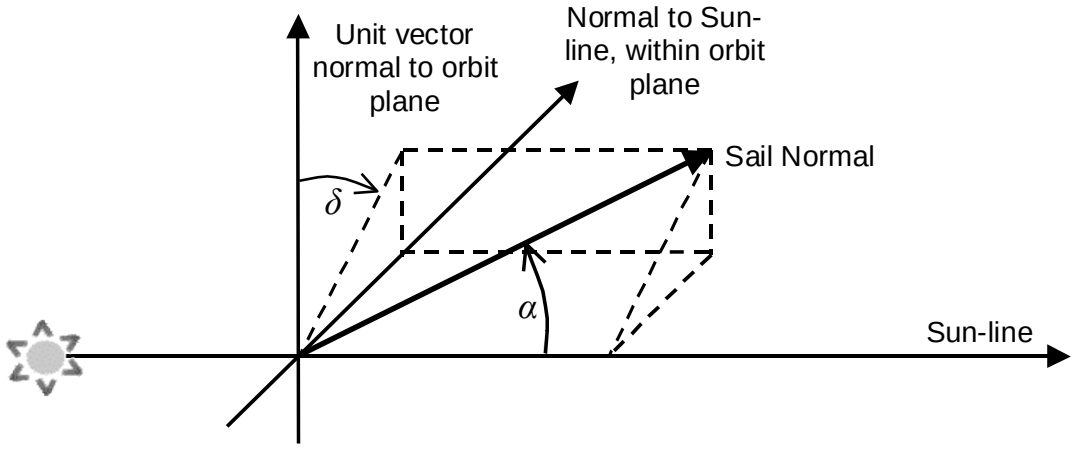

Figure 1.3 Orientation of the sail pitch and clock angles

By altering the orientation of the solar sail relative to the incoming photons, the solar radiation

pressure force vector can in principle be directed to any orientation within 90° of the Sun-line. As the pitch angle increases however, the magnitude of the solar radiation pressure force decreases due to the reduction in projected sail area and the reduction in the component of the solar radiation pressure force directed normal to the sail surface.

Solar sails are continuously accelerated so that their orbits are quite different from the usual ballistic arcs followed by conventional spacecraft. By choosing either a negative or positive sail pitch angle, with respect to the Sun-line, the solar sail will either spiral inwards towards the Sun, or outwards away from the Sun. Such 'tacking' provides solar sails with the agility to enable a wide range of mission applications. The rate at which the solar sail will spiral is of course a function of its characteristic acceleration. By an appropriate sequence of sail orientations, essentially any point in the solar system can be reached. However, for missions to rendezvous with planets, comets or asteroids, the sail orientation must be continuously altered along its trajectory to ensure that the solar sail matches the velocity of the target body when it is intercepted. Optimisation methods can be used in such instances to determine the best sail orientation time history to minimise the transfer time to the target body.

1.5 Further Reading

1.5.1 Historical interest

- Maxwell, J. C., Electricity and Magnetism, Oxford University Press, 1873
- Lebedew, P., 'The Physical Causes of Deviation from Newton's Law of Gravitation', Astrophysical Journal, **10**, 155-161, 1902
- Tsiolkovsky, K.E., Extension of Man into Outer Space, 1921 [also, Tsiolkovsky, K.E., Symposium Jet Propulsion, No. 2, United Scientific and Technical Presses, 1936]
- Tsander, K., From a Scientific Heritage, NASA Technical Translation TTF-541, 1967 [quoting a 1924 report by the author]
- Wiley, C., [pseudonym Sanders, R.] 'Clipper Ships of Space', Astounding Science Fiction, p. 135, May 1951
- Garwin, R.L., 'Solar Sailing - A Practical Method of Propulsion Within the Solar System', Jet Propulsion, **28**, 188-190, 1958
- Clarke, A.C., 'The Wind from the Sun', (first published 1963), see for example, Project Solar Sail, ed. Clarke, A.C., Penguin Group Books, New York, 1990
- Friedman, L., 'Solar Sailing - The Concept Made Realistic', AIAA-78-82, 16th AIAA Aerospace Sciences Meeting, Huntsville, January 1978
- Drexler, K.E., 'High Performance Solar Sails and Related Reflecting Devices', AIAA-79-1418, 4th Princeton/AIAA Conference on Space Manufacturing Facilities, Princeton, May 1979
- Logsdon, J.M., 'Missing Halley's Comet: The Politics of Big Science', ISIS, **80**, 302, 254-280,

1989

- Perret, A., LaBombard, E. & Miura, K., 'The Solar Sail Race to the Moon', IAF-89-539, 40th International Astronautical Federation Congress, Malaga, October 1989

1.5.2 Selected introductory papers

- Tsu, T.C., 'Interplanetary Travel by Solar Sail', American Rocket Society Journal, **29**, 422-427, 1959
- London, H.S., 'Some Exact Solutions of the Equations of Motion of a Solar Sail With a Constant Setting', American Rocket Society Journal, **30**, 198-200, 1960
- Sands, N., 'Escape from Planetary Gravitational Fields by use of Solar Sails', American Rocket Society Journal, **31**, 527-531, 1961
- Fimple, W.R., 'Generalised Three-Dimensional Trajectory Analysis of Planetary Escape by Solar Sail', American Rocket Society Journal, **32**, 883-887, 1962
- Wright, J.L. & Warmke, J.M., 'Solar Sail Mission Applications' AIAA-76-808, AIAA/AAS Astrodynamics Conference, San Diego, August 1976
- Van der Ha, J.C. and Modi, V.J., 'Long-term Evaluation of Three-Dimensional Heliocentric Solar Sail Trajectories with Arbitrary Fixed Sail Setting', Celestial Mechanics, **19**, 113-138, 1979
- Forward, R.L., 'Light-levitated Geostationary Cylindrical Orbits', Journal of Astronautical Sciences, **29**, 1, 73-80, 1981
- Svitek, T. et. al., 'Solar Sail Concept Study', IAF-ST-82-12, 33rd International Astronautical Congress, Paris, October 1982
- Staehle, R.L., 'An Expedition to Mars Employing Shuttle-Era Systems, Solar Sail and Aerocapture', Journal of the British Interplanetary Society, **35**, 327-335, 1982
- McInnes, C.R., 'Solar Sail Halo Trajectories: Dynamics and Applications', IAF-91-334, 42nd International Astronautical Congress, Montreal, October 1991
- Forward, R.L., 'Statite: A Spacecraft That Does Not Orbit', Journal of Spacecraft and Rockets, **28**, 5, 606-611, 1991
- McInnes, C.R. McDonald, A.J.C. Simmons, J.F.L. & MacDonald, E.W., 'Solar Sail Parking in Restricted Three-Body Systems', Journal of Guidance Dynamics and Control, **17**, 2, 399-406, 1994
- Maccone, C., 'Space Missions Outside the Solar System to Exploit the Gravitational Lens of the Sun', Journal of the British Interplanetary Society, **47**, 45-52, 1994
- Leipold, M., Borg, E., Lingner, S., Pabsch, A., Sachs, R. & Seboldt, W., 'Mercury Orbiter with a Solar Sail Spacecraft', Acta Astronautica, **35**, Suppl. 635-644, 1995
- Prado, J.Y., Perret, A., Pignolet, G. & Dandouras, I., 'Using a Solar Sail for a Plasma Storm Early Warning System', IAA-96-IAA.3.3.06, 47th International Astronautical Congress, October 1996
- Leipold, M., 'ODISSEE - A Proposal for Demonstration of a Solar Sail in Earth Orbit', IAA-

L98-1005, 3rd International Academy of Astronautics Conference on Low Cost Planetary Missions, Pasadena, April 1998

- McInnes, C.R., 'Mission Applications for High Performance Solar Sails', IAA-L.98-1006, 3rd IAA International Conference on Low-Cost Planetary Missions, Pasadena, April 1998
- Gershman, R. & Seybold, C., 'Propulsion Trades for Space Science Missions', IAA-L.98-1001, 3rd IAA International Conference on Low-Cost Planetary Missions, Pasadena, April 1998
- Carroll, K.A., 'Economical Planetary Space Science Using Small Solar Sail Propelled Spacecraft', Proceedings of the 10th Conference on Astronautics, CASI (Canadian Aeronautics and Space Institute), Ottawa, October 1998

1.5.2 Solar sailing books

- Polyakhova, E., Space Flight Using a Solar Sail -The Problems and the Prospects, Kosmicheskiy Polet Solnechnym Parusom, Moscow, 1986, [in Russian]
- Friedman, L., Star Sailing: Solar Sails and Interstellar Travel, Wiley Science Publications, New York, 1988
- Clark, A.C., [ed.] Project Solar Sail, Penguin Group Books, New York, 1990
- Wright, J.L., Space Sailing, Gordon and Breach Science Publishers, Philadelphia, 1992
- Souza, D.M., Space Sailing, Lerner Publications Company, Minneapolis, 1994
- M^{c}Innes, C.R., Solar Sailing: Technology, Dynamics and Mission Applications, Springer-Praxis, Chichester, 1999

2. Solar Radiation Pressure

The observation that light exerts a force on matter contradicts our everyday experiences. However, it is a common mechanism within the solar system. A prominent example of this is the tail of a comet. Comets have two distinct tails, an ion tail swept by the solar wind and a dust tail swept by solar radiation pressure. Interplanetary dust is also affected by solar radiation pressure. The Poynting-Robertson effect is a process whereby dust grains experience a transverse drag as well as radial light pressure. This is due to the relativistic aberration of light as the dust grains orbit the Sun. The resulting drag then causes dust to very slowly spiral inwards towards the Sun. Under certain conditions when the grains spiral close to the Sun, they begin to evaporate so reducing the ratio of their mass to cross-sectional area. The effect of solar radiation pressure then greatly increases, sometimes to the extent that light pressure can exceed solar gravity, thereby ejecting dust into interstellar space.

In 1619 Johannes Kepler using corpuscular theory, the then dominant theory of optics, proposed that comet tails are pushed outwards from the Sun due to sunlight as was a natural consequence of the theory. It is of interest and somewhat ironic to note that this early theory is qualitatively the same as the current view of solar radiation pressure, namely photons. Newton, accepting that Kepler's view was a possible explanation in 1687 attempted to explain the phenomenon solely within his theory of universal gravitation. Advancing the hypothesis that there was some ambient ether denser than the material of comet tails. The observed repulsion was thus merely due to buoyancy forces, with the Sun exerting only an attractive, gravitational force. In 1744 Euler returned to Kepler's original view, however Euler adopted the longitudinal wave theory of light due to Huygens. With this theory Euler was able to show that a longitudinal wave would exert a repulsive force on a body in its path.

In 1754 de Marian and du Fay made the first attempts to experimentally verify and measure radiation pressure. However, due to the effect of residual air currents in their apparatus these investigations proved inconclusive. It was until the beginning of the 20th century, with the superior experimental apparatus of the day that radiation pressure was finally characterised in the laboratory.

The current theoretical basis for the existence of radiation pressure came independent of the astronomical theories. The Scottish physicist James Clerk Maxwell predicted the existence of radiation pressure in 1873 as a consequence of his unified theory of electromagnetic radiation. Apparently independently of Maxwell, in 1876 Bartoli demonstrated the existence of radiation pressure as a consequence of the second law of thermodynamics. Also in 1873, Crookes mistakenly believed that he had demonstrated the existence of radiation pressure using his newly devised radiometer. Even today this device is occasionally used as a flawed demonstration of radiation pressure. Crookes radiometer consists of four vanes each of which is blackened on one side and silvered on the other.

These are attached to the arms of a rotor which is balanced on a vertical support in such a way that it can turn with very little friction. The mechanism is encased inside a clear glass bulb that has been pumped out to a high, but not perfect vacuum. When light is shown on the radiometer (often called a light-mill) the vanes turn with the black surfaces apparently being pushed away by the light. Crookes at first believed this demonstrated that light radiation pressure on the black vanes was turning it round just like water in a water mill. His paper reporting the device was refereed by James Clerk Maxwell, who accepted the explanation gave by Crookes. Maxwell, it appears, was delighted to see a demonstration of the effect of radiation pressure as predicted by his theory of electromagnetism. However, light falling on the black side should be absorbed, while light falling on the silver side of the vanes should be reflected. The net result is that there is twice as much radiation pressure on the metal side as on the black. In that case the mill is turning the wrong way. When this anomaly was realised other explanations for the radiometer effect were sought and some of the ones that people came up with are still mistakenly quoted as the correct one. It was clear that the black side would absorb heat from infrared radiation more than the silver side. This would cause the rarefied gas to be heated on the black side. The obvious explanation in that case, is that the pressure of the gas on the darker size increases with its temperature creating a higher force on that side of the vane. This force would push the rotor round. Maxwell analysed this theory carefully, presumably being wary about making a second mistake, discovering that in fact the warmer gas would simply expand in such a way that there would be no net force from this effect, just a steady flow of heat across the vanes. As a variation on this theme, it is sometimes said that the motion of the hot molecules on the black side of the vane provides the push. Again this is not correct and could only work if the mean free path between molecular collisions were as large as the container, but in fact it is typically less than a millimetre. Osborne Reynolds, better remembered for the “Reynolds number”, provided the correct solution to the problem. Early in 1879 Reynolds submitted a paper to the Royal Society in which he considered what he called "*thermal transpiration*", and also discussed the theory of the radiometer. By thermal transpiration Reynolds meant the flow of gas through porous plates caused by a temperature difference on the two sides of the plates. If the gas is initially at the same pressure on the two sides, there is a flow of gas from the colder to the hotter side, resulting in a higher pressure on the hotter side if the plates cannot move. Equilibrium is reached when the ratio of pressures on either side is the square root of the ratio of absolute temperatures. This is a counterintuitive effect due to tangential forces between the gas molecules and the sides of the narrow pores in the plates. The effect of these thermomolecular forces is very similar to the thermomechanical effects of superfluid liquid helium. The liquid, which lacks all viscosity, will climb the sides of its container towards a warmer region. If a thin capillary is dipped into the superfluid it flows up the tube at such speed that a fountain effect is produced at the other end. The vanes of a radiometer are not however porous. To explain the radiometer, therefore, one must focus attention not on the faces of the vanes, but on their edges. The faster molecules from the warmer side strike the

edges obliquely and impart a higher force than the colder molecules. Again these are the same thermomolecular forces that are responsible for thermal transpiration. The effect is also known as thermal creep since it causes gases to creep along a surface where there is a temperature gradient. The net movement of the vane due to the tangential forces around the edges is away from the warmer gas and towards the cooler gas with the gas passing round the edge in the opposite direction. The behaviour is just as if there were a greater force on the blackened side of the vane (which as Maxwell showed is not the case), but the explanation must be in terms of what happens not at the faces of the vanes but near their edges. Maxwell refereed Reynolds's paper, and so became aware of Reynolds's suggestion. Maxwell at once made a detailed mathematical analysis of the problem, and submitted his paper, "*On stresses in rarefied gases arising from inequalities of temperature*", for publication in the Philosophical Transactions; it appeared in 1879, shortly before his death. The paper gave due credit to Reynolds's suggestion that the effect is at the edges of the vanes, but criticised Reynolds's mathematical treatment. Reynolds's paper had not yet appeared (it was published in 1881), and Reynolds was incensed by the fact that Maxwell's paper had not only appeared first, but had criticised his unpublished work! Reynolds wanted his protest to be published by the Royal Society, but after Maxwell's death this was thought to be inappropriate.

The true experimental verification of the existence of radiation pressure and the verification of Maxwell's quantitative results came in 1900. At the University of Moscow, Peter Lebedew succeeded in isolating radiation pressure using a series of torsion balance experiments. Nichols and Hull at Dartmouth College, New Hampshire, also obtained independent verification in 1901.

2.1 Radiation Pressure – Electromagnetic Description

Within the electromagnetic description of light momentum is transported by electromagnetic waves. The electric field component of the wave, **E**, induces a current, **j**, in the sail, as shown in **Figure 2.1**.

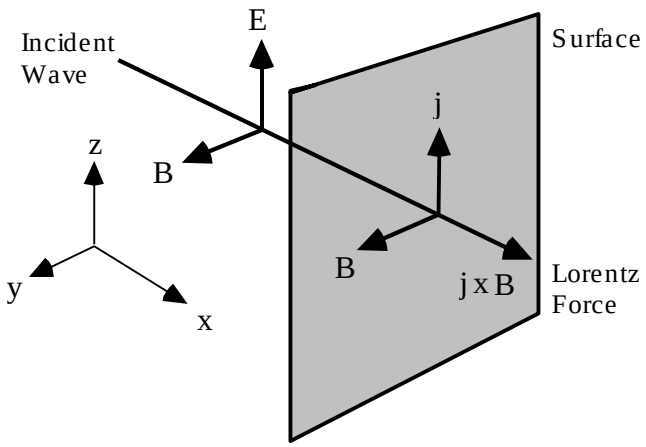

Figure 2.1 Electromagnetic radiation pressure

The magnetic component of the incident wave, **B**, generates a Lorentz force **j x B** in the direction of propagation of the wave. The induced current generates an electromagnetic wave, which is observed as the reflection of the incident wave. For a wave propagating along the x-axis the force exerted on a current element is then given by

$$df = j_z B_y \, dx \, dy \, dz \qquad [2.1]$$

where, j_Z is the current density induced in the surface of the reflector. The resulting pressure on the current element, defined as the force per unit area, can then be written as

$$dP = j_z B_y \, dx \qquad [2.2]$$

From Maxwell's equations of electrodynamics the current term in [2.2] can be replaced by field terms. Following this substitution it can be demonstrated that the time average pressure is then given by

$$\langle dP \rangle = -\frac{\partial}{\partial x}\left[\frac{1}{2}\varepsilon_o {E_z}^2 + \frac{1}{2\mu_o}{B_y}^2\right] dx \qquad [2.3]$$

The term in brackets is identified as the energy density, U, for the electric component, E, and magnetic component, B, of the incident wave, defined as

$$U = \frac{1}{2}\varepsilon_o E^2 + \frac{1}{2\mu_o}B^2 \qquad [2.4]$$

where, ε_O is the permittivity of free space and m_O is the permeability of free space. The pressure exerted on a surface of thickness Δl is obtained by integrating [2.3], as

$$\langle P \rangle = -\int_0^{\Delta l} \frac{\partial U}{\partial x} \, dx \qquad [2.5]$$

For a perfectly absorbing medium the pressure exerted on the surface is then given by the total energy density of the electromagnetic wave.

$$\langle P \rangle = \langle U \rangle \qquad [2.6]$$

Consider two plane waves, separated by a distance Δx and incident on a surface of area A, as shown in **Figure 2.2**. The volume of space between the two waves impinging on the surface is then $A\Delta x$. The spacing Δx between the waves is equivalent to $c\Delta t$, where Δt is the time of travel between the wave fronts. The energy density of the electromagnetic wave is thus

$$U = \frac{\Delta E}{A(c\Delta t)} \qquad [2.7]$$

where, ΔE is the energy contained within the volume element. In addition, the energy flux W across the surface can be written as

$$W = \frac{1}{A}\left[\frac{\Delta E}{\Delta t}\right] \qquad [2.8]$$

Therefore, it can be seen that

$$U = \frac{W}{c} \qquad [2.9]$$

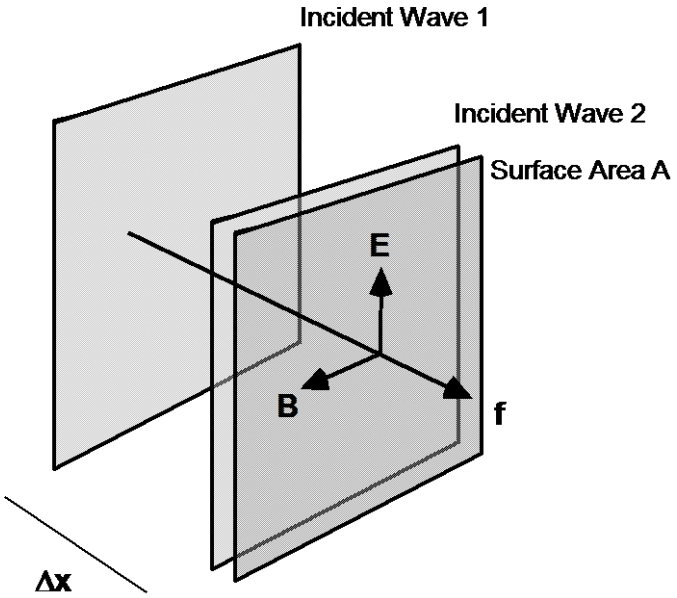

Figure 2.2 Energy density of an electromagnetic wave

This is the same expression as the radiation pressure derived from the quantum description of light.* It is therefore concluded that the quantum and electromagnetic description of radiation pressure are equivalent. For the electromagnetic description of light, the radiation pressure is the energy density of the electromagnetic wave, as stated by Maxwell in his 1873 work on Electricity and Magnetism – "*hence in a medium in which waves are propagated there is a pressure in the direction normal to the waves and numerically equal to the energy in unit volume*".

