Analytical Methodologies for Solar Sail Trajectory Design

By

Mr. Malcolm Macdonald B.Eng. AMRAeS

Dissertation submission to

Department of Aerospace Engineering, University of Glasgow

For the degree of Doctor of Philosophy

ProQuest Number: 10390730

ProQuest.

ProQuest 10390730

– Intentionally Blank –

– Intentionally Blank –

Is tric a bheothaich srad bheag teinne mór.

– Gaelic Proverb [*]

---

[*] *A small spark has often kindled a great fire.*

Intentionally Blank –

With increased international interest in solar sailing for future science missions comes the requirement to generate algorithms for effective orbit design, manoeuvring and control. Previously unexplained seasonal variations in sail escape times from Earth orbit are explained analytically and corroborated within a numerical trajectory model. Simple blended sail control algorithms are developed which are explicitly independent of time and provide near-optimal planetary escape trajectories, while maintaining a safe minimum altitude. It is also shown that the time until escape corresponding to the minimum sail acceleration requirement for shadow free Earth escape is independent of initial altitude.

Traditional trajectory optimisation methods are computationally intensive, requiring significant engineering judgement and experience. A new method of blending locally optimal control laws is thus developed for more complex applications. Each control law is prioritised by consideration of how efficiently it will use the solar sail and how far each orbital element is from its target value. The blended, locally optimal sail thrust vector is thus defined to use the sail efficiently, allowing the rapid generation of near-optimal trajectories. The blending method introduced is demonstrated for a complex orbit transfer about Mercury and for two planet-centred station-keeping applications. The new method is also demonstrated for three different heliocentric scenarios and is shown to closely match, or even out-perform some existing optimisation methods. Furthermore, the method is demonstrated as suitable for rapid mission analysis with an ideal, a non-ideal or optical degradation solar sail force model, while also providing an excellent initial guess for other optimisation methods. The blending algorithms used are explicitly independent of time and as such the control systems are suitable as on-board sail controllers.

– Intentionally Blank –

xiii

The reader should note that while symbols may be used for different purposes it will be clear within the context of a symbol which version is intended. For example, λ is the degradation constant of OSSD model and it is the geocentric east longitude sub-point of the spacecraft, while η is an efficiency factor or it is the cylinder shadow angle; the use of each will thus be clear by the context in which it is used.

Latin Symbols

Symbol	Description	Symbol	Description
A	Surface area	$\mathbf{f}_a$	Force exerted due to absorption
a	Semi-major axis	$\mathbf{f}_e$	Force exerted due to emission by re-radiation
a_{drag}	Atmospheric drag acceleration		
$\mathbf{a}_s, a_s$	Sail acceleration vector	$\mathbf{f}_i$	Force exerted due to incident photons
a_{s_c}	Sail characteristic acceleration	$\mathbf{f}_n, f_n$	Force exerted normal to sail surface
$\tilde{a}$	Absorption coefficient	$\mathbf{f}_r$	Force exerted due to reflected photons, $(= \mathbf{f}_{rs} + \mathbf{f}_{ru})$
$\mathbf{B}, B$	Magnetic component of the incident wave	$\mathbf{f}_{rs}$	Force exerted due to specularly reflected photons
B_b	Non-Lambertian coefficient of rear sail surface	$\mathbf{f}_{ru}$	Force exerted due to non-specularly reflected photons
B_f	Non-Lambertian coefficient of front sail surface	$\mathbf{f}_t, f_t$	Force exerted transverse to sail surface
b	Semi-minor axis	g	Modified equinoctial element
C_D	Drag coefficient	H_0	Reference density scale height
C_L	Lift coefficient	h	Modified equinoctial element
C_n^m	Gravitational harmonic coefficient	h	Height
D	Drag	h_0	Reference altitude
d	Degradation limit in OSSD model	I_0	Frequency integrated specific intensity
$\mathbf{E}, E$	Electric field component of electromagnetic wave	i	Inclination
		$\mathbf{j}$	Induced current in the sail surface
E	Energy	j	Current density induced in a reflector
E	Eccentric Anomaly	K	Eccentric longitude ($= \varpi + E$)
$\mathbf{f}, f$	force	k	Modified equinoctial element
f	Modified equinoctial element	L	True Longitude ($= \Omega + \omega + \nu = \varpi + \nu$)