For a perfectly reflecting surface the observed pressure is twice the value provided by [2.9] due to the momentum transferred to the surface by the incident photons and the reaction provided by reflected photons. Using [2.9] the solar radiation pressure exerted on a solar sail at the Earth's distance from the Sun (1 AU) can be calculated. Since the orbit of the Earth about the Sun is slightly elliptical the energy flux received at the Earth varies by approximately 3.5% during the year. An accepted mean value of the solar constant W_E is 1368 J s^{-1} m^{-2}. Thus, the pressure exerted on a perfectly reflecting solar sail at 1 AU is taken to be 9.12 x 10^{-6} N m^{-2}.

2.2 Force on a Perfectly Reflecting Solar Sail

A solar sail is an oriented surface such that the acceleration experienced by the surface is a function of the surfaces attitude. For a solar sail, as shown in **Figure 1.2**, of area A with a unit

* See, McInnes, C.R., Solar Sailing: Technology, Dynamics and Mission Applications, Springer-Praxis, Chichester, 1999

vector **n** directed normal to surface, the force exerted on the surface due to incident photons from direction $\mathbf{u}_i$ is given by

$$\mathbf{f}_i = PA(\mathbf{u}_i.\mathbf{n})\mathbf{u}_i \qquad [2.10]$$

where, $A(\mathbf{u}_i.\mathbf{n})$ is the projected surface area in direction $\mathbf{u}_i$, as shown in **Figure 1.2**. Similarly, the reflected photons will exert a force of equal magnitude on the surface, but in the specular reflected direction $-\mathbf{u}_r$.

$$\mathbf{f}_r = -PA(\mathbf{u}_i.\mathbf{n})\mathbf{u}_r \qquad [2.11]$$

Using the vector identity $\mathbf{u}_i - \mathbf{u}_r = 2(\mathbf{u}_i.\mathbf{n})\mathbf{n}$, the total force **f** exerted on the solar sail is therefore given by

$$\mathbf{f} = 2PA(\mathbf{u}_i.\mathbf{n})^2\mathbf{n} \qquad [2.12]$$

In order to calculate the pressure exerted on a body, the momentum transported by a flux of photons must also be considered. The energy flux W (the energy crossing unit area in unit time) at distance r from the Sun may be written in terms of the solar luminosity L_S and scaled by the Sun-Earth distance R_E as

$$W = W_E\left[\frac{R_E}{r}\right]^2 \qquad [2.13]$$

$$W_E = \frac{L_S}{4\pi {R_E}^2} \qquad [2.14]$$

where, W_E is the energy flux measured at the Earth's distance from the Sun. Thus, the total force may be written as

$$\mathbf{f} = \frac{2AW_E}{c}\left[\frac{R_E}{r}\right]^2(\mathbf{u}_i.\mathbf{n})^2\mathbf{n} \qquad [2.15]$$

The solar sail performance is parameterised by the total spacecraft mass per unit area ($^m/_A$). This constant is termed the sail loading σ and is a key design parameter. In addition, the sail pitch angle α will is defined as the angle between the sail normal and the incident radiation, as shown in **Figures 1.2 & 1.3**. Using these definitions the solar sail acceleration may now be written as

$$\mathbf{a} = \frac{2W_E}{c}\frac{1}{\sigma}\left[\frac{R_E}{r}\right]^2\cos^2\alpha\,\mathbf{n} \qquad [2.16]$$

The characteristic acceleration of the solar sail a_0 is again defined as the acceleration experienced at 1 AU, with the sail normal to the Sun such that $\alpha = 0$. The characteristic acceleration is an equivalent design parameter to the solar sail loading and may be conveniently written as

$$a_o = \frac{9.12\eta}{\sigma\left[g\,m^{-2}\right]} \quad \left[mm\,s^{-2}\right] \qquad [2.17]$$

where, η is again some overall efficiency of the solar sail used to account for the finite reflectivity of the sail film. Typically the total solar sail efficiency is of order 0.8 - 0.9.

For a solar sail in a heliocentric orbit the direction of incidence of the radiation $\mathbf{u}_i$ is defined by

the unit radial vector $\hat{\mathbf{r}}$ from the Sun to the solar sail. Furthermore, the solar sail acceleration may also be written in terms of the solar gravitational acceleration as

$$\mathbf{a} = \beta \frac{GM_s}{r^2} \left(\hat{\mathbf{r}}.\mathbf{n}\right)^2 \mathbf{n} \qquad [2.18]$$

where, M_S is the solar mass and G is the universal gravitational constant. The dimensionless sail parameter β will now be defined as the ratio of the solar radiation pressure acceleration to the solar gravitational acceleration. This parameter is normally referred to as the sail lightness number. As both the solar radiation pressure acceleration and the solar gravitational acceleration are assumed to have an inverse square variation, the lightness number is independent of the Sun-sail distance. Using [2.14], [2.16] and [2.18] the solar sail lightness number may be written as

$$\beta = \frac{\sigma^*}{\sigma} \qquad [2.19]$$

$$\sigma^* = \frac{L_S}{2\pi GM_S c} \qquad [2.20]$$

The critical solar sail loading parameter, σ^*, is found to be 1.53 gm^{-2}. This is a unique constant that is a function of the solar mass and the solar luminosity. With this mass per unit area the solar sail lightness number is one, such that the solar radiation pressure acceleration is exactly equal to the solar gravitational acceleration.

A more rigorous examination of the effect of radiation pressure on a surface can be found through the use of radiative transfer methods, as performed in the textbook by M^{c}Innes, 1999.

2.3 Radiation Pressure due to a Finite Solar Disc

While the variation of the solar radiation pressure with solar distance can be approximated by an inverse square variation relationship, this assumption breaks down at low solar radius, when the finite angular size of the solar disc must be considered. The modelling of the source of radiation pressure is distinct and independent from the modelling of solar radiation pressure force, which is dependent in the optical properties of the sail surface.

Initially we assume the solar disc has uniform brightness, such that an element of the solar disc will appear equally bright when viewed from any aspect angle. Thus, the specific intensity is time independent and isotropic across the solar disc and the solar radiation pressure exerted on a radially oriented, perfectly reflecting sail at a heliocentric distance r can be written as

$$P(r) = \frac{2}{c} \int_o^\infty \int_o^{2\pi} \int_o^{\theta_o} I_\nu \cos^2\theta \; \mathrm{d}\Omega \mathrm{d}\nu \quad , \quad \mathrm{d}\Omega = \sin\theta \; \mathrm{d}\theta \, \mathrm{d}\phi \qquad [2.21]$$

where, the angular radius of the solar disc θ_0 is given by $\sin^{-1}(R_S / r)$, as shown in **Figure 2.3**.

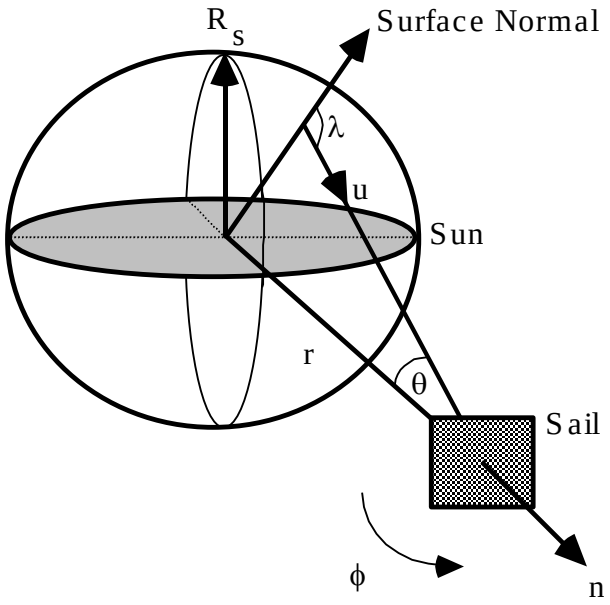

Figure 2.3 Solar radiation pressure due to a finite solar disc

Noting geometric symmetry about the azimuth and that the specific intensity is independent of r, [2.21] reduces to

$$P(r) = \frac{4\pi}{c} I_o \int_{\xi_0}^{1} \xi^2 d\xi \quad , \xi = \cos\theta \quad , \xi_o = \cos\theta_o \qquad [2.22]$$

where, I_O is the frequency integrated specific intensity. Performing this integration and substituting for x_O it is found that

$$P(r) = \frac{4\pi}{3c} I_o \left[1 - \left[1 - \left(\frac{R_S}{r} \right)^2 \right]^{3/2} \right] \qquad [2.23]$$

Which may be expanded in powers of $(R_S / r)^2$ and for $r \gg R_S$ as, to the first order,

$$P(r) = \frac{2\pi}{c} I_o \left[\frac{R_S}{r} \right]^2 + O(R_S / r)^4 \qquad [2.24]$$

However, at large values of r this expansion must match asymptotically with the expression for the radiation pressure from a distant point source.

$$P^*(r) = \frac{2}{c} \left[\frac{L_S}{4\pi {R_S}^2} \right] \left(\frac{R_S}{r} \right)^2 \qquad [2.25]$$

Thus by comparing [2.24] and [2.25] the frequency integrated specific intensity is identified as

$$I_o = \frac{L_S}{4\pi^2 {R_S}^2} \qquad [2.26]$$

Substituting for I_O in [2.23] an expression for the solar radiation pressure exerted on a radially oriented solar sail from a uniformly bright, finite angular sized solar disc is obtained as

$$P(r) = \frac{L_S}{3\pi c {R_S}^2}\left[1-\left[1-\left(\frac{R_S}{r}\right)^2\right]^{3/2}\right] \qquad [2.27]$$

A more useful way of expressing this is in terms of the usual inverse square law $P^*(r)$.

$$P(r) = P(r)^* F(r) \qquad [2.28]$$

$$F(r) = \frac{2}{3}\left(\frac{r}{R_S}\right)^2\left[1-\left[1-\left(\frac{R_S}{r}\right)^2\right]^{3/2}\right] \qquad [2.29]$$

The function $F(r)$ describes the deviation of the true solar radiation pressure from an inverse square law. $F(r)$ attains a minimum value at $r = R_S$, where $F(R_S) = {}^2/_3$, giving the greatest deviation of the solar radiation pressure from an inverse square law. As $r \to \infty$, $F(r) \to 1$ since the solar disc becomes more point-like. From **Figure 2.4** it is seen that $F(r)$ approaches 1 over a scale of order 10 solar radii (0.047 AU) so that the magnitude of the deviation from an inverse square form is small at large heliocentric distances.

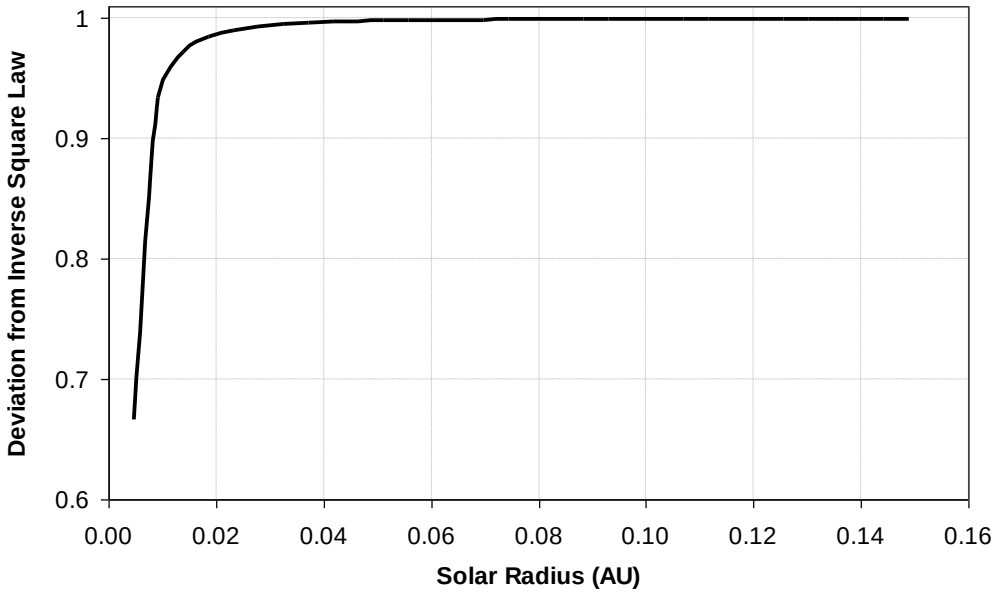

Figure 2.4 Deviation of the uniformly bright finite disc model from an inverse square model

Physically, the deviation from an inverse square law is due to photons from the solar limb intercepting the sail at an oblique angle to the sail surface, whereas photons from the centre of the disc are incident along the normal to the sail. The photons from the solar limb therefore transfer a smaller amount of momentum to the solar sail than those from the centre of the disc. At large heliocentric distances, however, photons from all parts of the solar disc are incident along near-parallel rays.

A more accurate model of the solar radiation pressure may be obtained by including solar

limb darkening in the functional form of the specific intensity rather than assuming a uniformly bright disc. Limb darkening is an effect due to the specific intensity of the solar radiation field having a directional dependence. Empirically, solar limb darkening has a complex functional form. However using an approximate model of the solar atmosphere an analytic expression for the limb darkening effect is obtained by McInnes, 1999. For the limb darkened specific intensity it is found that $F(R_S)$ = 0.708 so that the limb darkened solar radiation pressure deviates less from an inverse square form than the non-limb darkened pressure.

2.4 Solar Sail Force Models & Orbit Perturbations

The assumption that a solar sail is an ideal reflector is clearly not suitable for realistic trajectory design and mission analysis. The sail force can be parameterised through consideration of optics theory, considering sail reflection, absorption and re-radiation, however there are several assumptions made within this model, including the sail is perfectly flat. To account for these assumptions a numerical parametric force model was developed by JPL for the Halley rendezvous mission.* The JPL model is illustrated in **Figure 2.5**, where we see that at high pitch angles the force vector collapses towards zero. That is to say that the force vector cannot be directed at over 55.5° from the Sun-line, at sail pitch angle 72.6°. The sail force vector angle from the Sun-line is defined as the sail cone angle. As pitch is increased above 72.6° the cone angle falls back towards zero, along with the sail force magnitude.

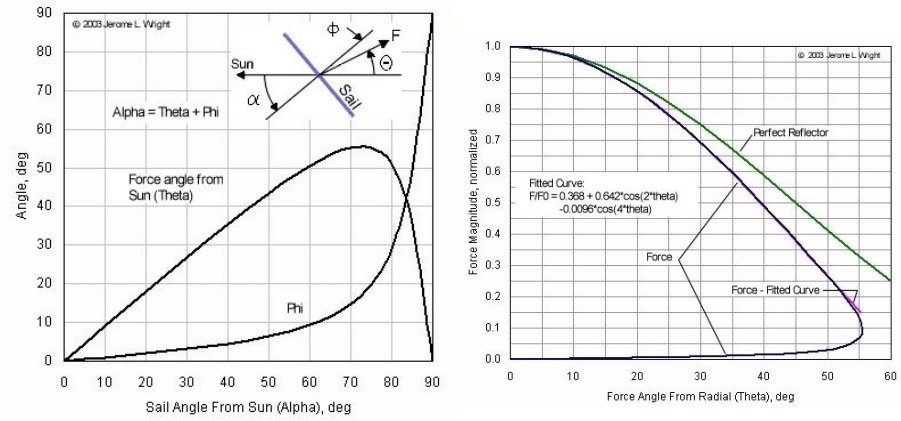

Figure 2.5 JPL numerical parametric sail force model (Images by Wright, 2003)

Both the theoretical and measured sail forces are discussed in some detail in the textbook by McInnes, 1999. We note however that both models apply standard optical theory, where non-specular reflections at > 6 – 10° are assumed to be of no use. A solar sail can however utilise all the reflected photons, no matter what angle they reflect at, as they will still exert a force on the sail surface. Recent work has been performed which shows the assumptions made within

* Wright, J.L., Space Sailing, Gordon and Breach Science Publishers, Philadelphia, 1992

both the traditional models significantly compromises the results, as reflection is highly symmetric about the specular line, that is to say there is no collapse in the force vector.*

Solar radiation pressure generates the largest force on a solar sail. However, secondary forces are also present. For example, the solar wind will exert a small pressure on the sail due to the momentum transported by solar wind protons. During periods of high solar wind speed the mean proton number density ρ at the Earth rises to approximately 4 x 10^6 m^{-3} with a wind speed v_w of order 700 km s^{-1}. The solar wind pressure P_W exerted on the solar sail can then be estimated from the momentum transported as

$$P_w \sim m_p \rho v_w^{\,2} \qquad [2.30]$$

where, m_p is the mass of a proton. Using the above parameters a solar wind pressure of order 3 x 10^{-9} N m^{-2} is obtained, which is nearly 10^{-4} less than the direct solar radiation pressure exerted on the sail at 1 AU. In addition, first order relativistic effects are proportional to the ratio of the solar sail speed to the speed of light, typically of order 10^{-4}. For solar sails in Earth orbit the secondary pressure due to radiation scattered from the Earth is also small, being at least three orders of magnitude less than that due to the direct solar radiation pressure. A significant perturbation in Earth orbit is residual atmospheric drag experienced at low altitudes. This perturbation is a strong function of solar activity, which is modulated in an 11-year cycle. Earth's residual atmosphere influences all satellites up to 500 – 1000 km. At periods of low solar activity the atmospheric drag and sail acceleration balance at ~ 430 km, as shown in **Figure 2.6**. At mean solar activity the balance points raises to ~ 560 km and on up to ~ 940 km at high activity periods. A safe mean altitude is > 800 km, although at times of solar maxima this may double. At Mars the minimum solar sail altitude is ~ 300 km, while at Venus it is ~ 900 km.

We note from **Figure 2.6** at no time can a solar sail deliver a capsule, such as a sample return capsule, to the orbit of the ISS for later retrieval. Instead the capsule must be delivered to a high altitude then recovered from there either through its own propulsion system or by a recovery craft.

Other perturbations of note in Earth orbit are lunar and solar gravity, as well as shadow events, which eliminate essential Sunlight and thus the thrust capability of a solar sail.

* Rogan et al, 2001, "Encounter 2001: Sailing to the Stars", SSC01-112, 15th Annual/USU Conference On Small Satellites

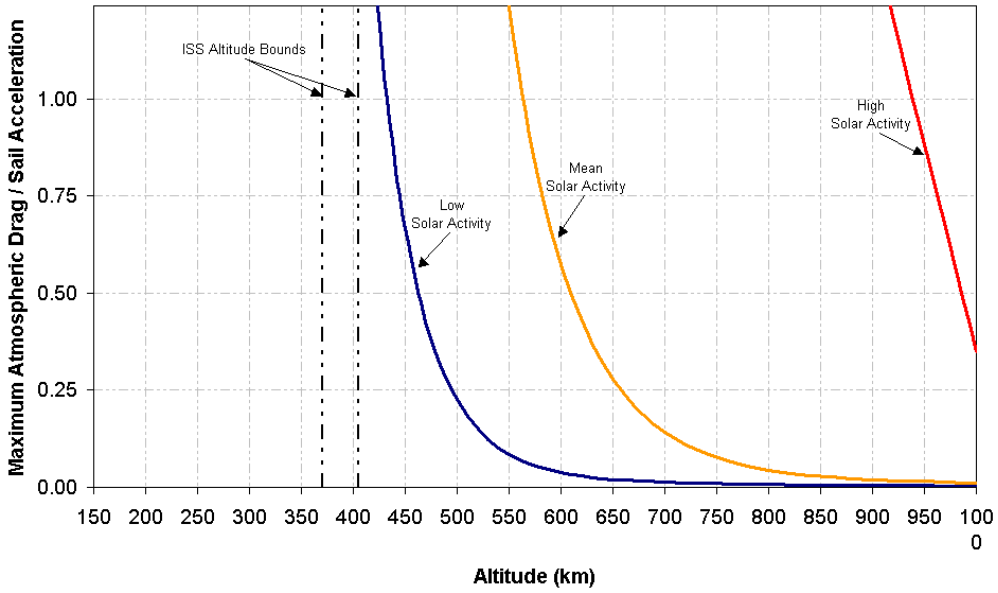

Figure 2.6 Effect of Earth's atmosphere on a solar sail at a range of altitudes

2.5 Further Reading

2.5.1 Historical interest

- Lebedew, P., 'The Physical Causes of Deviation from Newton's Law of Gravitation', Astrophysical Journal, **10**, 155-161, 1902
- Maxwell, J. C., Electricity and Magnetism, Oxford University Press, 1873
- Nichols, E.F. & Hull, G.F., 'The Experimental Investigation of the Pressure of Light', Astrophysical Journal, **15**, 60-65, 1902
- Nichols, E.F. & Hull, G.F., 'The Pressure due to Radiation', Physical Review, **17**, 260-50, 1903

2.5.2 Radiative transfer

- Mihalas, D. & Mihalas, B.W., Foundations of Radiation Hydrodynamics, Oxford University Press, 1984
- McInnes, C.R. & Brown, J.C., 'The Dynamics of Solar Sails with a Non-point Source of Radiation Pressure', Celestial Mechanics and Dynamical Astronomy, **9**, 249-264, 1990

2.5.3 Solar sail force model

- Georgevic, R. M., 'The Solar Radiation Pressure Force and Torques Model', Journal of Astronautical Sciences, **20**, 257-274, 1973
- Wright, J.L., *Space Sailing*, Gordon and Breach Science Publishers, 1992, (Appendix A, B)
- Rogan et al, 2001, "Encounter 2001: Sailing to the Stars", SSC01-112, 15th Annual/USU Conference On Small Satellites

3. Solar Sail Orbital Dynamics

3.1 Heliocentric Orbital Dynamics

In the first instance, the sun-centred orbital dynamics of solar sail spacecraft will be described, since this is traditionally where solar sailing is applied and where the bulk of the mission concepts exist, due to the large Δv requirements. The vector equation of motion of a solar sail spacecraft moving in a heliocentric orbit is defined by Eq. 3.1, where ***r*** is the position vector of the spacecraft with respect to the Sun, at time *t*. The gravitational parameter of the Sun is defined by μ. For an ideal sail, the thrust vector is aligned along the sail normal direction, **n**.

$$\frac{d^2\mathbf{r}}{dt^2} + \frac{\mu}{r^2}\hat{\mathbf{r}} = \beta\frac{\mu}{r^2}\left(\hat{\mathbf{r}}.\mathbf{n}\right)^2\mathbf{n} \qquad [3.1]$$

For the simple analysis to follow, it is appropriate to show how this equation is represented in two-dimensional polar coordinates, shown in Eq. 3.2 and 3.3. Here, θ is the azimuth angle from the x-axis, *r* is the distance of the spacecraft from the Sun. The pitch or cone angle, α is defined as the angle between the sail normal and the radial vector.

$$\frac{d^2r}{dt^2} - r\left[\frac{d\theta}{dt}\right]^2 = -\frac{\mu}{r^2} + \beta\frac{\mu}{r^2}\cos^3\alpha \qquad [3.2]$$

$$r\frac{d^2\theta}{dt^2} + 2r\left[\frac{dr}{dt}\right]\left[\frac{d\theta}{dt}\right] = \beta\frac{\mu}{r^2}\cos\alpha^2\sin\alpha \qquad [3.3]$$

The parameter, β is known as the sail lightness number and is defined as the ratio of the local solar radiation pressure acceleration produced by the sail to the local solar gravitational acceleration. This number is a useful parameter since it is independent of the solar distance and defines the performance of the sail. Another useful performance parameter is the sail characteristic acceleration, a_c, which is the solar radiation pressure induced acceleration of an ideal sail, pitched face on to the Sun ($\alpha = 0$) at 1 AU from the Sun. It is easy to calculate that a characteristic acceleration of 1.0 mm s^{-2} corresponds to a dimensionless sail lightness number of 0.1686. As has been discussed the solar sail characteristic acceleration is dependent on the surface area and reflectivity of the sail assembly and the mass of the spacecraft.

3.2 Conic Section Orbits

When the sail pitch angle is fixed at zero degrees, with the sail thrust vector always aligned

along the Sun-line, then families of conic section orbits can be produced. These are not Keplerian orbits in the strictest sense, since the solar gravity is effectively reduced by the solar radiation pressure force acting on the spacecraft in a directly opposing direction. They are effectively conic section orbits with a modified gravitational parameter, $\tilde{\mu} = \mu(1-\beta)$, since solar gravity and solar radiation pressure both vary as the inverse square of the solar distance. The orbit equation can then be written as Eq. 3.4, where *h* is the orbital angular momentum per unit mass. Similar equations can be derived for the other orbital parameters, with the modified gravitational parameter.

$$r = \frac{h^2/\tilde{\mu}}{1 + e\cos f} \qquad [3.4]$$

Figure 3.1 shows the family of these orbits for a range of sail lightness numbers. For a lightness number of zero, so with no solar sail, the base orbit is circular and Keplerian. When the solar sail is deployed and pitched to a Sun facing attitude then the modified orbit is elliptical for lightness numbers between zero and 0.5, with an increasing eccentricity for higher increasing sail performance. With a lightness number of exactly 0.5 then there is a transition from an elliptical orbit to a hyperbolic orbit through the parabolic orbit, which shows the lightness number necessary for escape. When the lightness number increases to between 0.5 and unity we have a hyperbolic orbit, which can again be defined using equations used for hyperbolic Keplerian orbits, but with the modified gravitational parameter, $\tilde{\mu}$. When the lightness number is exactly unity, then we have the interesting situation where the solar gravity is completely cancelled out. This could enable rectilinear orbits or solar levitators, providing the anti-Sunward pointing sail thrust is maintained, via zero pitch control. With extremely high performance sails exhibiting lightness numbers of greater than unity, the Sun now becomes placed at the opposite focus of the conic section. In this regard, we have the exotic situation whereby solar radiation pressure now becomes the primary force, with the solar gravity acting as a perturbation force.

3.3 Logarithmic Spiral Trajectory

When the solar sail thrust is orientated at a fixed pitch angle to the Sun-line, with a low sail performance, we obtain a logarithmic spiral trajectory. The radial component of the sail thrust reduces the effective gravitational force on the sail as for zero pitch, however the component of thrust in the transverse direction acts to increase (or decrease) the orbital angular momentum of the orbit.

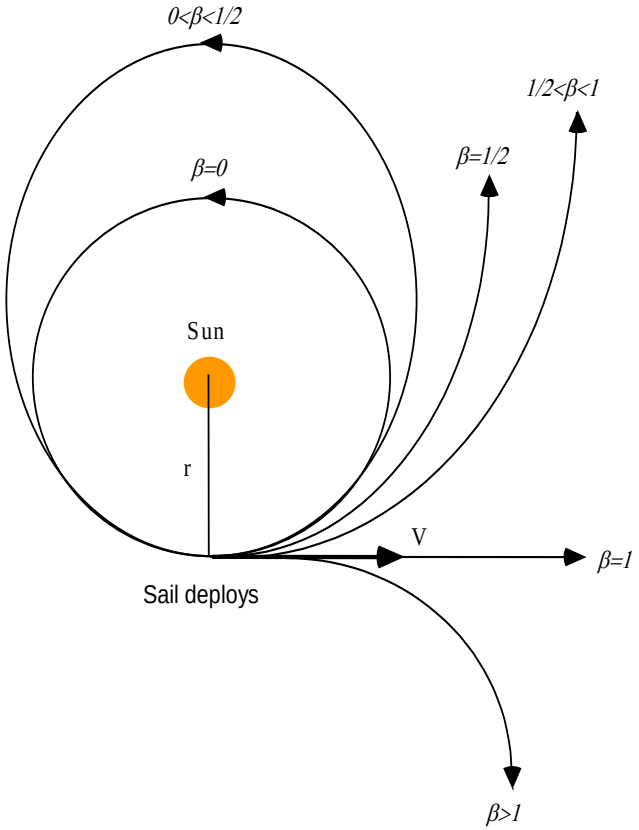

Figure 3.1 Conic section orbits with zero pitch

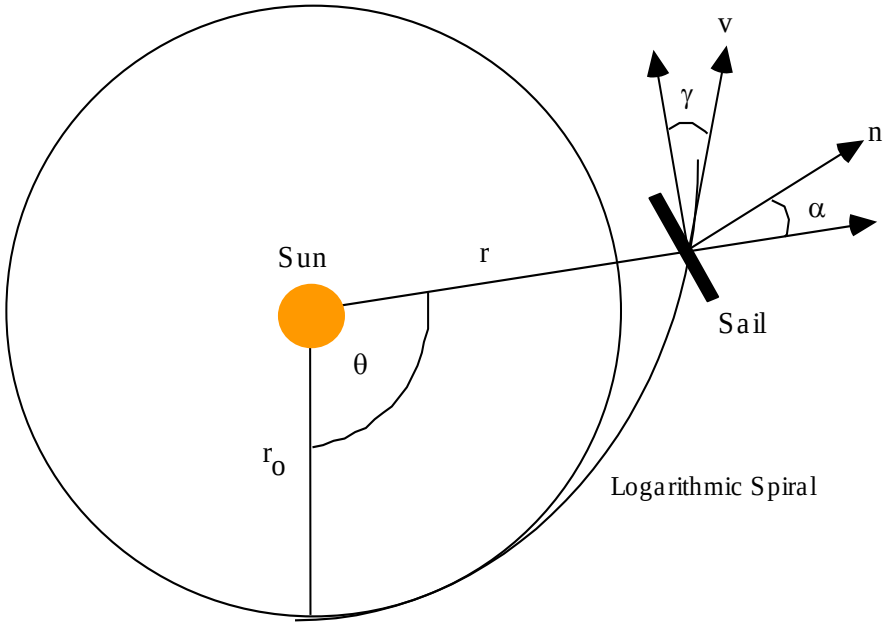

Figure 3.2 Logarithmic spiral trajectory

It should be reiterated here that the solar sail can spiral away from the Sun by directing a component of thrust along the velocity vector in order to increase the orbital angular momentum and so spiral away from the Sun. Similarly, negative pitch angle can direct a component of the thrust against the velocity vector, so decreasing the orbital angular momentum as so spiral towards the Sun.

For fixed sail pitch angle in a two-dimensional planar case, a particular solution to the polar equations of motion in Eqs. 3.2 and 3.3 can be found to Eq. 3.5.

$$r(\theta) = r_o \exp(\theta \tan\gamma) \qquad [3.5]$$

Where r_o is the initial distance from the Sun, and γ is the spiral angle, the angle between the transverse direction and the velocity. For a specific logarithmic spiral, the spiral angle will be fixed, as shown in **Figure 3.2**. If we differentiate the solution and substitute, we can obtain the velocity magnitude in Eq. 3.6, where it can be seen that the local solar sail speed is always less than the local circular orbit speed.

$$v(r) = \sqrt{\frac{\mu}{r}}\left[1 - \beta\cos^2\alpha(\cos\alpha - \sin\alpha\tan\gamma)\right]^{1/2} \qquad [3.6]$$

This means that coplanar transfer between two circular orbits cannot be achieved without applying a hyperbolic excess at launch to place the solar sail onto the logarithmic spiral, and circularising the orbit on arrival at the final circular orbit. These discontinuities pose problems in the practical application of logarithmic spirals to orbit transfers. However, the logarithmic spiral often provides a good first guess for a more advanced control law that can deal with this two-point boundary value problem. By combining the radial and tranverse components of velocity, and substituting these into the equations of motion, it can be shown that an implicit relationship exists between the sail pitch angle and spiral angle, as in Eq. 3.7, the result of this is depicted in **Figure 3.3**. For the lower branch of solutions, it is observed that there appears to be an optimum sail pitch angle that maximises the spiral angle for each sail lightness number. This leads into the definition of the optimum pitch angle derived from optimum control theory, to be defined later.

$$\frac{\sin\gamma\cos\gamma}{2 - \sin^2\gamma} = \frac{\beta\cos^2\alpha\sin\alpha}{1 - \beta\cos^3\alpha} \qquad [3.7]$$

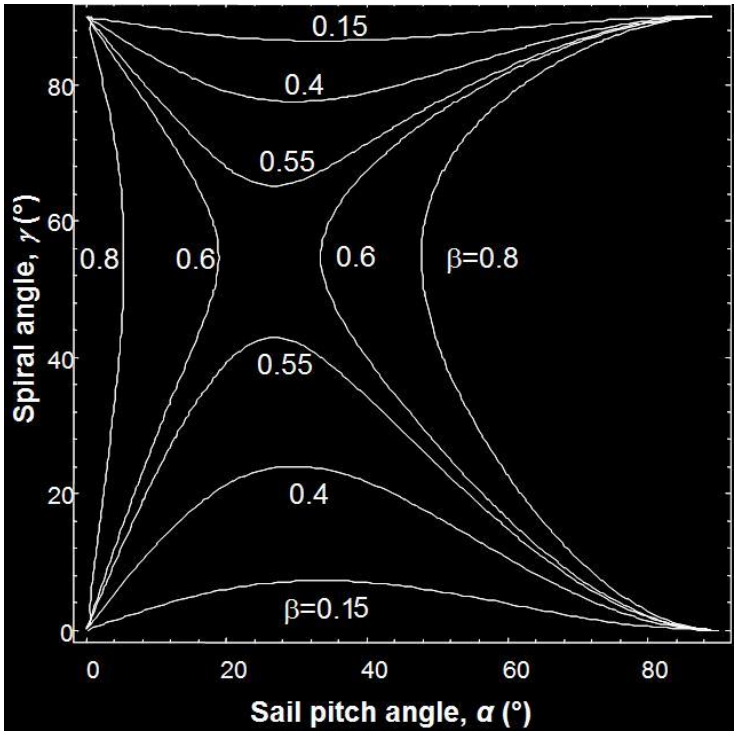

Figure 3.3 Implicit relationship between sail pitch angle and logarithmic spiral angle, for contours of sail lightness number [McInnes, 1999]

3.4 Optimal Sail Force Vector

As has been described previously, the sail thrust vector is defined by the cone and clock angles in the radial-transverse-normal (RTN) frame, as shown in **Figure 3.4**. The thrust vector (sail normal) is defined by Eq. 3.8.

$$\mathbf{n} = \cos\alpha\,\hat{\mathbf{r}} + \sin\alpha\cos\delta\,\hat{\mathbf{p}} + \sin\alpha\sin\delta\,\hat{\mathbf{p}}\times\hat{\mathbf{r}} \qquad [3.8]$$

In order to optimise the sail control angles, we define a required direction, **q** along which the component of the sail thrust is to be maximised, shown in Eq. 3.9.

$$\mathbf{q} = \cos\tilde{\alpha}\,\hat{\mathbf{r}} + \sin\tilde{\alpha}\cos\tilde{\delta}\,\hat{\mathbf{p}} + \sin\tilde{\alpha}\sin\tilde{\delta}\,\hat{\mathbf{p}}\times\hat{\mathbf{r}} \qquad [3.9]$$

The force in this required direction is given by Eq. 3.10 and 3.11. If we differentiate Eq. 3.11 with respect to the cone angle and find the turning points, $\partial f_q / \partial\alpha = 0$, then we can obtain the optimal sail cone angle, which maximises the force in the required direction, shown in Eq. 3.12.

$$f_q = 2PA(\mathbf{n}.\hat{\mathbf{r}})^2(\mathbf{n}.\mathbf{q}) \quad [3.10]$$

$$f_q = 2PA\cos^2\alpha\left[\cos\alpha\cos\tilde{\alpha} + \sin\alpha\sin\tilde{\alpha}\cos\left(\delta - \tilde{\delta}\right)\right] \quad [3.11]$$

$$\tan\alpha^* = \frac{-3+\sqrt{9+8\tan^2\tilde{\alpha}}}{4\tan\tilde{\alpha}} \quad [3.12]$$

Figure 3.5 depicts the process of optimisation of the sail cone angle. It is found that the optimal sail cone angle lags behind the required cone angle due to the reduction in the total force magnitude as the cone angle is increased, which varies as $\cos^2\alpha$.

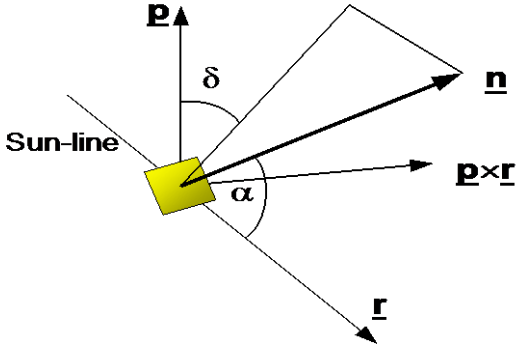

Figure 3.4 Definition of sail control angles

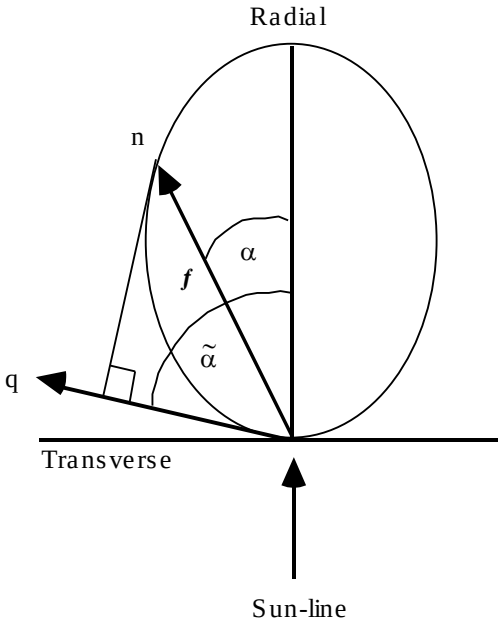

Figure 3.5 Optimisation of the sail cone angle [McInnes, 1999]

It is easy to show that the sail cone angle which maximises the force in the transverse direction, so that $\tilde{\alpha} = 90^{\circ}$, is given by $tan\,\alpha^{*} = 1/\sqrt{2}$. This optimal angle is 35.26°, which enables the maximum increase in angular momentum. It is observed in **Figure 3.3**, that the pitch angle which maximises the logarithmic spiral angle, for low sail lightness numbers, is also close to this optimal angle.