l	Mean longitude ($= \varpi + M$)	r_a	Radius of apocentre
M	Mean anomaly	r_p	Radius of pericentre
$\mathbf{m}$	Unit vector in the direction of the total sail force	$\hat{\mathbf{r}}$	Unit radial vector from the Sun to the solar sail
m	Mass	$\tilde{r}$	Reflection coefficient
m_0	Rest mass of a body	S_n^m	Gravitational harmonic coefficient
m_p	Mass attached to the sail assembly, i.e. the spacecraft mass	s	Specular reflection coefficient
m_s	Solar sail film mass	s	Auxiliary (positive) variable
N	Normal perturbation component, RTN axes	T	Temperature
		T	Transformation matrix
$\mathbf{n}$	Unit vector normal to sail surface	T	Transverse perturbation component, RTN axes
n	The number of integration steps		
n	Mean Motion ($= \sqrt{\mu/a^3}$)	$\mathbf{t}$	Sail transverse unit vector normal to $\mathbf{u}$
		t	Thickness
O_{ECI}	ECI reference frame	t	Time
O_{RTN}	RTN reference frame	U	Energy density
O_{SUN}	Sun-line reference frame	$\mathbf{u}_i$	Direction of incident photons impinging sail surface
P	Pressure, i.e. the solar radiation pressure	$\mathbf{u}_r$	Direction of specularly reflected photons from sail surface
P	Legendre polynomial with argument $u = \sin\phi$	$\mathbf{u}_i$	Component of incident photons impinging sail surface
p	Momentum	$\mathbf{v}, v$	Velocity
p	Semi-latus rectum	v	Pericentre velocity
R	Radial perturbation component, RTN axes	W	Solar irradiance
R_e	Earth's distance from the Sun	W_σ	Weight of control law σ, or constant by which A^{n}D score is multiplied to find the weight.
R_m	Mercury's distance from the Sun		
$\mathbf{r}, r$	Orbit radius	w	Auxiliary (positive) variable
r_h	Radius of Spacecraft from Sun		

Symbol	Description	Symbol	Description
α	Sail pitch angle	ρ	Density
α	Auxiliary (positive) variable	ρ_0	Reference density
$\tilde{\alpha}$	Pitch angle of the ideal force vector	ς	Critical shadow parameter
α_h	Pitch of the sail with respect to a hypersonic flow	ς_p	Critical penumbra shadow parameter
		ς_u	Critical umbra shadow parameter
β	Sail lightness number	Σ	Solar Radiation Dose
δ	Sail clock angle	$\hat{\Sigma}$	Half-Life Radiation Dose, in OSSD model
ε	Emissivity coefficient		
ε_b	Emissivity coefficient of rear of sail	σ	Arbitrary orbit element
ε_f	Emissivity coefficient of front of sail	σ	Sail loading
η	Efficiency	σ_s	Sail assembly loading, i.e. mass per unit area of the sail assembly
η	Cylinder shadow angle		
λ	Degradation constant of OSSD model	τ	Transmission coefficient
λ	Geocentric east longitude sub-point	τ	Auxiliary (positive) variable
λ_b	Blended thrust vector orientation	Φ	Geopotential function
λ_D	Debye length	φ	Centre-line angle
λ_p	OSSD degradation constant	ψ	Shadow angle
λ_σ	Function of orbit elements / Locally optimal thrust vector orientation	ψ_p	Penumbra shadow angle
		ψ_u	Umbra shadow angle
θ	Sail cone angle	ω	Argument of pericentre
θ	Argument of latitude ($= \omega + v$)	Ω	Longitude of ascending node
θ_0	Angular radius of the solar disc $= \sin^{-1}(R_S / r)$	Ω_{sun}	Longitude of ascending node of solar terminator ground-track
θ_p	Penumbra cone angle	ϖ	Longitude of pericentre
θ_u	Umbra cone angle	ϑ	Angle measured from the first point of Aries to the planet, within the ecliptic plane
v	Frequency of a photon		
v	True anomaly		

Constant	Value	Description
$c\left(=\sqrt{\dfrac{1}{\varepsilon_0\mu_0}}\right)$	2.99792458×10^8 m s^{-1}	Speed of light
G	$(6.6742 \pm 0.0010) \times 10^{-11}$ N m^2 kg^{-2}	Universal gravitational constant
h	6.6262×10^{-27} erg s $= 6.6262 \times 10^{-34}$ J s	Plank's constant
k	1.3807×10^{-23}	Boltzmann's Constant
L_S	3.827×10^{26} W	Solar Luminosity
M_S	1.989×10^{30} kg	Solar mass
R_E	6378.1 km	Earth radii (at Equator)
R_M	2439.7 km	Mercury radii (at Equator)
R_S	696000 km	Solar radii
W_e	1367.6 W m^{-2}	Mean solar irradiance measured at the Earth's distance from the Sun
α_{opt}	35.264 deg	Optimal fixed sail pitch angle
ε	0.01 deg	Obliquity of the Ecliptic (Mercury)
ε	23.45 deg	Obliquity of the Ecliptic (Earth)
ε_0	8.8542×10^{-12} C^2 N^{-1} m^{-2}	Permittivity of free space
μ	3.986032×10^{14} m^3 s^{-2}	Gravitational parameter (Earth)
μ	2.203956×10^{13} m^3 s^{-2}	Gravitational parameter (Mercury)
μ	1.327124×10^{20} m^3 s^{-2}	Gravitational parameter (Sun)
μ_0	1.2566×10^{-6} N A^{-2} m^{-1}	Permeability of free space
$\sigma*$	1.53 g m^{-2}	Critical solar sail loading parameter
$\tilde{\sigma}\left(=\dfrac{2\pi^5 k^4}{15h^3c^2}\right)$	5.670×10^{-8} J K^{-4} m^{-2} s^{-1}	Stefan-Boltzmann constant