3.5 Lagrange Variational Equations

The cartesian or polar equations of motion can alternatively be represented in terms of the Lagrange Variational Equations, shown in Eqs. 3.13-18, where *p* is the semi-latus rectum, and *n* is the orbital mean motion. As will be shown, these equations are particularly useful when we want to maximise the rate of change of a particular orbital element. This is useful when we want to modify one orbital element, while leaving the other time-averaged elements unchanged.

$$\frac{da}{df} = \frac{2pr^2}{\mu\left(1-e^2\right)^2}\left[Se\sin f + T\frac{p}{r}\right] \qquad [3.13]$$

$$\frac{de}{df} = \frac{r^2}{\mu}\left[S\sin f + T\left(1+\frac{r}{p}\right)\cos f + T\frac{r}{p}e\right] \qquad [3.14]$$

$$\frac{di}{df} = \frac{r^3}{\mu p}\cos(f+\omega)W \qquad [3.15]$$

$$\frac{d\Omega}{df} = \frac{r^3}{\mu p\sin i}\sin(f+\omega)W \qquad [3.16]$$

$$\frac{d\omega}{df} = -\frac{d\Omega}{df}\cos i + \frac{r^2}{\mu e}\left[-R\cos f + T\left(1+\frac{r}{p}\right)\sin f\right] \qquad [3.17]$$

$$\frac{dt}{df} = \frac{r^2}{\sqrt{\mu p}}\left[1-\frac{r^2}{\mu e}\left[S\cos f - T\left(1+\frac{r}{p}\right)\sin f\right]\right] \qquad [3.18]$$

In the definition of the RTN frame, the radial (*S*), transverse (*T*), and normal (*W*) components of the solar sail thrust are defined by Eqs. 3.19, 3.20, and 3.21.

$$S = \beta\frac{\mu}{r^2}\cos\alpha^3 \qquad [3.19]$$

$$T = \beta \frac{\mu}{r^2} \cos\alpha^2 \sin\alpha \sin\delta \qquad [3.20]$$

$$W = \beta \frac{\mu}{r^2} \cos\alpha^2 \sin\alpha \cos\delta \qquad [3.21]$$

It is of interest to note that variational equations can be combined to find variational equations for other orbital elements, such as the radius of periapsis and apoapsis. For example, if the radius of apoapsis, $r_A = a(1+e)$ is differentiated, we obtain Eq. 3.22.

$$\frac{dr_A}{df} = \frac{da}{df}(1+e) + a\frac{de}{df} \qquad [3.22]$$

3.6 Locally Optimal Control Laws

We will now discuss how the variational equations can be used to formulate a set of control laws that maximise the rate of change of a particular orbital element. These control laws cannot guarantee global optimality, and they are often termed closed loop methods. Global optimality requires the use of numerical methods, and even then, the true optimum solution is hard to attain. A general formulation of the Lagrange variational equations is given by Eq. 3.23, where $\mathbf{f} = (S,T,W)$ is the solar radiation pressure thrust of the sail, and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ is a vector of functions of the solar sail orbital elements. In optimal control theory, this required force vector is often called the primer vector, and is illustrated in **Figure 3.6**.

$$\frac{dZ}{df} = \boldsymbol{\lambda}(Z).\,\mathbf{f} \qquad [3.23]$$

The force component along the primer vector is given by Eq. 3.24. With reference to **Figure 3.6**, for 2D planar trajectories, where $W = 0$, we obtain Eq. 3.25. In a similar procedure to **section 3.4**, if we maximise the thrust in the required direction, $\partial f_\lambda / \partial\alpha = 0$, then, we obtain the equivalent in Eq. 3.26.

$$f_\lambda = 2PA(\mathbf{n}.\hat{\mathbf{r}})^2\,\mathbf{n}.\boldsymbol{\lambda} \qquad [3.24]$$

$$\tan\tilde{\alpha} = \left[\frac{\lambda_2}{\lambda_1}\right] \qquad \tilde{\delta} = \frac{\pi}{2} \qquad [3.25]$$

$$\tan\alpha^{*} = \frac{-3+\sqrt{9+8\tan^{2}\tilde{\alpha}}}{4\tan\tilde{\alpha}} \qquad [3.26]$$

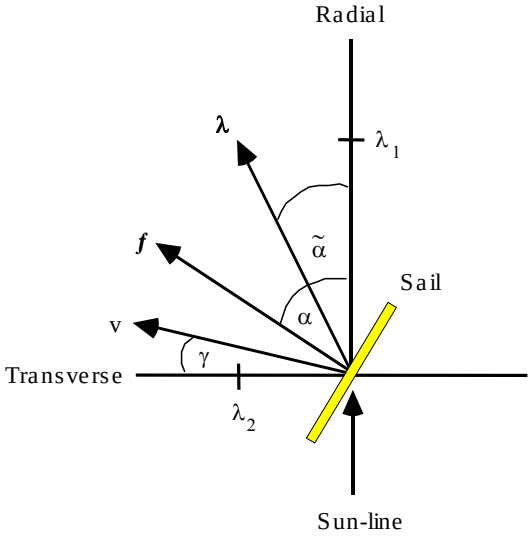

Figure 3.6 Required force vector for locally optimal trajectories

The components of the primer vector can be obtained by taking the variational equation of interest and, ignoring common scaling factors, the components of the primer vector are found using the dot product. Therefore, the primer vector components that maximise the first variational equation for the rate of change of semi-major axis is shown in Eq. 3.27 and 3.28. The required cone angle is then given by Eq. 3.29, through substitution of the orbit equation for *r*. The optimum cone angle is then found via Eq. 3.26.

$$\lambda_1 = e\sin f \qquad [3.27]$$

$$\lambda_2 = p/r \qquad [3.28]$$

$$\tan\tilde{\alpha} = \frac{1+e\cos f}{e\sin f} \qquad [3.29]$$

The effect of this control law that maximises the rate of change of semi-major axis is shown in **Figure 3.7**, which shows the actual sail pitch/cone angle profile necessary to increase the semi-major axis at a maximum rate.

To maximise the rate of change of the orbit eccentricity, the components of the primer vector are found to be Eqs. 3.30 and 3.31, which gives a required cone angle shown in Eq. 3.32. An example of this control law is shown in **Figure 3.8**. Equivalent control laws can be derived for the other planar orbital elements, such as argument of periapsis, and radius of

periaspsis and apoapsis. It has been found that the eccentricity control law also gives near-optimal changes in the radius of apoapsis.

$$\lambda_1 = \sin f \qquad [3.30]$$

$$\lambda_2 = \left[1 + \frac{r}{p}\right]\cos f + e\frac{r}{p} \qquad [3.31]$$

$$\tan\tilde{\alpha} = \frac{2 + e\cos f}{1 + e\cos f}\cot f + \frac{e\cos\,\mathrm{ec}f}{1 + e\cos f} \qquad [3.32]$$

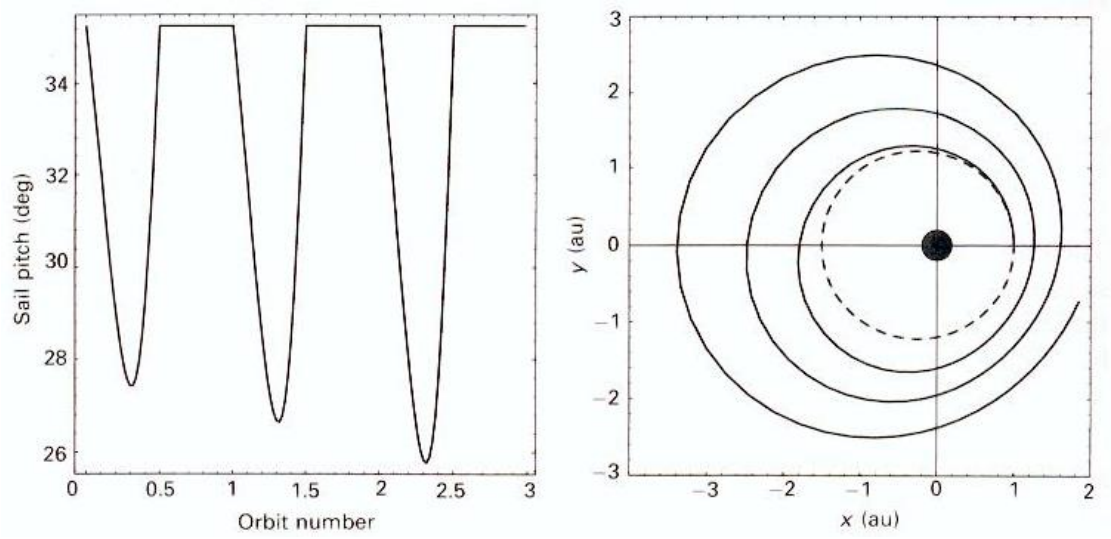

Figure 3.7 Maximising the rate of change of the semi-major axis [McInnes, 1999, constraint on maximum pitch angle]

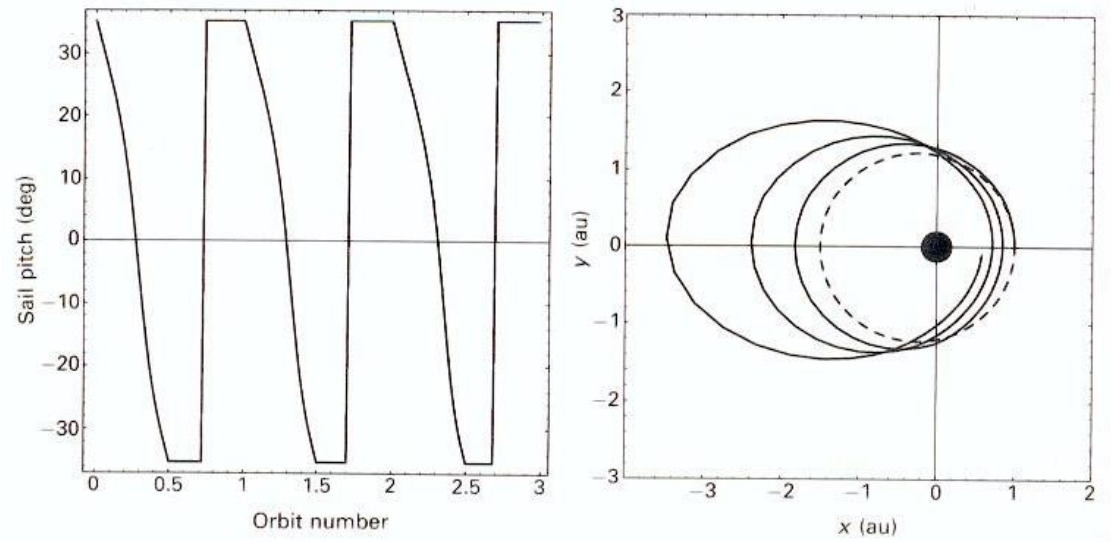

Figure 3.8 Maximising the rate of change of the eccentricity [McInnes, 1999, constraint on maximum pitch angle]

Changes in the out-of-plane orbital elements, such as inclination and right ascension of the ascending node, can be effected by the use of simple switching functions. For maximum rate of change of inclination, the solar sail thrust can be directed alternately above and below the orbit plane every half-orbit, as defined by Eq. 3.33. *sign* represents +1 or -1, depending on the sign of the function in square brackets.

$$\text{sign}W = \text{sign}\left[\cos(f+\omega)\right] \qquad [3.33]$$

The optimal sail cone angle that maximises the out-of-plane sail thrust is 35.26°, given by Eq. 3.34, for the sail clock angle control law in Eq. 3.35. The effect of this control law is generate a 'cranking' orbital manoeuvre, as depicted in **Figure 3.9**.

$$\alpha^* = \tan^{-1}\left[\frac{1}{\sqrt{2}}\right] \qquad [3.34]$$

$$\delta^* = \frac{\pi}{2}\left(1 - \text{sign}\left[\cos(f+\omega)\right]\right) \qquad [3.35]$$

If we substitute the control law into the inclination variational equation, Eq. 3.15, and integrate over one orbit, we can obtain the change in inclination per orbit. We find that the change in inclination per orbit is independent of orbit radius and only depends on the sail lightness number, shown in Eq. 3.36.

$$\Delta i = 88.2\beta \quad [degrees\ per\ orbit] \qquad [3.36]$$

However, closer orbits to the Sun have shorter orbit periods, and so the time to achieve an overall inclination change is shorter. In addition, the constant radial force component means that the orbit period will increase slightly during cranking.

A simple switching function can be used in a similar manner as above to increase the right ascension of ascending node angle, as in Eq. 3.37. This control law is implemented with a 90° phase shift to the inclination control law. The steering law for the cone and clock angle is shown in Eq. 3.38 and 3.39, respectively.

$$\mathrm{sign}W = \mathrm{sign}\left[\sin(f+\omega)\right] \qquad [3.37]$$

$$\alpha^* = \tan^{-1}\left[\frac{1}{\sqrt{2}}\right] \qquad [3.38]$$

$$\delta^* = \frac{\pi}{2}\left(1 - \mathrm{sign}\left[\sin(f+\omega)\right]\right) \qquad [3.39]$$

Again, by integrating the variational equation over one orbit, we find that the change in ascending node angle per orbit is independent of solar radius, as shown in Eq. 3.40. It is only dependent on the sail performance and orbit inclination. However, the overall change is faster at closer radii, due to the shorter orbit period, as for inclination changes above. The change per orbit is much larger at low inclinations, due to the definition of ascending node angle.

$$\Delta\Omega = \frac{88.2\beta}{\sin i} \quad \left[\textit{degrees per orbit}\ \right] \qquad [3.40]$$

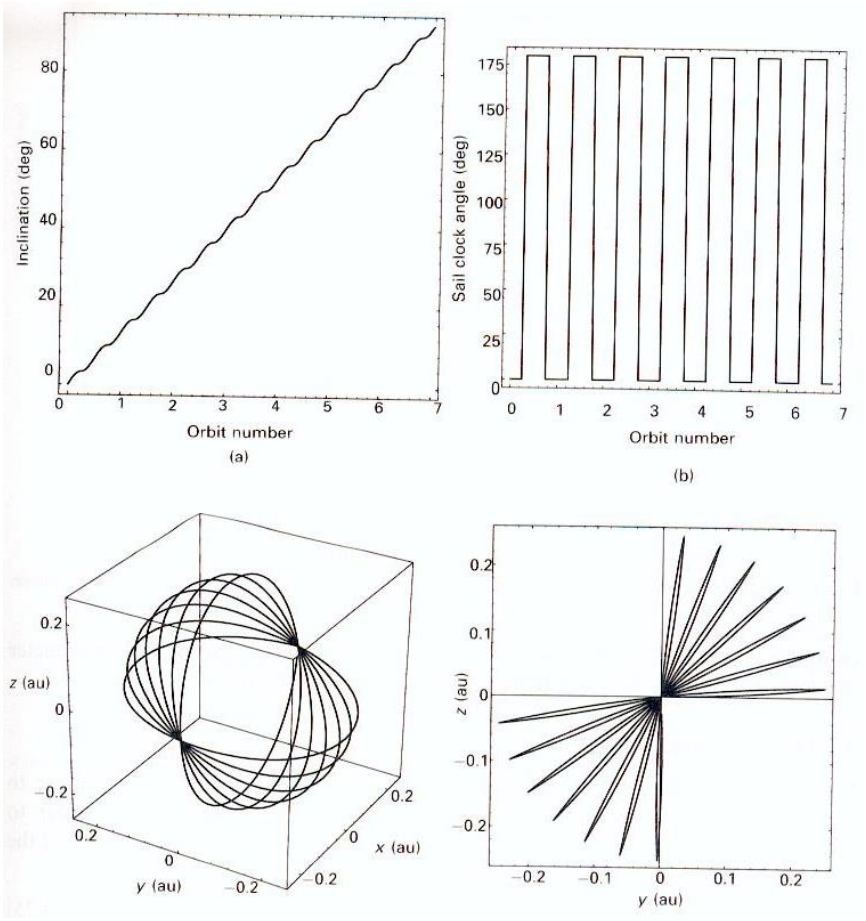

Figure 3.9 Cranking orbit produced by inclination control law [McInnes, 1999]

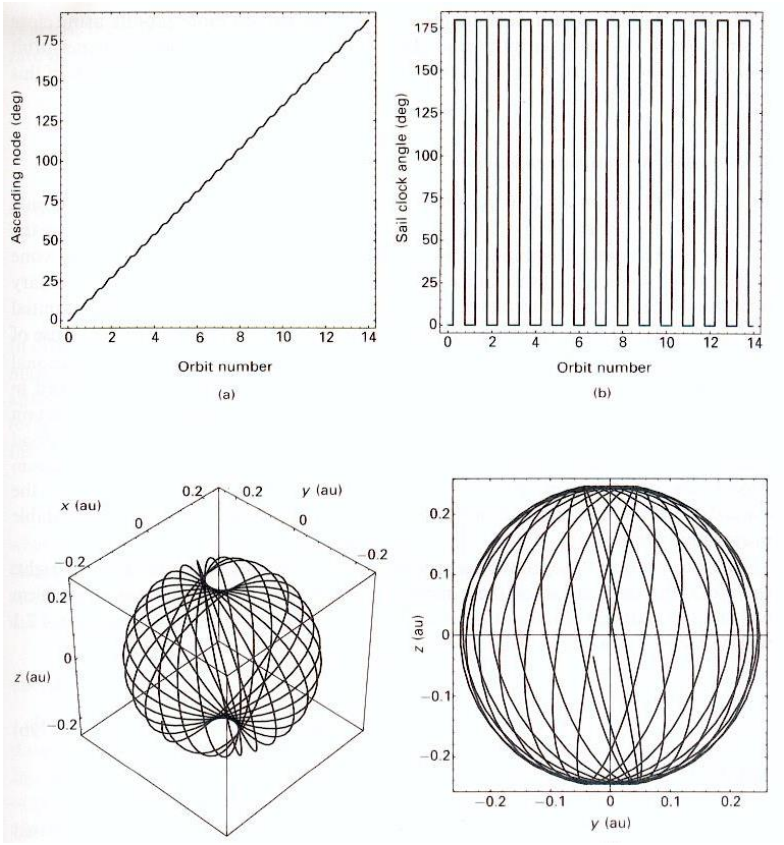

Figure 3.10 Rotation of line of nodes through right ascension of ascending node angle control law [McInnes, 1999]

3.7 True Minimum Time Trajectories

The control laws described can often provide near-optimal results, but they can never produce true-minimum time solar sail trajectories. The locally optimal control only optimises the instantaneous thrust vector in a closed loop fashion. To obtain true-optimal trajectories, which fully match the two-point boundary conditions for transfer and rendezvous, we must resort to numerical methods and optimal control theory. Since optimal control theory is a complex field, the method will be briefly outlined in this section. The reader is referred to the many books available on the subject.

Firstly, the vector equation of motion in Eq. 3.1 is re-formulated in terms of two first order differential equations, Eq. 3.41 and 3.42.

$$\dot{\mathbf{r}} = \mathbf{v} \qquad [3.41]$$

$$\dot{\mathbf{v}} = -\frac{\mu}{r^2}\hat{\mathbf{r}} + \beta\frac{\mu}{r^2}\left(\hat{\mathbf{r}}.\mathbf{n}\right)^2\mathbf{n} \qquad [3.42]$$

Next, we formulate a function called the Hamiltonian, from the equations of motion, shown in Eq. 3.43. Here, $\mathbf{p_r}$ and $\mathbf{p_v}$ are the momentum vectors of the co-states of position and velocity, respectively. The velocity co-state is also called the primer vector in optimal control theory. The rate of change of the co-state vectors can be obtained by differentiating the Hamiltonian function, as Eqs. 3.44 and 3.45, to give Eqs. 3.46 and 3.47.

$$H = \mathbf{p}_r.\mathbf{v} - \frac{\mu}{r^2}\mathbf{p}_v.\hat{\mathbf{r}} + \beta\frac{\mu}{r^2}\left(\hat{\mathbf{r}}.\mathbf{n}\right)^2\mathbf{p}_v.\mathbf{n} \qquad [3.43]$$

$$\dot{\mathbf{p}}_r = -\frac{\partial H}{\partial \mathbf{r}} \qquad [3.44]$$

$$\dot{\mathbf{p}}_v = -\frac{\partial H}{\partial \mathbf{v}} \qquad [3.45]$$

$$\dot{\mathbf{p}}_r = \frac{\mu}{r^3}\mathbf{p}_v - \frac{3\mu}{r^5}\left(\mathbf{p}_r.\mathbf{r}\right)\mathbf{r} + 2\beta\frac{\mu}{r^3}\left(\hat{\mathbf{r}}.\mathbf{n}\right)\left(\mathbf{p}_v.\mathbf{n}\right)\left(\mathbf{n} + 2\left(\hat{\mathbf{r}}.\mathbf{n}\right)\hat{\mathbf{r}}\right) \qquad [3.46]$$

$$\dot{\mathbf{p}}_v = -\mathbf{p}_r \qquad [3.47]$$

For a solar sail, minimisation of the transfer time is accomplished by maximising the Hamiltonian at all points in the trajectory, through an appropriate choice of the sail cone and clock angles. To maximise the Hamiltonian, the sail clock angle must be aligned along the primer vector, and the optimal sail clock angle is aligned such that the component of the sail force along the primer vector is maximised, as in **section 3.4**, via Eq. 3.48.

$$\tan\alpha^* = \frac{-3 + \sqrt{9 + 8\tan^2\tilde{\alpha}}}{4\tan\tilde{\alpha}} \qquad [3.48]$$

The co-states are required as a function of time. This can be achieved by numerical integration of Eqs. 3.46 and 3.47 along with the equations of motion. However, the boundary conditions of the co-states are not known, unlike the equations of motion. To ensure the optimality of the launch date, the transversality condition in Eq. 3.49 must also be satisfied.

$$\left[\beta\frac{\mu}{r^2}\left(\hat{\mathbf{r}}.\mathbf{n}\right)^2\mathbf{p}_v.\mathbf{n}\right]_{t=t_o} = \left[\beta\frac{\mu}{r^2}\left(\hat{\mathbf{r}}.\mathbf{n}\right)^2\mathbf{p}_v.\mathbf{n}\right]_{t=t_f} \qquad [3.49]$$

Iterative numerical methods have to be used to select the co-state boundary conditions that satisfy the orbit boundary conditions and the tranversality condition. Small changes in the solar sail control profile have a very small effect on the transfer time, such that convergence to the true-minimum time solution is often difficult. An initial guess of the co-states must be supplied to ensure convergence to the optimal solution, however these co-states are non-intuitive to the user and the problem is often highly sensitive to these. For this reason, this Calculus of Variations based method is classed as an Indirect method. It is possible that locally optimal control laws could be used to provide an initial guess for the iterative process. The Calculus of Variations forms the basis of the NASA/JPL VARITOP low-thrust trajectory optimisation tool, which has been used very successfully, but only by experts who have many years of experience of using the method. Indirect trajectory optimisation is something of a 'black-art'.