Acronym	Meaning
A$^{\pi}$D	Accessibility and Deficit
ASAP-5	Ariane Structure for Auxiliary Payloads, on Ariane 5
CAD	Computer Aided Design
CFRP	Carbon Fibre Reinforced Plastic
DLR	Deutschen Zentrum für Luft- und Raumfahrt (German Centre for Air and Spaceflight)
ECI	Earth-Centred Inertial (axes)
EGM	Earth Gravity Model
ENEAS	Exploration of Near-Earth Asteroids with a Sailcraft
ENEAS+SR	Exploration of Near-Earth Asteroids with a Sailcraft plus Sample Return
ESA	European Space Agency
GEO	Geostationary Orbit
GTO	Geostationary Transfer Orbit
IAE	Inflatable Antenna Experiment
ICBM	Inter-Continental Ballistic Missile
IHP	Interstellar Heliopause Probe
ISAS	Institute of Space and Astronautical Science
ISS	International Space Station
JD	Julian Day
JPL	NASA Jet Propulsion Laboratory
MLI	Multi-Layer Insulation
NASA	National Aeronautics ands Space Administration, the space agency of the U.S.A.
NEP	Nuclear Electric Propulsion
NOAA	National Oceanic and Atmospheric Administration, an agency of the US Department of Commerce.
OAST	NASA Office of Aeronautics and Space Technology
ODISSEE	Orbital Demonstration of an Innovative, Solar Sail driven Expandable structure Experiment
OSSD	Optical Solar Sail Degradation
REP	Radioisotope Electric Propulsion
RTN	Radial, Transverse and Normal (axes)
SCI-A	ESA Payload and Advanced Concepts Office

SEP	Solar Electric Propulsion
SMART	Small Missions for Advanced Research in Technology
SPA	Solar Photonic Assist
SPO	Solar Polar Orbiter
SQP	Sequential Quadratic Programming
SRD	Solar Radiation Dose
SRP	Solar Radiation Pressure
ST	Space Technology, normally followed by a designation number of 5 or over.
TRL	Technology Readiness Level
TRS	Technology Reference Studies

– Intentionally Blank –

– Intentionally Blank –

1   Introduction

Harnessing the power of the Sun to propel a spacecraft may appear somewhat ambitious and the observation that light exerts a force contradicts our everyday experiences. However, it is an accepted phenomenon that the quantum packets of energy which compose Sunlight, that is to say photons, perturb the orbit attitude of spacecraft through conservation of momentum; this perturbation is known as solar radiation pressure (SRP). To be exact, the electromagnetic energy from the Sun pushes the spacecraft and from Newton's second law momentum is transferred when the energy strikes and when it is reflected. The concept of solar sailing is thus the use of these quantum packets of energy, i.e. SRP, to propel a spacecraft, potentially providing a continuous acceleration limited only by the lifetime of the sail materials in the space environment. The momentum carried by individual photons is extremely small. At best a solar sail will experience 9 N of force per square kilometre of sail located in Earth orbit, thus to provide a suitably large momentum transfer one requires the sail to have a large surface area while maintaining as low a mass as possible. Adding the impulse due to incident and reflected photons it is found that the idealised thrust vector is directed normal to the surface of the sail, hence by controlling the orientation of the sail relative to the Sun orbital angular momentum can be gained or reduced. Using momentum gained by reflecting these quantum packets of energy the sail slowly but continuously accelerates to accomplish a wide-range of potential missions.[1]