While optimal control theory can provide true-minimum time trajectories, it is often easier to parameterise the trajectory in an open-loop fashion. Then, gradient-based, non-linear programming methods are used to solve the constrained parameter optimisation, two-point boundary value problem by iteratively selecting a discretised cone and clock angle control history, that satisfies the boundary conditions and orbit constraints, whilst minimising the transfer time. Thus, these methods are termed Direct methods. As will be discussed in **section 3.10**, the cone and clock angles can be characterised by interpolation between a set of discrete points along the trajectory. As the number of optimised parameters is increased, then the control profile increasingly approximates the true-optimal continuous control profile of the Indirect method of optimal control theory. An example of a minimum-time trajectory to Mercury is shown in **Figure 3.11**, along with the optimal control profile in **Figure 3.12**. This trajectory was generated using a Direct parameter optimisation method using 50 control nodes, equally spaced in time and linearly interpolated between. This 2.8 year rendezvous spiral, with a solar sail characteristic acceleration of 0.25 mm s^{-2} (β = 0.042), is very close to the true-optimal trajectory, since the smooth oscillatory control profile closely approximates a continuous control profile, with adequate control resolution. The optimisation method is based on Sequential Quadratic Programming and is described in more detail in **section 3.10**.

3.8 Non-Keplerian Orbits

Solar sails do not require propellant mass, and so exhibit an effectively infinite specific impulse over long mission durations. The continuous thrust produced enables quite exotic Non-Keplerian Orbits (NKOs). One example of a family of NKOs will be described here, for a more detailed discussion the reader is referred to McInnes (1999, Chapter 5). The constant solar radiation pressure force can be used in two (and three) body systems to displace a circular (or, in principal, any) orbit above or below the Sun-centred plane, and artificially maintain a desired orbit period.

The dimensions and period of this Displaced NKO are defined by the sail lightness number and cone angle (the clock angle in this case is zero, for out-of-plane thrusting), and vice versa. **Figure 3.13** shows how this displaced NKO can be produced through balancing solar gravity with the centrifugal and solar radiation pressure force. In the rotating frame of reference, the equation of motion, is given by Eq. 3.50. The left hand side of the equation contains the kinematic, coriolis and centripetal acceleration terms, and the right hand side contains the solar sail acceleration and gravitational acceleration, defined by Eq. 3.51 and the gravitational potential in Eq. 3.52.

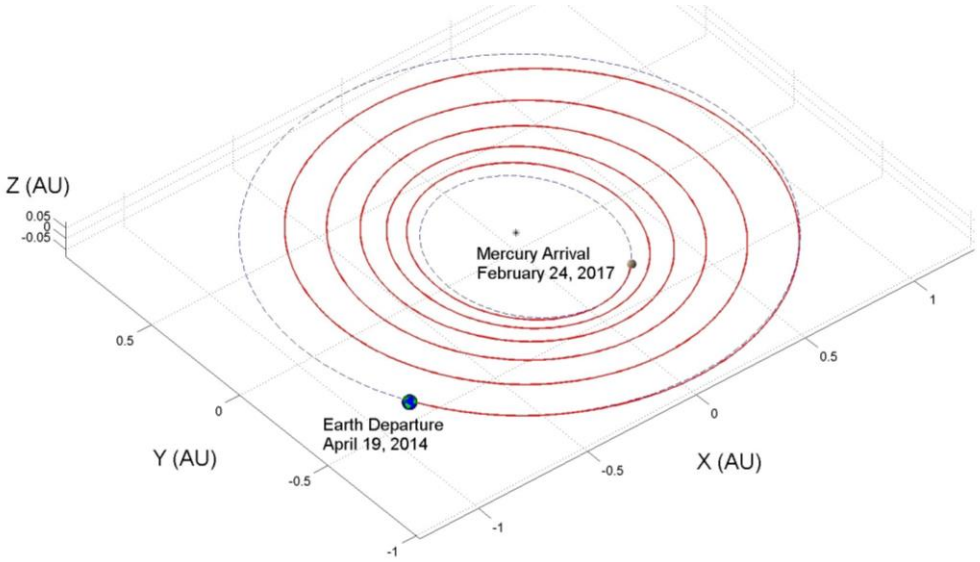

Figure 3.11 Minimum-time rendezvous trajectory to Mercury, generated using a Direct parameter optimisation method (**section 3.10**) [Hughes et al, 2003 and 2004]

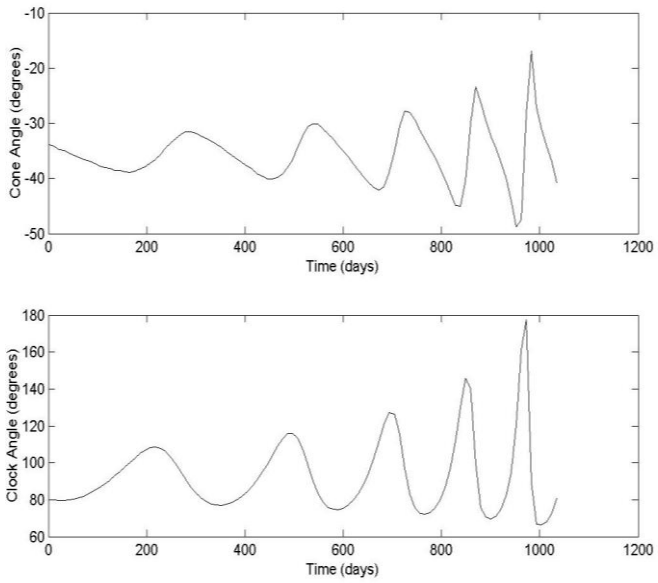

Figure 3.12 Minimum-time Mercury rendezvous control profile (50 parameterised control nodes with linear interpolation between each discrete node) [Hughes et al, 2003 and 2004]

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\omega} \times \frac{d\mathbf{r}}{dt} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) = \mathbf{a} - \nabla V \qquad [3.50]$$

$$\mathbf{a} = \beta \frac{\mu}{r^2} (\hat{\mathbf{r}}.\mathbf{n})^2 \mathbf{n} \qquad [3.51]$$

$$V = -\frac{\mu}{r} \qquad [3.52]$$

We can define a modified potential, $U=V+\Phi$, by noting that the centripetal acceleration term is conservative, such that it is defined by a scalar potential in Eq. 3.53, and therefore the centripetal acceleration term becomes Eq. 3.54. The equation of motion in the rotating frame now becomes the reduced form as Eq. 3.55.

$$\Phi = -\frac{1}{2}|\boldsymbol{\omega} \times \mathbf{r}|^2 \qquad [3.53]$$

$$\nabla\Phi = \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \qquad [3.54]$$

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\omega} \times \frac{d\mathbf{r}}{dt} + \nabla U = \beta \frac{\mu}{r^2} (\hat{\mathbf{r}}.\mathbf{n})^2 \mathbf{n} \qquad [3.55]$$

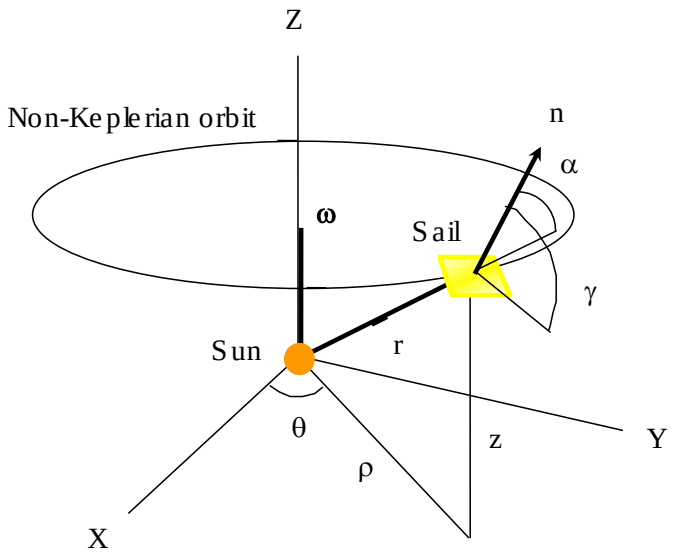

Figure 3.13 Sun-centred Non-Keplerian Orbit (inertial frame)

For an equilibrium solution to this equation, we take the vector product of Eq. 3.55 with **n**, to obtain Eq. 3.56, where ε is an arbitrary scalar multiplier. Since **n** is a unit vector of magnitude unity then Eq. 3.56 results in a definition of **n** for equilibrium in Eq. 3.57.

$$\nabla U \times \mathbf{n} = 0 \quad \Rightarrow \quad \mathbf{n} = \varepsilon \nabla U \qquad [3.56]$$

$$\mathbf{n} = \frac{\nabla U}{|\nabla U|} \qquad [3.57]$$

For a circular displaced NKO, we require only uniform azimuthal motion, such that there is no component of **n** in the azimuthal direction (clock angle δ = 0). **n** is then only in the plane defined by $\hat{\mathbf{r}}$ and $\boldsymbol{\omega}$, so we take the vector and scalar products of Eq. 3.57 with $\hat{\mathbf{r}}$, to obtain the required sail cone angle in Eq. 3.58. We can also do the same with **n** to obtain the required sail lightness number in Eq. 3.59.

$$\tan\alpha = \frac{|\hat{\mathbf{r}} \times \nabla U|}{\hat{\mathbf{r}}.\nabla U} \qquad [3.58]$$

$$\beta = \frac{r^2}{\mu} \frac{\nabla U.\mathbf{n}}{(\hat{\mathbf{r}}.\mathbf{n})^2} \qquad [3.59]$$

In the cylindrical coordinate system of **Figure 3.13**, the modified potential is given by Eq. 3.60. Then, evaluating the potential gradient and using Eqs. 3.58 and 3.59, we obtain, Eq. 3.61 and 3.62 in the cylindrical coordinate system, where ω is the required orbital angular velocity of the displaced NKO, and $\tilde{\omega}$ is the angular velocity of a circular orbit of radius *r*.

$$U = -\left[\frac{1}{2}\rho^2\omega^2 + \frac{\mu}{r}\right] \qquad [3.60]$$

$$\tan\alpha = \frac{(z/\rho)(\omega/\tilde{\omega})^2}{(z/\rho)^2 + \left[1 - (\omega/\tilde{\omega})^2\right]} \quad , \quad \tilde{\omega}^2 = \frac{\mu}{r^3} \qquad [3.61]$$

$$\beta = \left[1 + \left(\frac{z}{\rho}\right)^2\right]^{1/2} \frac{\left[(z/\rho)^2 + \left(1 - (\omega/\tilde{\omega})^2\right)^2\right]^{3/2}}{\left[(z/\rho)^2 + \left(1 - (\omega/\tilde{\omega})^2\right)\right]^2} \qquad [3.62]$$

These equations can then be used to determine the required sail cone angle and lightness

number for a given NKO of dimension z by ρ and period $T = 2\pi/\omega$. Nested cylinders and tori of sail lightness number are generated by Eq. 3.61, shown in **Figure 3.14**, for one-year NKOs. A complete treatment of these families of NKOs is given in McInnes (1999), along with an investigation of their stability and control. Planet centred NKOs can also be generated by noting the planet-centred form of the sail acceleration vector, shown in **section 3.9**. Again, this is discussed in depth in McInnes (1999). Using the methods of **section 3.10**, transfers to these displaced NKOs can be optimised, as shown in the 210 day transfer to a one year Earth-synchronous NKO in **Figure 3.15**. High performance sails may use these NKOs as unique vantage points for IR telescopes or for permanently viewing the solar poles.

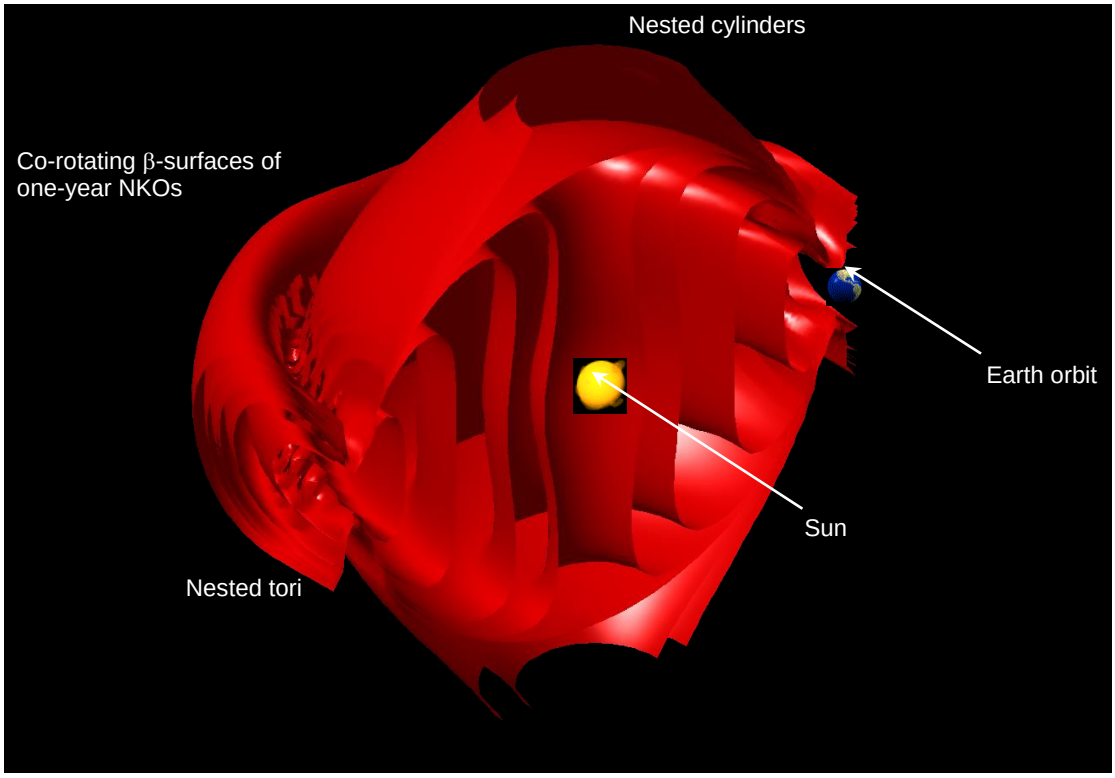

Figure 3.14 Surfaces of equal lightness number for 1 year displaced two-body NKOs

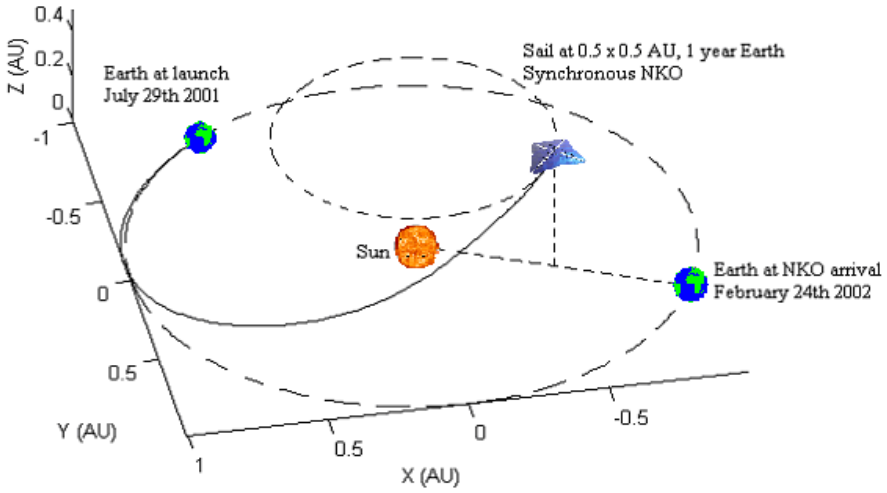

Figure 3.15 Minimum-time transfer to 0.5 x 0.5 AU, 1 year NKO [Hughes and McInnes, 2001]

The Lagrange points are a well known consequence of the restricted three-body problem. However, the addition of the solar sail enables a large family of Artificial Lagrange points, similar to the displaced NKOs described earlier, but with much lower sail performance requirements. The 2-body NKOs above require very high lightness numbers, but these 3-body NKOs would only require near to mid-term performance sails. The orbit geometry is depicted in **Figure 3.16**, and again, a detailed analysis of these 3-body NKOs is provided in McInnes (1999). Equivalent equations for the required sail lightness number and cone angle can be derived, with a clock angle of zero. Contours of sail lightness number are shown in **Figure 3.17**, for families of artificial Lagrange points in the Earth-Sun-Sail system. The contours correspond to sail lightness numbers of 0.01 (contour 1) to 1.02 (contour 8). The figure also shows the forbidden regions where there are no equilibrium solutions, as is also observed for 2-body NKOs in **Figure 3.14**.

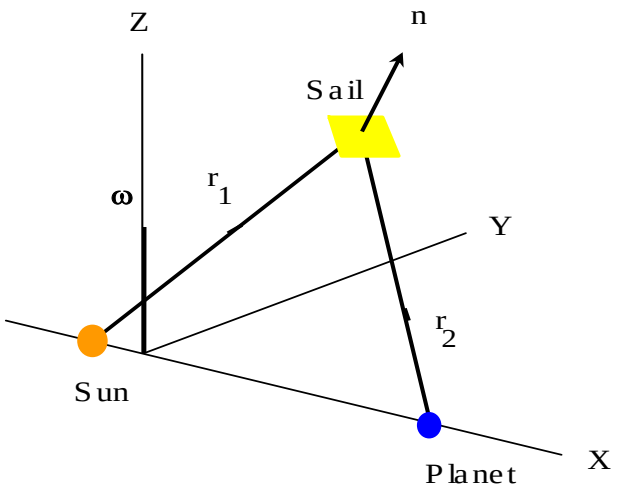

Figure 3.16 Solar sail in restricted three-body system

Near-term applications of 3-body NKOs are envisaged in the Geostorm and Polar Observer missions. Geostorm was a mission concept studied by NASA/JPL, in which a modest solar sail spacecraft is placed Sunward of the classical Earth-Sun L_1 point. The use of a magnetometer to detect solar wind polarity would enable a twofold increase in the warning time of conventional L_1 Halo orbiters such as SOHO. The Polar Observer mission would use an artificial Lagrange point displaced above the ecliptic plane, high above one of the Earth's poles. This would provide real-time views of the polar latitudes for studying climate change, or could provide line-of-sight telecommunications for high-latitude groundstations and users.

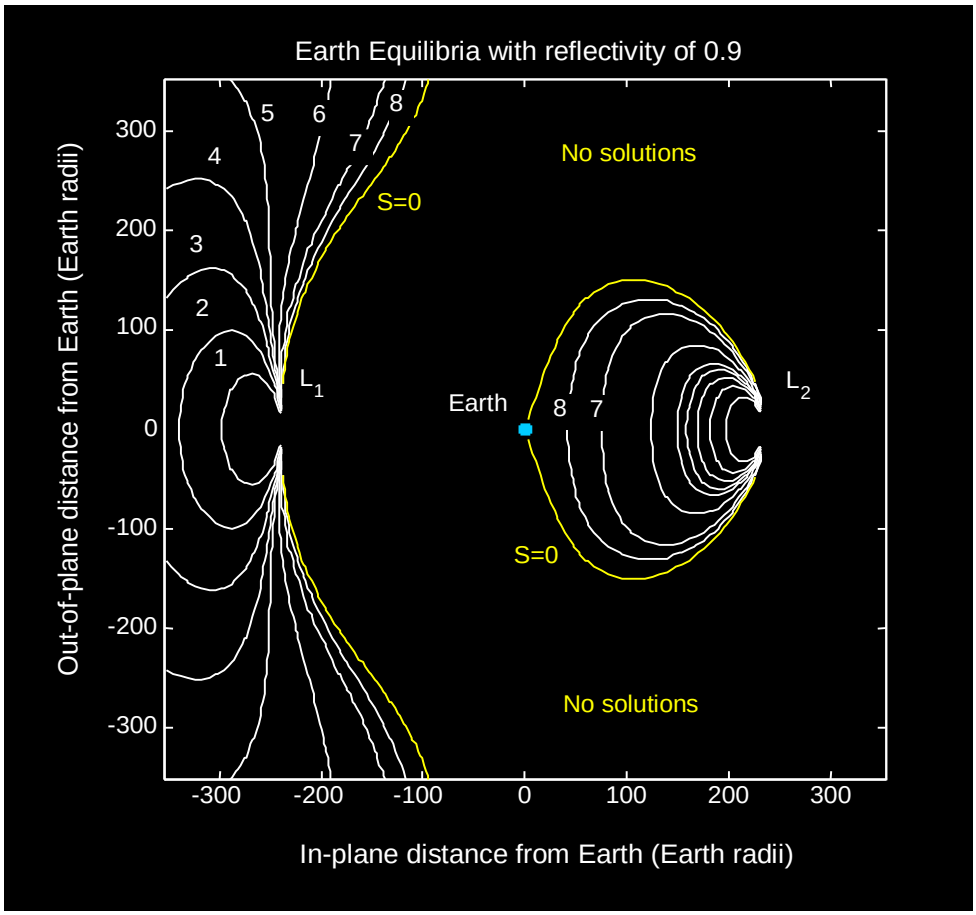

Figure 3.17 Contours of sail lightness number for families of artificial Lagrange equilibria about the Earth [McInnes, 1999]

3.9 Planet-Centred Orbits

Earth escape trajectories using solar sailing will now be discussed. In particular trajectories using a variety of simple sail steering laws, under simplistic assumptions. Sackett and

Edelbaum investigated minimum time escape trajectories in 1978 using a method of averaging to reduce the computational effort significantly. This method is only valid when the change in orbit elements over the averaged interval is small, thus the method is not valid up to escape and is instead propagated to a sub-escape point. It was found that the use of locally optimal control laws could generate escape trajectories within 1 – 3.5 % of globally optimal for Earth escape from high Earth orbits, thus we will focus on sub-optimal trajectories rather than attempting to generate true globally optimal trajectories, which would be a very intensive and complex process. A significantly more detailed analysis of locally optimal solar sailing planetary escape trajectories can be found in the journal paper by Macdonald.* However it is important to note that the results in that paper do not significantly vary from those that will be presented here.