1.1   An Historical Perspective

Johannes Kepler in 1619 proposed that comet tails are pushed outwards from the Sun due to sunlight. This is one of the first recorded observations that light may exert a force; however the mechanism behind such a force was unclear. In 1690 Christiaan Huygens published *"Traité de la Luminère"* in which he proposed that light travelled as a wave, supported by the observation that two intersecting beams of light did not bounce off each other, as would be expected if they were composed of particles. Fourteen years later Isaac Newton proposed the corpuscular theory of light in his 1704 publication *"Optics"*, believing that light could not be of a wave nature even though in this same work he observed Newton's rings, which were actually an experimental demonstration of the wave nature of light. In many ways Newton's optics work appeared to contradict that of Huygens, yet it is of interest and somewhat ironic to note that these two early theories are qualitatively the same as the current electromagnetic and quantum views of light. In 1754 de Marian and du Fay made the first attempts to experimentally verify and measure radiation pressure. These investigations however proved inconclusive and it wasn't until the beginning of the 20th century that radiation pressure was finally demonstrated in the laboratory. In 1873 James Clerk Maxwell predicted the existence of radiation pressure as a consequence of his unified theory of electromagnetic radiation.[2] Apparently independent of Maxwell, in 1876 Bartoli demonstrated the existence of radiation pressure as a consequence of the second law of thermodynamics. Furthermore, 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 even though the paradox was correctly solved by Maxwell in his 1879 paper *"On stresses in rarefied gases arising from inequalities of temperature"*, in the

Philosophical Transactions. It is of interest that Crookes radiometer was recently used by Thomas Gold when he published a note questioning the concept of light pressure and solar sailing.[†] Gold concluded that solar sailing would not work as it contradicted the laws of thermodynamics, a curious conclusion given that the second law of thermodynamics can be used to demonstrate light pressure. Furthermore, Gold neglected to suitably acknowledge either Reynold's solution or Maxwell's more mathematical solution to the paradox. Moreover, solar radiation pressure has been used in the past to aid the attitude control of various space missions, both in geostationary orbit (GEO) and interplanetary space. Most notably, Mariner 10 used a small *"kite"* (31 cm × 76 cm in area) for manoeuvring by using the pressure of sunlight for attitude control. By using the ballast solar sail for attitude control manoeuvring the Mariner 10 project was able to extend the planned life of the mission and increase mission science returns.[3 - 5] It is of interest to note that the NASA websites which discuss the Mariner 10 mission state the solar wind was used to control spacecraft altitude. This is an error, it was solar radiation pressure.[‡]

The first experimental verification of the existence of radiation pressure and the verification of Maxwell's results came in 1900. At the University of Moscow, Peter Lebedew succeeded in isolating radiation pressure using a series of torsion balance experiments.[6] Nichols and Hull at Dartmouth College, New Hampshire, obtained independent verification in 1901.[7, 8] Around this period a number of science fiction authors wrote of spaceships propelled by mirrors, notably the French authors Faure and Graffigny in 1889. However, it was not until the early 20th century that the idea of a solar sail was accurately articulated. Solar sailing as an engineering principle can be traced back to the Father of Astronautics, Konstanty Eduardowicz Ciołkowski and Fridrikh Arturovitch Tsander, see Figure 1.[9, 10] There is some uncertainty regarding the dates of Ciołkowski's writings on the potential use of photonic pressure for space propulsion. However, it is known that he received a government pension in 1920 and continued to work and write about space. It is within the early part of this period of his life, in 1921 perhaps, which he first conceived of space propulsion using light. Upon the publication of the works of Herman Oberth in 1923, Ciołkowski's works were revised and published more widely, enabling him to gain his due international recognition. Ciołkowski is considered to be the father of cosmonautics and human space flight, his visionary ideas about the future of humanity in space were glorious and far ahead of their time. Much of Ciołkowski's work was performed prior to the first powered flight by the Wright brothers in 1903, although Ciołkowski had designed a monoplane as early as 1894, which subsequently flew in 1915. In the same year as the Wright brothers' first flight Ciołkowski determined correctly that the Earth's escape velocity, into orbit, was 8 km s[-1], using the Rocket Equation he had derived.[11] Inspired by Ciołkowski, Tsander in 1924 wrote *"For flight in interplanetary space I am working on the idea of flying, using tremendous mirrors of very thin sheets, capable of achieving favourable*

[†] Gold's article was posted in June 2003 on a web site of *"e-print physics archives"* http://arxiv.org/html/physics/0306050 and reviewed by New Scientist, Issue 2402, July 2003.

[‡] Further information can be found at http://nssdc.gsfc.nasa.gov/nmc/tmp/1973-085A.html, http://pds.jpl.nasa.gov/planets/welcome/m10.htm and http://www.jpl.nasa.gov/missions/past/mariner10.html.
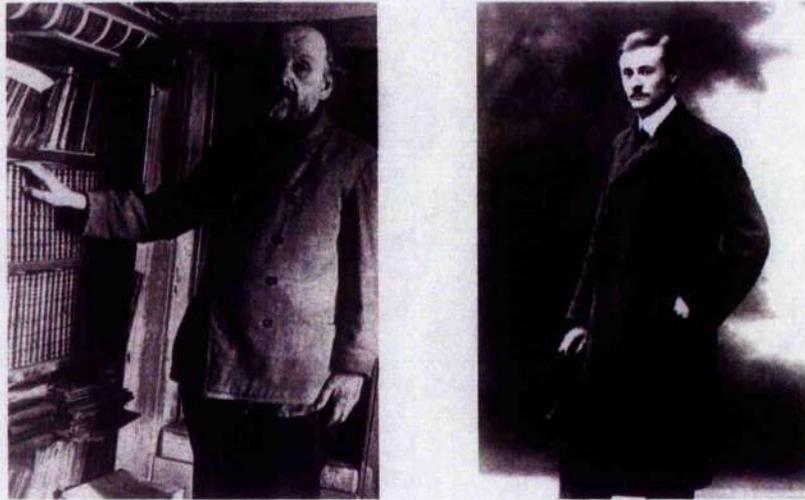
Figure 1 Konstanty Eduardowicz Ciołkowski (left, from www.astronautyka.planty.pl) and Fridrikh Arturovitch Tsander (right, from www.daviddarling.info).

results."[10] Today this statement is widely, though not universally, bestowed the credit as the beginning of solar sailing as an engineering principle.

In 1923 the German rocket pioneer Herman Julius Oberth proposed the concept of reflectors in Earth orbit (Spiegelrakete, or Mirror rocket) to illuminate northern regions of Earth and for influencing weather patterns.[12] This is a similar concept to that behind the Znamya experiments in the early 1990's which will be briefly discussed in Section 1.2. It was this work which caused the works of Ciołkowski to be revised and published more widely. In 1929 Oberth extended his earlier concept for several applications of orbit transfer, manoeuvring and attitude control (Spiegelführung, or Mirror guidance) using mirrors in Earth orbit.[13] This work has a clear parallel with that of Tsander's from 1924.

Following the initial work by Ciołkowski, Tsander and Oberth the concept of solar sailing appears to have remained largely dormant for over thirty years. In the 1950s the concept was re-invigorated and published once again in popular literature, this time in North America. 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.[1, 14] In Reference 14 Wiley discusses 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. In 1958 Richard Garwin, then at the IBM Watson laboratory of Columbia University, authored a solar sail paper in the journal Jet Propulsion where he coined the term "*solar sailing*".[15]

Subsequent to 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.[16 - 23] For a fixed sail orientation several authors have shown that solar sail heliocentric orbits are of the form of logarithmic spirals.[18, 24] A range of applications for solar sailing have been extensively discussed since the 1960's, however planet-centred trajectory design has been largely restricted to escape manoeuvres or relatively simplistic orbit manoeuvring, such as Lunar fly-

by's or orbit inclination change.[25 - 27] In the early 1960's Sands and Fimple investigated planetary escape trajectories for solar sailing using initially circular orbits and analytical techniques which necessitated many simplifying approximations, including a fixed solar position and omitting orbit restrictions such as negative altitude.[19, 21] Fimple used a locally optimal energy-gain control strategy, based on the earlier generalised conclusion by Irving that such a strategy is sufficiently close to the optimal for most low-thrust propulsion systems.[28] It is also worth noting that Lawden mathematically showed for a low-thrust motor that little advantage was to be gained by implementation of a more complex thrust program than aligning the motor thrust with the velocity vector.[29] In the late 1970's this generalised statement about low-thrust propulsion was shown to hold for solar sailing, with an overall efficiency drop of between 1 and 3.5 % for escape from high Earth orbits.[30] In 1978 Sackett and Edelbaum presented optimal Earth sub-escape and orbit transfer manoeuvres for solar sail propulsion, using a method of orbit averaging to reduce the number of computations needed while still giving good performance estimates through the solution of a two-point boundary value problem.[31, 32] During this work a characteristic rapid increase in orbit eccentricity was noted in most trajectories and it was found that often the optimal solution resulted in a negative altitude perigee passage. A minimum altitude constraint was not included in the optimisation tool; however a penalty function was developed to ensure an adequate radius of perigee was maintained throughout the trajectory without significant loss of optimality. Several papers have since been published discussing the use of solar sail propulsion for Earth escape using locally optimal techniques and while the inclusion of a rotating sun-line has become common place, much of the analysis performed continues to make significant simplifications, such as implementing a spherical gravity model or neglecting shadow, 3rd body gravity effects or even setting minimum perigee values.[33, 34] Recently however, a more complete Earth escape trajectory analysis has been performed by Leipold, where the above perturbations were included, while continuing to neglect a minimum altitude constraint.[35] Note also that recent work has produced extremal steering strategies for simulation and optimisation of Earth – Moon transfer trajectories using solar sailing.[36] This work resulted in the solution of a weak stability boundary problem and generated realistic orbital mechanics solutions for the transfer, with the inclusion of all relevant perturbations.