As solar radiation pressure force cannot be directed sunward, solar sail escape trajectories are quite different from escape trajectories using other forms of low thrust propulsion. In particular, almost no energy can be gained while the solar sail is moving sunward and of course while the solar sail is in shadow. Such issues can be partially alleviated by using polar escape trajectories, which are normal to the Sun-line. These orbits can have an almost fixed Sun aspect angle and can be chosen to be free of eclipse periods under certain conditions.§ However, it will be shown that polar escape trajectories are in fact still less efficient than equatorial trajectories. The same steering laws that will be derived for escape trajectories can also be used for capture spirals.

The general equation of motion for a perfectly reflecting solar sail in a planet-centred orbit is

$$\frac{d^2 r}{dt^2} + \mu \frac{\mathbf{r}}{r^3} = \kappa (\mathbf{l} \cdot \mathbf{n})^2 \mathbf{n} \qquad [3.63]$$

where, κ is the magnitude of the solar sail acceleration and $\mathbf{l}$ is the unit vector defined along the Sun-line. For Earth-centred trajectories κ is the solar sail characteristic acceleration. Due to the scale of the spiral orbit relative to the Earth-Sun distance, the magnitude of the solar sail acceleration can be assumed to be constant in this analysis. In addition, the Sun-line will have a slow annual rotation of ~ 0.986° per day due to the motion of the Earth about the Sun. However, as the Earth's orbit about the Sun is slightly elliptical this rate will vary during the year.

Escape trajectories are found to require long spiral times due to the high local gravitational acceleration in Earth orbit. Furthermore, for effective escape manoeuvres it will be shown that rapid solar sail rotation rates may be required, placing large demands on the sail attitude control actuators. For these reasons Earth escape trajectories are in fact best avoided if possible by using direct injection into an interplanetary trajectory using a chemical kick-stage. Clearly such direct escape strategies have cost implications. However, by

* Macdonald M., McInnes C. R., 'Realistic Earth Escape Strategies for Solar Sailing', Journal of Guidance, Control, and Dynamics, Jan.-Feb. 2005.

avoiding long escape spirals and rapid rotations of the sail, solar sail designs can be optimised for the more benign conditions of interplanetary cruise.

3.9.1 Earth Escape – On / Off Switching

We will first consider a simple On / Off switching of the solar radiation pressure force using the sail pitch attitude. The sail is oriented to face the Sun for one half of the orbit and will then be oriented edge-on to the Sun for the other half orbit, as shown in **Figure 3.18**.

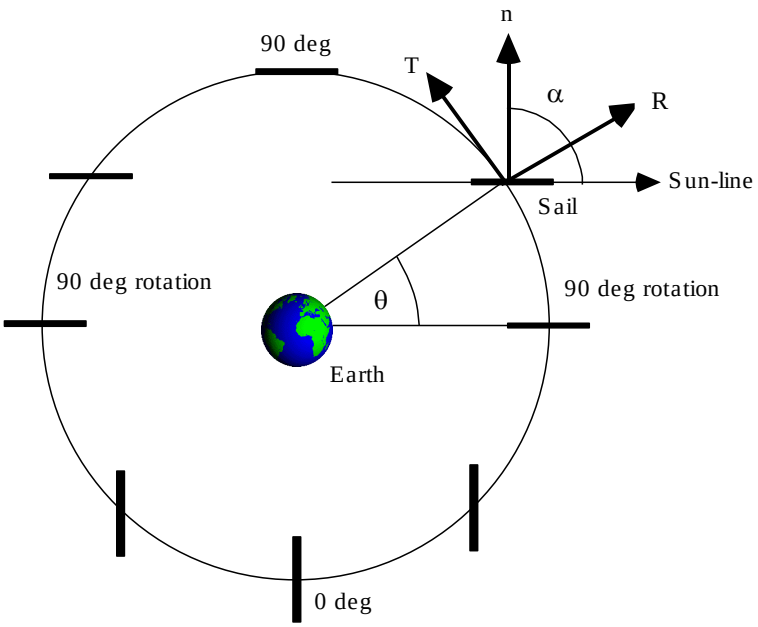

Figure 3.18 On / Off Steering law

Although the switching profile is simple, a rapid 90° rotation manoeuvre is required twice per orbit. The duration of these rotations must be short relative to the orbit period to avoid loss of efficiency. Therefore, low altitude starting orbits are undesirable due to the long spiral time, short orbit period and the presence of long periods of eclipse. A further consequence of the asymmetry of the switching law is an increase in orbit eccentricity with the solar sail orbit expanding normal to the Sun-line.

The increase in semi-major axis per orbit can be estimated using the variational equations of motion of the classical orbit elements. For a small orbit eccentricity the variational equation for semi-major axis becomes

$$\frac{da}{df} = \frac{2a^3}{\mu} T(f) \qquad [3.64]$$

Neglecting any periods of eclipse the transverse acceleration experienced by the solar sail is

$$T(f) = 0 \qquad (0 \le f \le \pi) \qquad [3.65]$$

$$T(f) = \kappa \cos\left[f - \frac{3\pi}{2} \right] \qquad (\pi \le f \le 2\pi) \qquad [3.66]$$

Therefore, the change in semi-major axis over one orbit is obtained by integrating the transverse acceleration using

$$\Delta a = \frac{2a^3}{\mu}\int_0^{2\pi} T(f)\ df \qquad [3.67]$$

Therefore, integrating over the half orbit when the solar sail is facing the Sun it is found that

$$\Delta a = \frac{4\kappa}{\mu}a^3 \qquad [3.68]$$

Since the semi-major axis gain is a strong function of the initial semi-major axis, the benefit of a high starting orbit is again seen. In the current and following analysis the steering laws will be illustrated using a solar sail with a lightness number of 0.17, corresponding to a characteristic acceleration of 1 mm s^{-2}. The spiral manoeuvre will begin from an idealised geostationary orbit (GEO), corresponding to an initial semi-major axis of 42 241 km and an eccentricity of zero. A moderate performance solar sail and a high starting orbit are used to obtain a reasonably quick increase in semi-major axis. Each steering law presented will be used for a period 72 hours to illustrate the variation of the solar sail orbital elements and the sail pitch angle as a function orbit number. An example orbit-raising manoeuvre is shown in **Figure 3.19** using this simple switching law. It can be seen that rapid sail rotations are required and that energy is gained during only half of each orbit. The method is therefore inefficient in gaining energy, but requires only a simple Sun-facing attitude for half of each orbit.

3.9.2 Earth Escape – Orbit Rate Steering

Given the limitations of the switching law a more efficient orbit-raising scheme is considered. Here the solar sail rotates at one half of the orbit rate, allowing some energy to be gained during the half orbit when the solar sail is moving sunward. Since the orbit rate varies as the semi-major axis and eccentricity of the orbit increases, the sail pitch rate will not in general be constant. The sail pitch rate is related to the orbit rate by

$$\frac{d\alpha}{dt} = \frac{1}{2}\frac{df}{dt} \qquad [3.69]$$

Integrating and ensuring the correct phasing of the sail pitch angle with the Sun-line, the pitch angle will be defined as

$$\alpha = \frac{\pi}{4} + \frac{f}{2} \qquad [3.70]$$

as illustrated in **Figure 3.20**.

It is found that the transverse solar radiation pressure acceleration may be written as

$$T(f) = \kappa\left|\cos^3\left[\frac{\pi}{4} + \frac{f}{2}\right]\right| \quad (0 \le f \le 2\pi) \qquad [3.71]$$

Substituting for the transverse acceleration and integrating, the change in semi-major axis per orbit is found to be

$$\Delta a = \frac{16\kappa}{3\mu} a^3 \qquad [3.72]$$

which is clearly greater than that for the simple switching law.

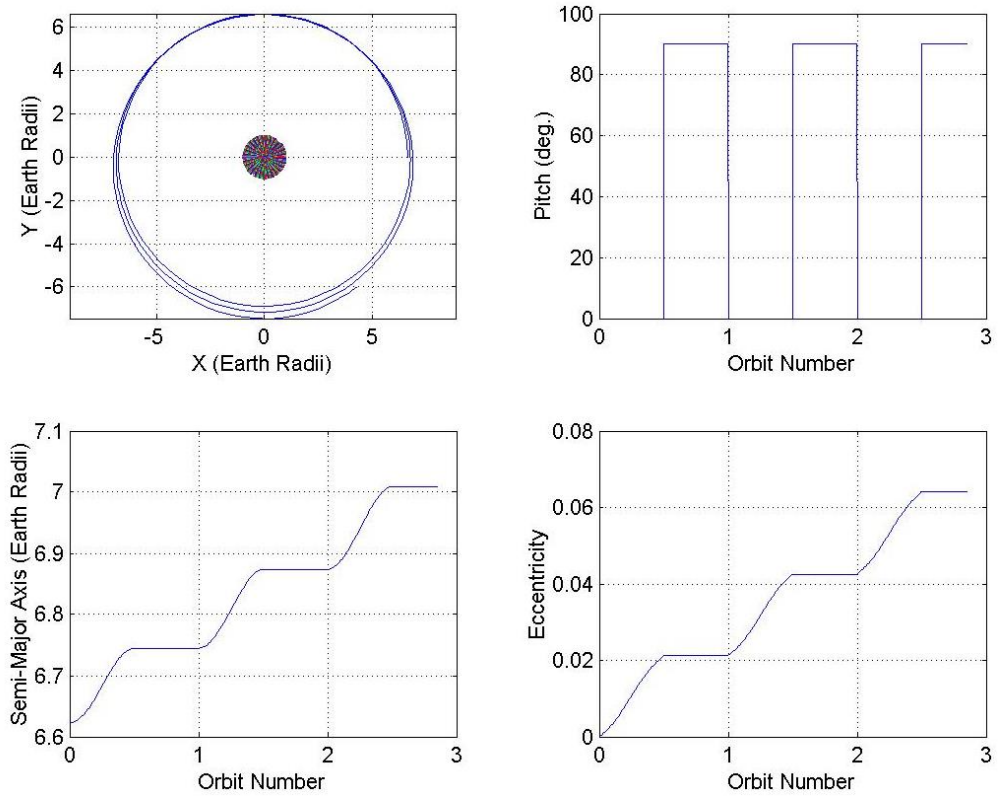

Figure 3.19 On / Off orbit-raising manoeuvre

An orbit-raising manoeuvre using this steering law is shown in **Figure 3.21**. In this case a rapid 180º rotation is required once per orbit, again imposing demands on the sail attitude control actuators, particularly for low altitude starting orbits. An alternative to the rapid rotation manoeuvres is to use a sail that is reflective on both sides. In this case the sail rotation rate is almost constant, slowly decreasing as the orbit period increases. While this alleviates the need for rapid rotations, additional mass is clearly added to the sail thus reducing its performance. It can also be seen from **Figure 3.21** that the energy gain during the sunward part of the orbit remains small. The main benefit of the orbit rate steering law is to align the solar radiation pressure force vector more closely to the solar sail velocity vector than is possible with the switching law.

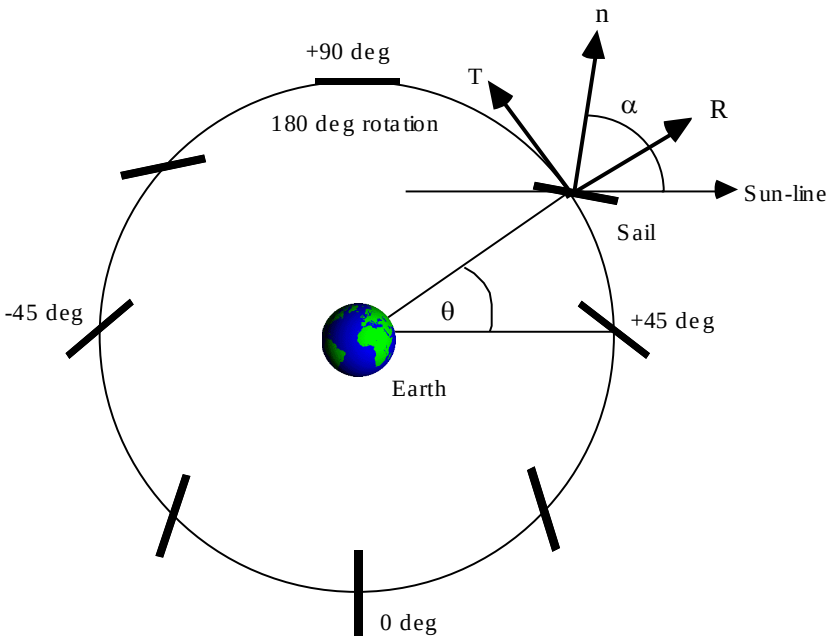

Figure 3.20 Orbit rate steering

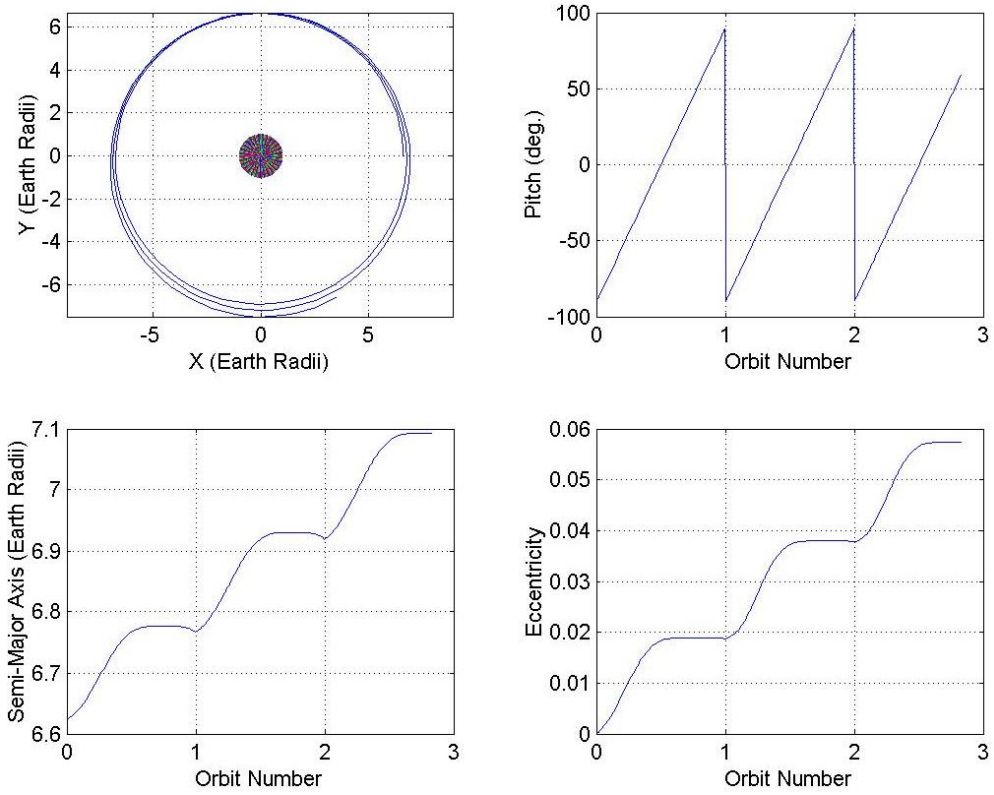

Figure 3.21 Orbit rate orbit-raising manoeuvre

3.9.3 Earth Escape – Locally Optimal Steering

To maximise the instantaneous rate of energy gain a locally optimal steering law can be derived, where the sail thrust vector is maximised along the velocity vector. Taking the scalar product of the solar sail equation of motion, [3.63], with the solar sail velocity vector $\mathbf{v}$ it can be seen that

$$\frac{d^2\mathbf{r}}{dt^2}.\mathbf{v} + \mu\frac{\mathbf{r}.\mathbf{v}}{r^3} = \kappa(\mathbf{l}.\mathbf{n})^2\mathbf{n}.\mathbf{v} \qquad [3.73]$$

However the left side of this may be written as

$$\frac{d^2\mathbf{r}}{dt^2}.\mathbf{v} + \mu\frac{\mathbf{r}.\mathbf{v}}{r^3} = \frac{d}{dt}\left[\frac{1}{2}\mathbf{v}.\mathbf{v} - \frac{\mu}{r}\right] \qquad [3.74]$$

which represents the rate of change of total orbit energy E. Therefore, the instantaneous rate of change of orbit energy is given by

$$\frac{dE}{dt} = \kappa(\mathbf{l}.\mathbf{n})^2\mathbf{n}.\mathbf{v} \qquad [3.75]$$

To maximise the energy gain at all points along the trajectory the sail pitch angle must now be chosen to maximise the component of solar radiation pressure force along the solar sail velocity vector. Defining ψ as the angle of the solar sail velocity vector relative to the Sun-line, as shown in **Figure 3.22**, it can be seen that

$$(\mathbf{l}.\mathbf{n})^2\mathbf{n}.\mathbf{v} = \cos^2\alpha\cos(\psi - \alpha) \qquad [3.76]$$

Thus, to maximise the instantaneous rate of change of orbit energy the turning point must be determined. This corresponds to finding the solar sail pitch angle that maximises the orbit energy gain for each value of ψ. The turning point is found from

$$\frac{d}{d\alpha}\left[\cos^2\alpha\cos(\psi - \alpha)\right] = 0 \qquad [3.77]$$

which yields the optimal sail pitch angle as

$$\alpha^* = \frac{1}{2}\left[\psi - \sin^{-1}\left(\frac{\sin\psi}{3}\right)\right] \qquad [3.78]$$

Using this locally optimal steering law the transverse acceleration of the solar sail is given by

$$T(f) = \kappa\cos^2\alpha^*\left|\sin(\alpha^* - f)\right| \qquad [3.79]$$

For a near circular orbit, $\psi \sim f + \pi/2$, such that that the transverse acceleration is a function of true anomaly only. Thus the change in semi-major axis per orbit is obtained as

$$\Delta a = \frac{5.52\kappa}{\mu}a^3 \qquad [3.80]$$

An orbit-raising manoeuvre is shown in **Figure 3.22** using this locally optimal steering law. Again, in this case either a rapid 180° rotation is required once per orbit, or the sail must be reflective on both sides. There is little energy gain during the sunward part of each orbit, which also coincides with the time for the sail rotation. It can be seen from **Figure 3.23** that the sail pitch angle profile is similar to that of the orbit rate steering law. As the eccentricity of the orbit grows the pitch rate varies to track the rotations of the solar sail velocity vector.