Early 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, though of course the level of assumed technology status is crucial in such comparisons.[37] These 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.[1] In the early 1970's the development of the Space Shuttle and the technological advances associated with deployable structures and thin films suggested that perhaps solar sailing was ready to move beyond paper studies.[38 - 41] In 1974 NASA funded a low-level study of solar sailing at the Battelle laboratories in Ohio which gave positive recommendations for further investigation.[42] The Battelle laboratories recommendations were acted upon at NASA-JPL in an Advanced Mission Concepts Study for Office of Aeronautics and Space Technology (OAST) in FY1976.[43] During the continuation of the Battelle laboratories study Jerome Wright discovered a trajectory that would allow a relatively high-performance

solar sail[§] to rendezvous with comet Halley at its perihelion in the mid-1980's by spiralling towards the Sun and then changing the orbit inclination by almost 180 deg.[44] The flight time of four years would allow for a late 1981 or early 1982 launch, however the required level of performance suggests the study was always over optimistic. A seven to eight year mission had been envisaged using solar-electric ion propulsion, requiring a launch as early as 1977. These positive results prompted NASA-JPL to initiate an engineering assessment study of the potential readiness of solar sailing, following which a formal proposal was put to NASA management on 30 September 1976. At the same time a companion study and technology development program for Advanced Solar Electric Prolusion was initiated in order to allow it to be evaluated as a competitor for the Halley mission. During the initial design study an 800-m per side, three-axis stabilised, square solar sail configuration was envisaged, Figure 2, but then dropped in May 1977 due to the high risks associated with deployment of such a massive structure. The design work progressed to focus on a spin stabilised heliogyro configuration, Figure 2. The heliogyro concept, which was to use twelve 7.5 km long blades of film rather than a single sheet of sail film, had been developed by Richard MacNeal and John Hedgepath.[45-48] The heliogyro could be more easily deployed than the square solar sail by simply unrolling the individual blades of the spinning structure. As a result of this design study the structural dynamics and control of the heliogyro were characterised and potential sail films manufactured and evaluated.[49, 50] As a result of the Advanced Solar Electric Prolusion companion study NASA selected the Solar Electric Propulsion (SEP) system in September 1977 upon its merits of being a less, but still considerable risk for a comet Halley rendezvous.[51] A short time later the SEP rendezvous mission was also dropped due to escalating cost estimates.[52]

1.2   Recent Technology Developments and Activities

A true solar sail has yet to fly (as of 10 May 2005), however significant steps have been taken since the beginning of the 1990's. Furthermore, the Planetary Society has confirmed a launch effort scheduled for 31 May 2005 (as of 10 May 2005) in an attempt to fly the first ever solar sail.

On 4 February 1993, under the guidance of Vladimir Syromiatnikov, the Russian Space Regatta Consortium deployed a 20-m spinning reflector, Znamya 2.5 shown in Figure 3, following the undocking of Progress M-40 from the space station Mir.[1] Znamya 2.5 followed Znamya 2, which was also an in-orbit deployment. Observed from Mir, Znamya 2 and 2.5 showed that spin deployment could be controlled by passive means, while Znamya 2.5 illuminated a spot on the surface of the Earth which would other wise have been in darkness. Znamya is Russian for banner or flag and the experiment 2.5 was a realisation of the idea first proposed by Oberth in 1923 as a Spiegelrakete.

In May 1996 NASA's Special Payloads Division in Goddard Space Flight Centre flew the Spartan mission 207, a large deployable reflector shown in Figure 3, during the Shuttle mission STS-77. The 14-m Inflatable

---

[§] The comet Halley solar sail had a required characteristic acceleration of ≥ 1 mm s^{-2}, see Reference 50. Characteristic acceleration will be defined later in Section 1.3.2.
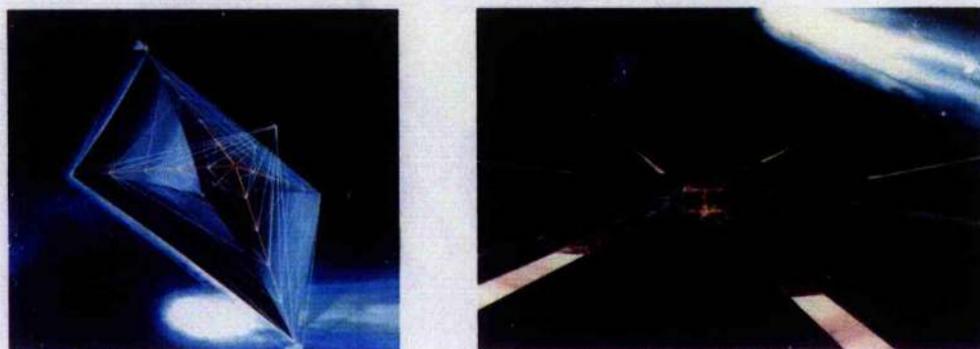
Figure 2 Comet Halley solar sails. Image credit NASA – JPL.

Antenna Experiment (IAE) was primarily designed as a radio frequency reflector, but the promise of inflatable technology towards solar sailing was clearly demonstrated despite mission anomalies.[53]

In Köln in December 1999 the German space agency, DLR, in association with ESA and INVENT GmbH deployed a square 20-m solar sail, shown in Figure 4.[54, 55] This deployment now forms the basis for a future ESA in-orbit deployment demonstration perhaps in early 2006, visualised in Figure 4. Within Europe, to date, solar sail technology development has been driven forward by DLR with the development of sail boom technology, shown in Figure 5.[56] No comparable or competitive design has been significantly funded to date (as of 13 April 2005) by any source within Europe.

Currently, NASA is funding the development of two competing solar sail hardware designs, one by Able Engineering, now part of ATK and the other by L'Garde, both of which underwent 10-m deployment tests in vacuum chambers in mid-2004, as shown in Figure 6.[57 - 59] In April 2005 the ATK sail design underwent a 20-m environmental test using a 2 μm film substrate, also shown in Figure 6. The ATK April 2005 test was conducted in a large thermal vacuum chamber under ambient space conditions at NASA's Glenn Research Center. The sail attitude control system was also tested. The L'Garde sail design is scheduled to be similarly tested during the summer of 2005 (as of 10 May 2005). The two NASA funded sail concepts will be developed to technology readiness level (TRL) 6 then one concept will be selected as the preferred NASA solar sail design. This approach allows NASA to ensure it develops a scaleable technology that can be demonstrated and provide heritage for future science missions. NASA is also funding other hardware component development studies and software models, towards the New Millennium Program's Space Technology 9 Project, ST9.[60] ST9 will be a system-level, technology-validation experiment having a launch date in the 2007 to 2008 time frame.

In August 2004 the Institute of Space and Astronautical Science, ISAS, in Japan deployed two small solar sails in space from an S-310 sounding rocket, as shown in Figure 7. The long-term aim however of the ISAS work is not solar sailing, rather the deployment of large structures to enable high-power SEP as part of a hybrid system working in co-operation with the solar sail. None the less, the technology cross-over and similarities are obvious for pure solar sailing and the stated short-term aim is the development of solar sail deployment capability.[61 - 66]
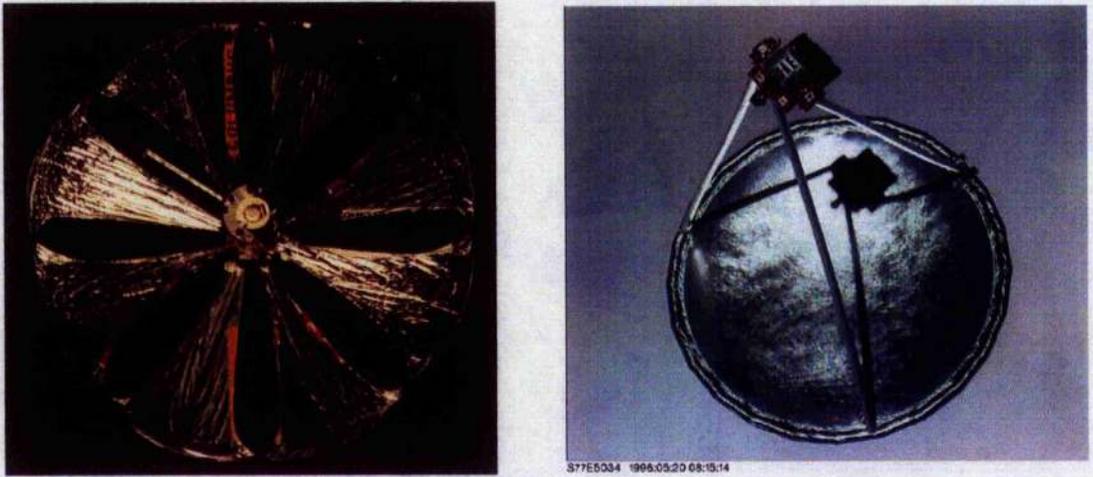
Figure 3 Znamya 2.5, viewed from Mir (left, Image credit Russian Space Regatta Consortium) and the Inflatable Antenna Experiment, viewed from STS-77 (right; Image credit NASA / L'Garde).
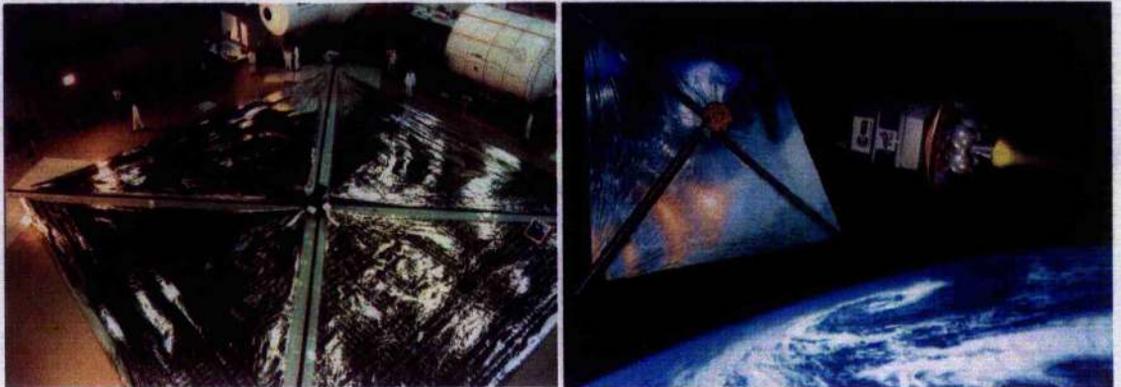
Figure 4 DLR solar sail ground deployment test (left) and visualisation of in-orbit deployemnt test (right). Image credit DLR.