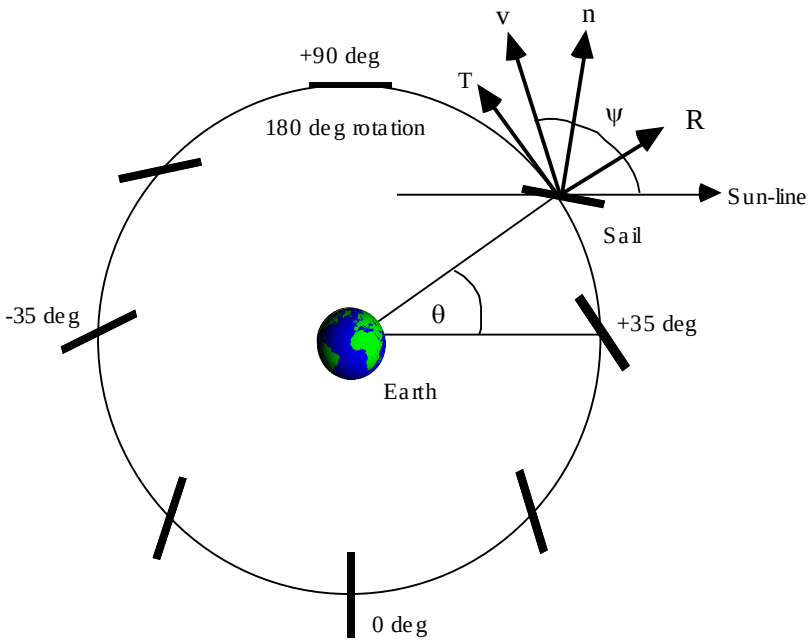

Figure 3.22 Locally optimal steering (Image adapted from McInnes, 1999)

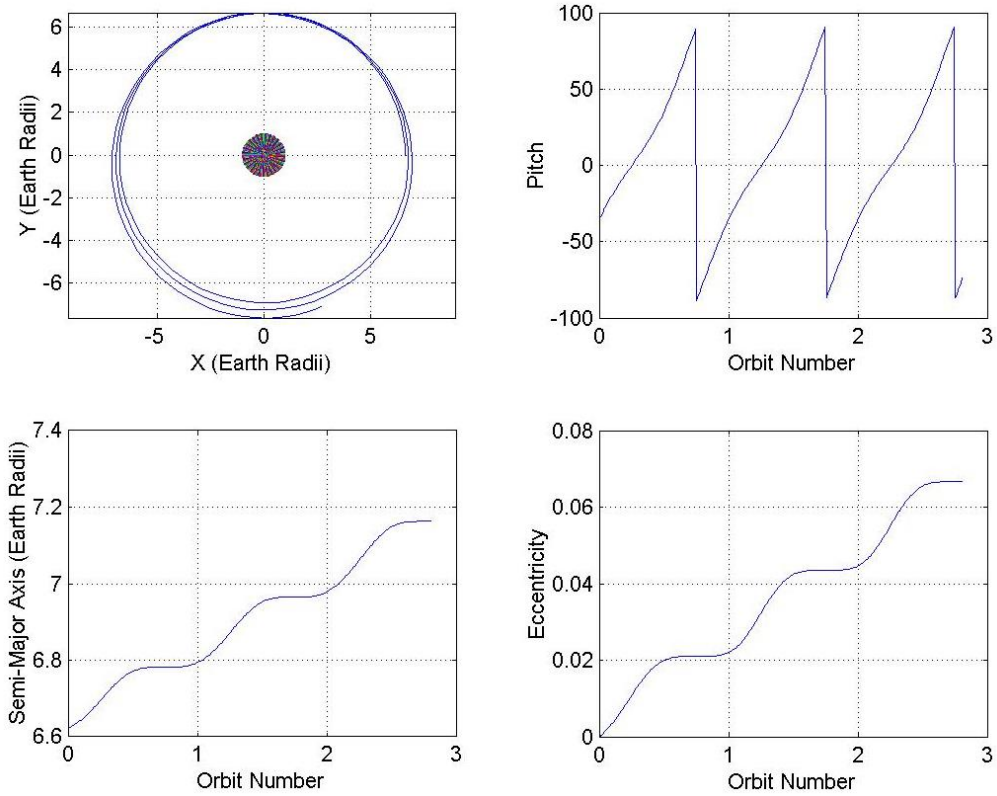

Figure 3.23 Locally optimal orbit-raising manoeuvre

A longer 60-day escape spiral from geostationary orbit is shown in **Figure 3.24** with a rotating Sun-line. It can be seen that as the orbit eccentricity grows, the major axis of the orbit remains normal to the Sun-line, while the sail pitch corrects for the ever-varying orbit size and shape. We also note that the escape asymptote is away from the Sun.

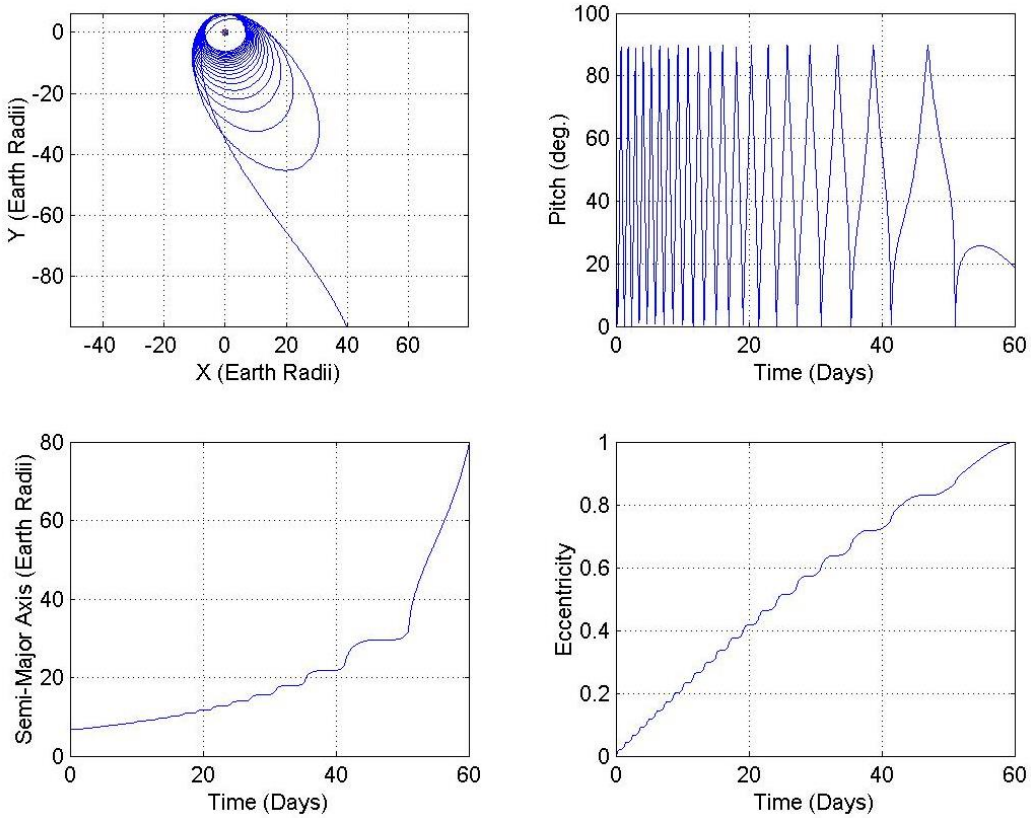

Figure 3.24 60-day locally optimal escape spiral

3.9.4 Earth Escape – Polar Orbit

To avoid rapid sail rotations and enable energy gain at all points along the trajectory, polar orbit escape can be considered with the solar sail orbit plane normal to the Sun-line. In this case the transverse acceleration is constant,

$$T(f) = \kappa \cos^2 \alpha \sin \alpha \qquad [3.81]$$

For a near circular orbit the semi-major axis change per orbit is then obtained as

$$\Delta a = \frac{4\pi\kappa}{\mu} a^3 \cos^2 \alpha \sin \alpha \qquad [3.82]$$

Hence, the semi-major axis gain is maximised if the transverse acceleration is maximised using the sail pitch angle $\alpha = \tan^{-1}\left(1/\sqrt{2}\right)$. Then the change in semi-major axis per orbit is given by

$$\Delta a = \frac{8\pi\kappa}{3\sqrt{3}\mu} a^3 \qquad [3.83]$$

The change in eccentricity during the spiral manoeuvre may also be obtained as

$$\Delta e = \frac{2a^2}{\mu} \int_0^{2\pi} T(f) \cos f \; df \qquad [3.84]$$

However, as the transverse force is constant it is found that Δe ~ 0, such that the orbit remains quasi-circular as the solar sail spirals outwards. Although the sail pitch angle remains constant, the sail must still roll through 360° once per orbit to keep the transverse acceleration tangent to the trajectory. This is significantly less demanding than the rotation manoeuvres required for the other sail steering laws discussed.

An orbit-raising manoeuvre using polar orbit escape is shown in **Figure 3.25**. Although an initial polar orbit at geostationary altitude is impractical, it will be used for comparison with the earlier solar sail steering laws. It can be seen that the semi-major axis monotonically increases demonstrating the constant orbit energy gain as the solar sail spirals outwards. It is also clear that the orbit remains quasi-circular, unlike other steering laws where the eccentricity increases due to the asymmetry of the transverse solar radiation pressure force experienced by the solar sail during each orbit. Since the orbit remains quasi-circular, the variational equation for semi-major axis may be used to obtain a closed form solution for the outward spiral. The variational equation may be written as

$$\int_{a_o}^{a} \frac{da}{a^3} = \frac{2T}{\mu} \int_{f_o}^{f} df \qquad [3.85]$$

Therefore, integrating and re-arranging the orbit equation is obtained as

$$a(f) = a_o \left[1 - \frac{4a_o{}^2}{\mu} \kappa \cos^2 \alpha \sin \alpha (f - f_o) \right]^{-1/2} \qquad [3.86]$$

We can thus estimate the number of orbit required to attain escape from GEO radius, with a sail characteristic acceleration of 1 mm s^{-2} as

$$\frac{1}{2\pi \left(\frac{4a_0^2}{\mu} \kappa \cos^2 \alpha \sin \alpha \right)} = 23.09 \qquad [3.87]$$

Continuing the propagation shown in **Figure 3.26** it is found that escape occurs after 23.62 orbits, as shown in **Figure 3.27**. We note from **Figure 3.27** that the rate of spin of the sail varies during escape, but since e < 0.5 for the majority of the trajectory the attitude control demands remain low.

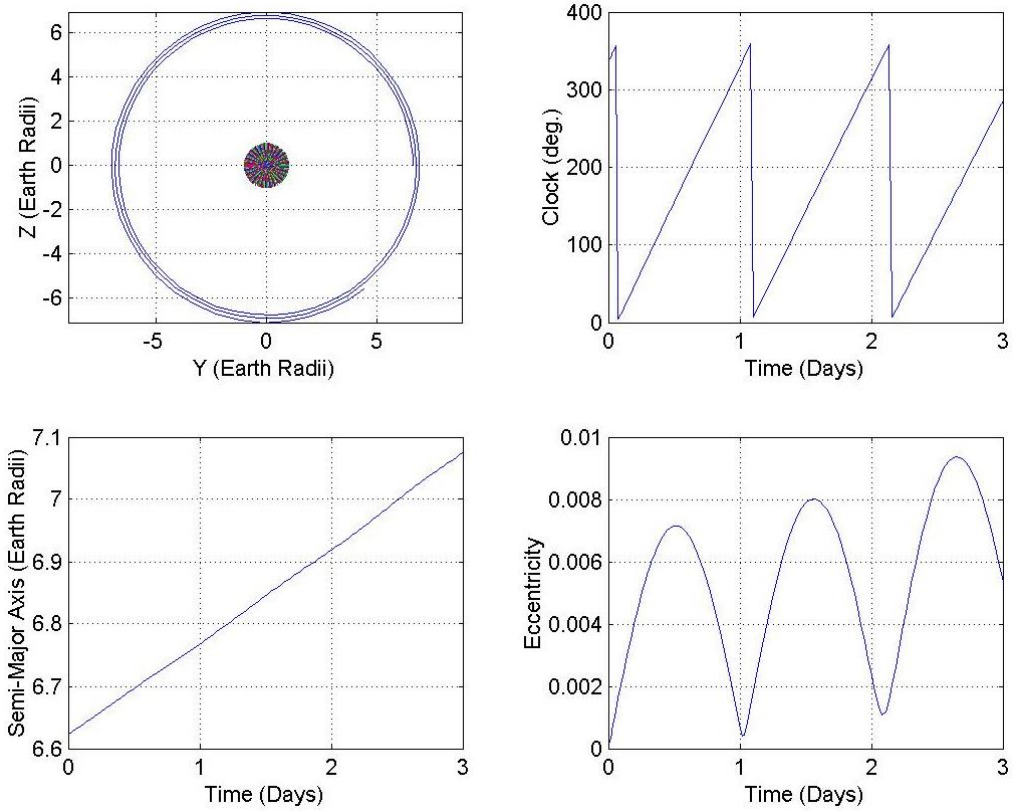

Figure 3.25 Polar orbit-raising manoeuvre

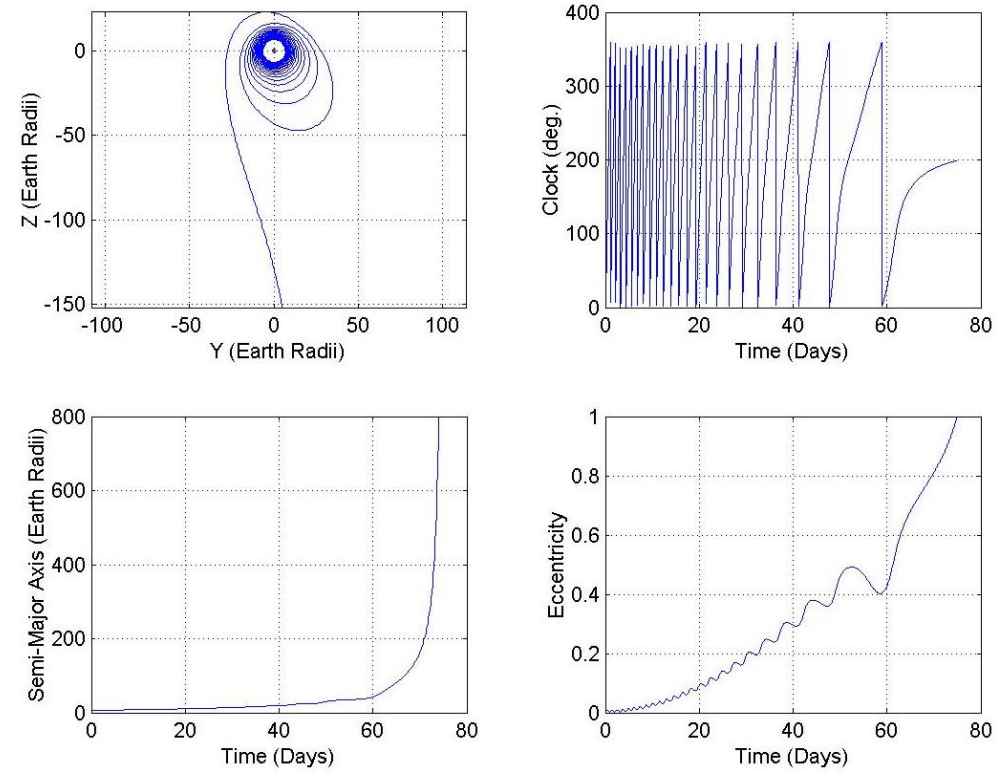

Figure 3.26 Polar orbit raising until escape

As a solar sail passes through a planet's shadow cone the solar radiation flux over the sail surface drops to zero; as does the thrust level, so that a secondary attitude control system may be required during shadow passage. This secondary system could take any of the standard forms, however all of these would increase system mass and correspondingly decrease sail performance. Additionally, shadow events will impart severe thermal loads on the sail systems that will dynamically excite the structure, and thus stress the sail, requiring heavier booms and/or thicker film coatings which further degrade sail performance. Eclipse will also cause large charging swings, it is thus attractive to be able to generate planetary escape trajectories that avoid planetary occultation of the sail – Sun-line; this would potentially enable a reduction in sail assembly loading and a corresponding increase in sail acceleration or payload capability.

Using a serious of blended control algorithms outlined by Macdonald*, which simplify down to the locally optimal controller discussed in **Section 3.9.3**, the required sail characteristic acceleration for escape from an Earth polar orbit at a range of altitudes was found, **Figure 3.27**. We define the initial orbit such that the initial orbit normal is aligned with the Earth-Sun line and calculation start epoch is set at Vernal Equinox. The orbit model utilised only considers perturbations due to the sail acceleration; the introduction of other perturbations such as gravitational harmonics or a more realistic sail force model significantly prolongs calculation time and from experience typically alters escape time by between three and five percent.

The required sail characteristic acceleration for a range of initial altitudes from 800 km to 25000 km is shown in **Figure 3.27**, where we see an exponential increase in sail acceleration requirements as altitude is decreased in order to maintain a shadow free escape trajectory. This exponential curve is analogous to the well-documented exponential reduction in escape time as initial altitude is increased for a given sail acceleration or the exponential reduction in escape time for a given altitude as sail acceleration is increased. The corresponding minimum time for shadow free escape from each altitude is shown in **Figure 3.28**. As would be expected from the exponential curve of required sail accelerations the minimum escape time for shadow free trajectories is essentially independent of initial altitude; as the required sail acceleration varies exponentially thus maintaining a constant escape time. The mean escape time was found to be 141.46 days, the standard deviation in the escape time data is 6.1 days.

* Macdonald M., McInnes C. R., 'Realistic Earth Escape Strategies for Solar Sailing', Journal of Guidance, Control, and Dynamics, Jan.-Feb. 2005.

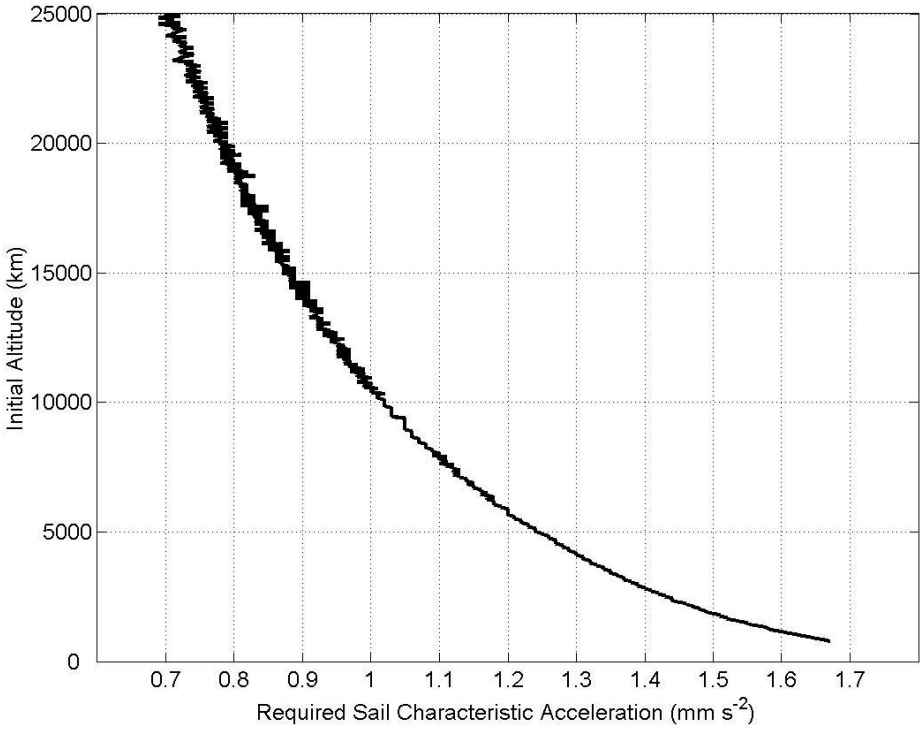

Figure 3.27 Required sail characteristic acceleration for shadow free Earth escape, from a polar orbit

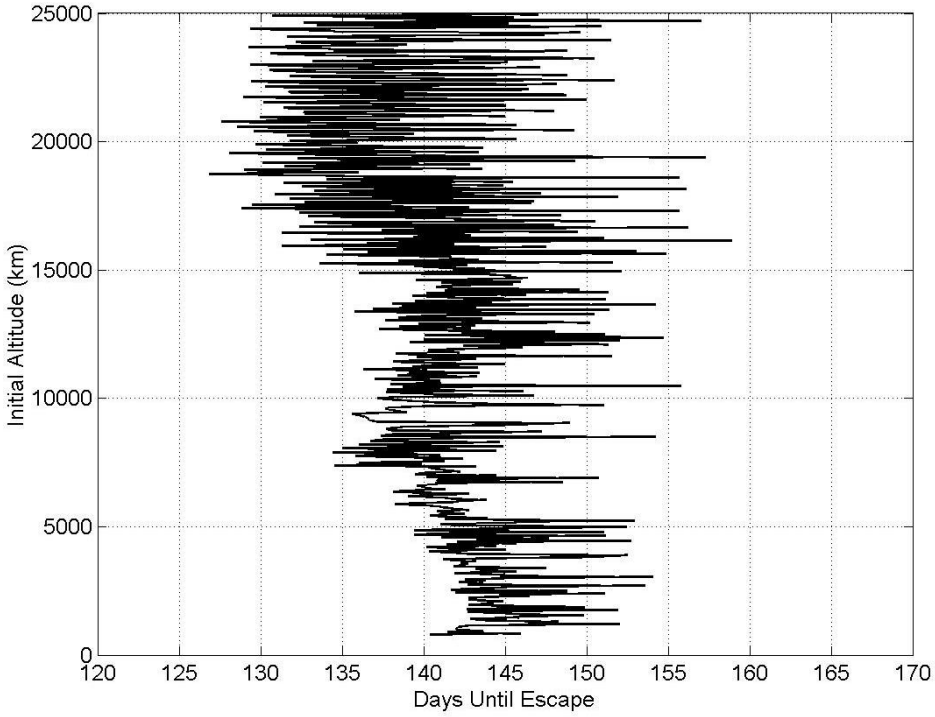

Figure 3.28 The minimum shadow free escape time from each altitude in **Figure 3.27**

3.10 Heliocentric Optimisation Software

The Indirect methods of Optimal Control Theory have been described in **section 3.7**, where the Calculus of Variations is used to obtain true-optimal trajectories. As has been noted, this method is highly sensitive to the initial co-state estimates and as such requires many years of experience to use successfully. A number of closed loop analytical methods exist, such as those used in **section 3.6** and **3.9**. These methods are simple to use, but can never be truly optimal. It should be noted that these control laws for each element can be blended together using weighting functions to potentially create optimal changes in all six orbital elements. However, this still does not solve the two-point boundary value problem. Other methods, such as two-variable asymptotic expansion have also been proposed for low-thrust trajectory optimisation in the past. A plethora of Direct, open-loop optmisation methods exist, where the control history over the entire trajectory is optimised in a 'top-down' approach. In addition to the control profile, the state vectors at multiple points along the trajectory also become constrained parameters in a process known as Collocation. Multiple-shooting methods propagate adjacent trajectory segments backwards and forwards in time in an iterative process in an attempt to get the segment boundary states to match. Many Direct parameter optimisation methods use Non-linear programming algorithms such as Sequential Quadratic Programming to optimise the parameters. Some of these methods are quite robust, but they all suffer a drawback in that they are deterministic, gradient-based, local-search methods. Local search methods can converge to an optimal solution to a high degree of accuracy, however, an initial guess of the control angles within the region of the global optimum needs to be provided to ensure convergence on the global minimum in transfer time. The initial guess can be provided through discretisation of solutions from analytical control laws, or from Global-search methods. True global optimisation is the Holy Grail of trajectory analysis, but there has been little success so far. Deterministic Branch and Bound methods have been developed, such as the Multi-level Coordinate search method, which divide the global search space into successively smaller regions. Stochastic processes which use analogies from the biological or physical world are widely available. Simulated Annealing uses the physical concept of annealing, whereby liquids freeze and metals recrystalise, based on the Metropolis algorithm. This is not often applied to trajectory optimisation. Genetic Algorithms are based on the evolutionary concept of natural selection and survival of the fittest. They have been widely used for trajectory optimisation problems, and have been used by the Authors for solar sail trajectory optimisation.

The main optimisation algorithm used at the University of Glasgow for solar sail trajectory analysis is Sequential Quadratic Programming (SQP). The generalised optimal control problem is defined by Eq. 3.88. Where, J is the objective function which must be minimised subject to the constraints, $c_i = 0$, for i = 1 to n, where n is the total number of constraints, $\underline{x}$ is the state vector, $\underline{u}$ is the control vector, and t_f is the terminal time. This is then transcribed to a Non-linear Programming problem by discretisation of variables, and then solved using NPSOL 5.0 or SNOPT 6.0 in Fortran77. These subroutines are based on a SQP

algorithm coded at Stanford Unversity. SQP employs a Newton or quasi-Newon approximation to the Kursh-Kuhn-Tucker conditions of optimality. This results in a sub-problem of minimising a quadratic approximation to the function of Lagrange multipliers. These Lagrange multipliers incorporate the objective function and constraints of the trajectory optimisation problem.

$$J(\underline{x}(t_f),\underline{u},t_f) = f(\underline{x}(t_f),\underline{u},t_f) + \int_0^{t_f} g(\underline{x}(t),\underline{u},t)dt \qquad [3.88]$$

The trouble with SQP is that it only finds local minima in an, often, extremely complex search space. As has been mentioned, an initial guess of the control parameters close to the global solution must be supplied to ensure convergence to the near-globally optimal solution. We have used a manual iterative method of successively visualising the trajectory until the trajectory end-point is within proximity of the desired end-point to enable the SQP algorithm to produce a feasible solution. When engineering insight cannot be used then we resort to the global search properties of a Genetic Algorithm, or employ an analytical closed loop control law to generate the initial guess.

The Genetic Algorithm (GA) simulates the biological concept of natural selection and has been used with very good results, in many applications. For low-thrust trajectory optimisation problems, the parameter search space can contain multiple maxima and minima, or 'hills and valleys'. However, the GA searches from a population of points. As the population evolves over successive generations, a number of genetic operators are executed. Tournament Selection compares two individuals at random and allows the one with the highest fitness go on to reproduce. This process is repeated until the whole population is covered. Single-Point Crossover facilitates information exchange between the selected parents and randomly interchanges sections of the parent's genomes. Mutation helps maintain genetic diversity in the population by randomly 'flipping' a bit in the chromosome. One special operator used was Elitism, which copies the fittest individual from one generation to the next to prevent it inadvertently being lost. The GA has a good global optimisation capability, but poor convergence properties, so we used SQP to refine the solution. In addition, it is difficult to incorporate constraints in the single-valued GA fitness function. We use a penalty function method to achieve this, although other authors have used the concept of Pareto-fronts with some success.

For heliocentric optimisation, the trajectory is divided into segments of equal time. The control variables to be optimised are then the discrete nodes at the segment boundaries. The instantaneous controls are then obtained by linear interpolation between the nodes across each segment. It would also be possible to use Hermite or Cubic polynomials, but if the nodal resolution is high enough then linear interpolation more than adequately

approximates a continuous profile. A simpler control representation would be to have the cone and clock angles fixed relative to the Sun-line in each segment; this may be more easily accomplished in practice through passive means. These control representations are illustrated in **Figure 3.29**.

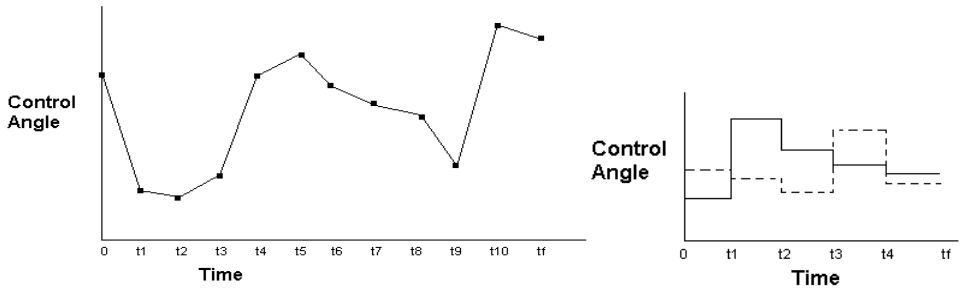

Figure 3.29 Two methods of discretised control representation

To summarise the heliocentric optimisation method, NPSOL or SNOPT is used to null the errors between the spacecraft and target end-point state vectors (constraints), whilst minimising transfer time (objective function). On the few occasions when engineering insight fails to find the control angles initial guess, then the GA or an analytical control law is used *apriori*. The variational equations are formulated in Modified Equinoctial Orbital Elements to avoid singularities, and then integrated using an Adams-Moulton-Bashforth adaptive step-size, variable order method at each SQP function evaluation. Full target ephemeris data is supplied by a cubic polynomial approximation to J2000. In the solar sail mathematical model, it is assumed that 2-body dynamics are prevalent (although third-body perturbations have been modelled). The sail is usually modelled as an ideal, flat, perfectly reflecting surface. The Sun is assumed to be a point source of radiation and a point mass. The solar wind is negligible, and so photon pressure dominates.

An example of an Earth-Mercury trajectory generated using the above method is shown in **Figure 3.11**. Many more trajectories to a variety of solar system targets have been generated at the University of Glasgow. **Figure 3.30** shows a dual asteroid rendezvous trajectory with a characteristic acceleration of 1.0 mm s^{-2}. The solar sail can enable long multiple-target, high Δv missions such as this due to the elimination of propellant mass. An exciting application of solar sailing is enabled by the greatly increased photon pressure at close solar distances. The sail can spiral down to as close as 0.25 AU with conventional coating and then pitch face on to effect a manoeuvre known as the solar photonic assist. The acceleration at 0.25 AU is 16 times higher than at 1 AU, enabling a considerable velocity change. The solar sail can then rapidly escape the solar system. An application of this is the Interstellar Heliopause Probe mission, which could reach the boundaries of interstellar space at 200 AU in 25 years, shown in **Figure 3.31**. Another unique application of the solar sail is to

spiral down to a close solar orbit, then initiate an orbit cranking manoeuvre to obtain a true Solar polar orbit, as shown in **Figure 3.32**.

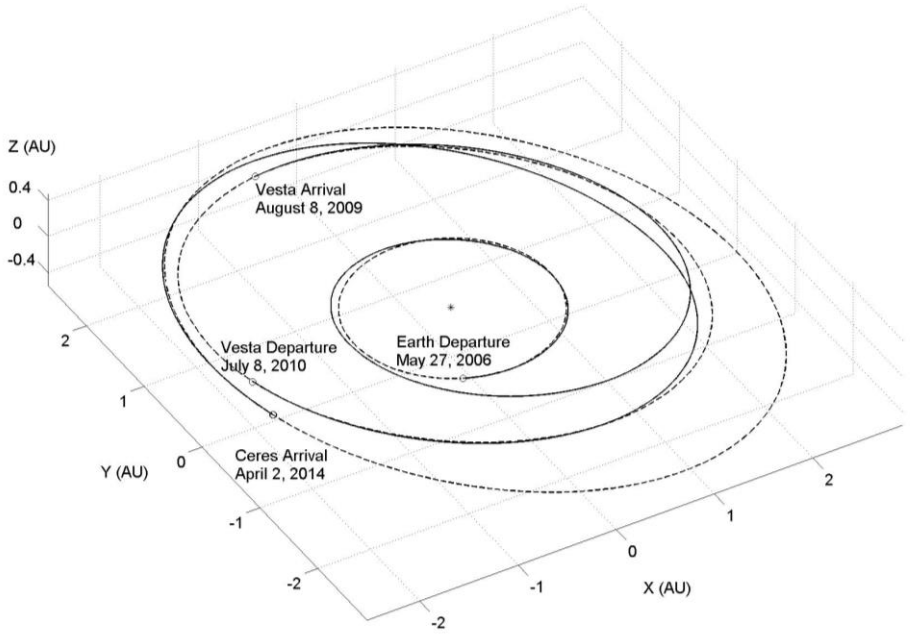

Figure 3.30 1.0 mm s^{-2} Earth-Vesta-Ceres dual asteroid rendezvous trajectory [Hughes and McInnes, 2004]

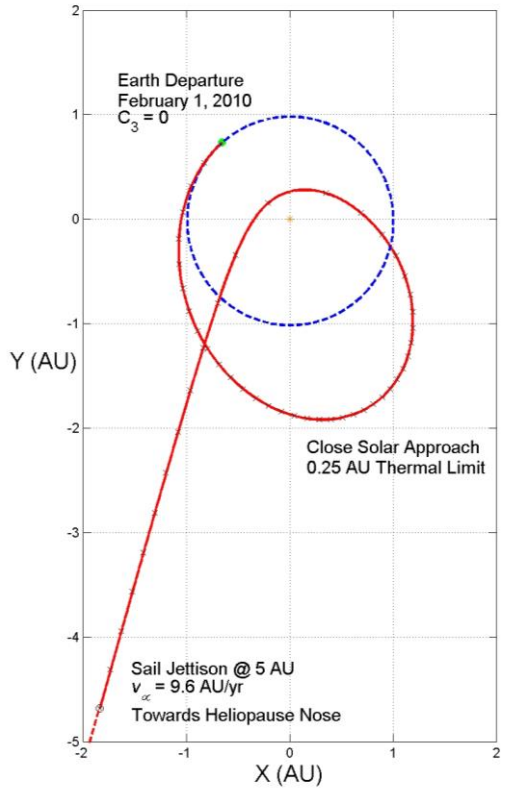

Figure 3.31 1.5 mm s^{-2} 25 year Interstellar Heliopause trajectory [Hughes]

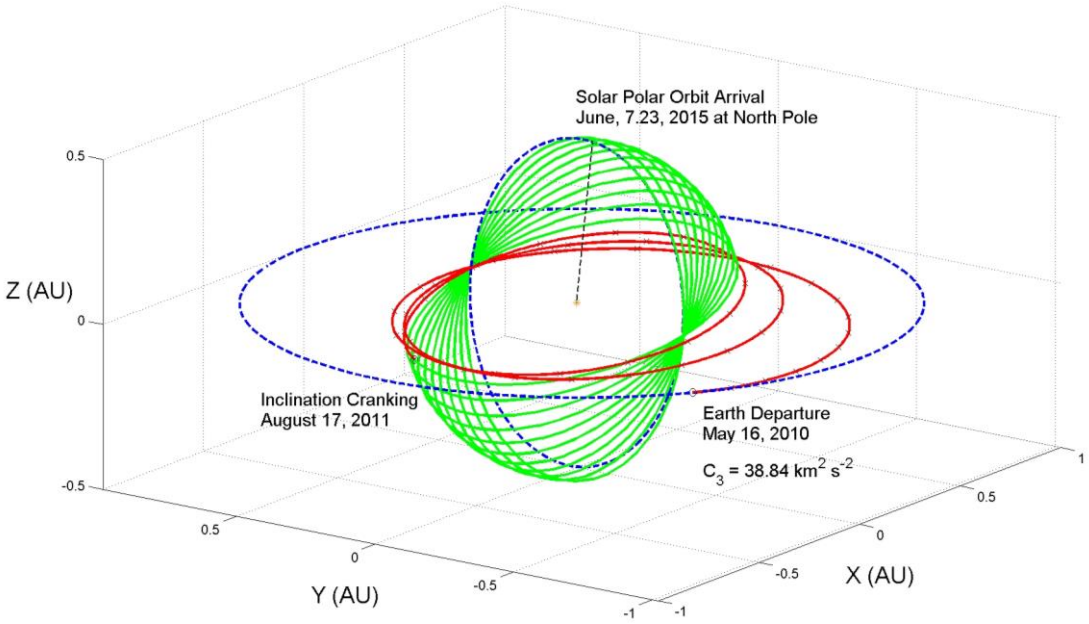

Figure 3.32 0.42 mm s^{-2} 5 year Solar Polar Orbiter transfer trajectory [Hughes]

3.10 Further Reading

3.10.1 Sun-centred trajectories

- Bacon, R.H., 'Logarithmic Spiral - an Ideal Trajectory for an Interplanetary Vehicle with Engines of Low Sustained Thrust', American Journal of Physics, **27**, 12-18, 1959
- Tsu, T.C., 'Interplanetary Travel by Solar Sail', American Rocket Society Journal, **29**, 422-427, 1959
- London, H.S., 'Some Exact Solutions of the Equations of Motion of a Solar Sail With a Constant Setting', Journal of the American Rocket Society, **30**, 198-200, 1960
- Kiefer, J.W., 'Feasibility Considerations for a Solar-Powered Multi-Mission Probe', Proceedings of the 15th International Astronautical Congress, Warsaw, **1**, 383-416, 1965.
- Hughes, G. W., and McInnes, C. R., "Solar Sail Hybrid Trajectory Optimization for Non-Keplerian Orbit Transfers," *Journal of Guidance, Control and Dynamics*, Vol. 25, No. 3, May-June 2002, pp. 602-604.
- Hughes, G. W., and McInnes, C. R., "Small-Body Encounters Using Solar Sail Propulsion," *Journal of Spacecraft and Rockets*, Vol. 41, No.1, January-February, 2004, pp. 140-150.
- Hughes, G. W., Macdonald, M., McInnes, C. R., Atzei, A., and Falkner, P., "Analysis of a Solar Sail Mercury Sample Return Mission," 55th International Astronautical Congress, Vancouver, Canada, October 4-8, 2004.
- Hughes, G. W., Macdonald, M., McInnes, C. R., Atzei, A., and Falkner, P., "Terrestrial Planet Sample Return Missions Using Solar Sail Propulsion," 5th IAA International Conference on Low-Cost Planetary Missions, ESA/ESTEC, The Netherlands, September 24-26, 2003.
- Sauer, C.G., 'A Comparison of Solar Sail and Ion Drive Trajectories for a Halley's Comet Rendezvous', AAS-77-4, AAS/AIAA Astrodynamics Conference, September 1977
- Van der Ha, J.C. & Modi, V.J., 'Long-term Evaluation of Three-Dimensional Heliocentric Solar Sail Trajectories with Arbitrary Fixed Sail Setting', Celestial Mechanics, **19**, 113-138, 1979
- Koblik, V.V., et. al., 'Controlled Solar Sailing Transfer Flights into Near-Sun Orbits Under Restrictions on Sail Temperature', Cosmic Research, **34**, 572-578, 1996

3.10.2 Minimum time trajectories

- Zhukov, A.N. & Lebedev, V.N., 'Variational Problem of Transfer Between Heliocentric Circular Orbits by Means of a Solar Sail', Cosmic Research, **2**, 41-44, 1964
- Sauer, C.G., 'Optimum Solar Sail Interplanetary Trajectories', AIAA-76-792, AAS/AIAA Astrodynamics Conference, August 1976
- Green, A.J., 'Optimal Escape Trajectory From a High Earth Orbit by Use of Solar Radiation Pressure', T-652, Master of Science Thesis, Massachusetts Institute of Technology, 1977

- Sackett, L.L. and Edelbaum, T.N., 'Optimal Solar Sail Spiral to Escape', AAS/AIAA Astrodynamics Conference, September 1977
- Sun, H. & Bryson, A.E., 'Minimum Time Solar Sailing from Geosynchronous Orbit to the Sun-Earth L_2 Point', AIAA-92-4657, AAS/AIAA Astrodynamics Conference, August 1992
- Simon, K. & Zakharov, Y., 'Optimisation of Interplanetary Trajectories with Solar Sail', IAF-95-A.2.08, 46th International Astronautical Federation Congress, October 1995

3.10.3 Planet-centred trajectories

- Sands, N., 'Escape from Planetary Gravitational Fields by use of Solar Sails', American Rocket Society Journal, **31**, 527-531, 1961
- Fimple, W.R., 'Generalised Three-Dimensional Trajectory Analysis of Planetary Escape by Solar Sail', American Rocket Society Journal, **32**, 883-887, 1962
- Isayev, Y.N. & Kunitsyn, A.L., 'To the Problem of Satellite's Perturbed Motion Under the Influence of Solar Radiation Pressure, Celestial Mechanics, **6**, 44-51, 1972
- Van der Ha, J.C. & Modi, V.J., 'Solar Pressure Induced Orbital Perturbations and Control of a Satellite in an Arbitrary Orbit', AIAA-77-32, AIAA 15th Aerospace Sciences Meeting, January 1977
- Green, A.J., 1977, “Optimal Escape Trajectories From a High Earth Orbit by Use of Solar Radiation Pressure”, T-652, Master of Science Thesis, Massachusetts Institute of Technology.
- Sackett, L.L., Edelbaum, T.N., 1978, “Optimal Solar Sail Spiral to Escape”, Advances in Astronautical Sciences, AAS/AIAA Astrodynamics Conference, A78 31-901.
- Fekete. T.A., et. al., 'Trajectory Design for Solar Sailing from Low-Earth Orbit to the Moon', AAS-92-184, AAS/AIAA Spaceflight Mechanics Meeting, February 1992
- McInnes C. R., Macdonald M., Angelopolous V., Alexander D., ‘GEOSAIL: Exploring the Geomagnetic Tail Using a Small Solar Sail ’ Journal of Spacecraft and Rockets, Vol. 38, No. 4, July-August 2001.
- Macdonald M., McInnes C. R., ‘Analytic Control Laws for Near-Optimal Geocentric Solar Sail Transfers (AAS 01-472)’, Advances in the Astronautical Sciences, Vol. 109, No. 3, pp. 2393-2413, 2001.
- Alexander D., Sandman A. W., McInnes C. R., Macdonald M., Ayon J., Murphy N., Angelopoulos V, ‘GEOSAIL: A Novel Magnetospheric Space Mission Utilizing Solar Sails’, IAC-02-IAA.11.1.04, 53rd International Astronautical Congress, World Space Congress, Houston, Texas USA, 10-19 October 2002.
- Macdonald M., McInnes C. R., Alexander D., Sandman A., ‘GeoSail: Exploring the Magnetosphere Using a Low-Cost Solar Sail’, Proceedings of Fifth IAA International Conference on Low-Cost Planetary Missions, ESA Special Publication SP-542, pp. 341-349, September 2003.
- Macdonald M., McInnes C. R., ‘Realistic Earth Escape Strategies for Solar Sailing’, Journal

of Guidance, Control, and Dynamics, Jan.-Feb. 2005.

3.10.4 Miscellaneous

- Roy, A.E., Orbital Motion, Adam Hilger, Bristol, 1982
- Battin, R.H., An Introduction to the Methods and Mathematics of Astrodynamics, AIAA Education Series, New York, 1987