Figure 5 DLR developed carbon fibre reinforced plastic solar sail booms. Image credit DLR.
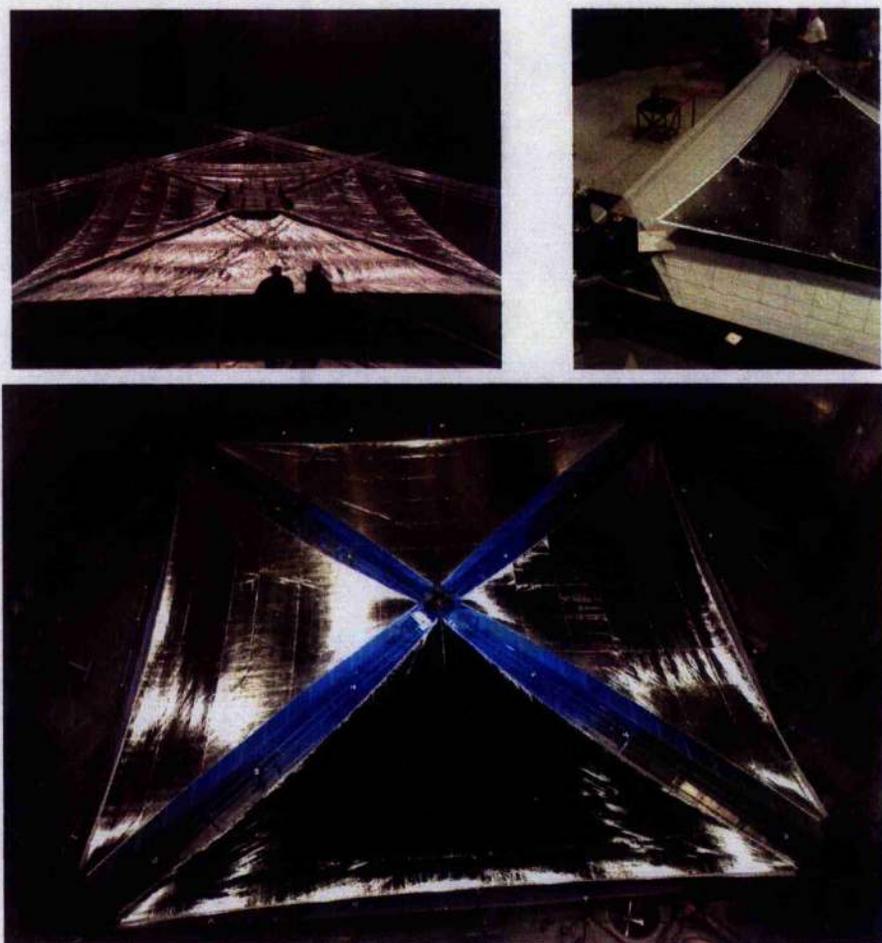
Figure 6   10-m solar sail deployment test performed by L'Garde in the 30-m vacuum chamber at NASA's Glenn Research Centre, Plum Brook Station (top-left).   Able Engineering (now ATK) 10-m sail test in the 15-m vacuum chamber at NASA's Langley Research Centre (top-right).   ATK 20-m sail test in the 30-m vacuum chamber at NASA's Glenn Research Centre, Plum Brook Station (bottom).   Image credit NASA.

Figure 7   ISAS sail deployment on S-310 sounding rocket.   Image credit ISAS.

In addition to the work of government funded agencies several private enterprises are seeking to advance solar sailing as a viable propulsion system. Most notably among these is the Planetary Society's COSMOS-1 solar sail, Figure 8, which as stated above is scheduled for launch on 31 May 2005. The COSMOS-1 sail will be launched from a submerged Russian submarine in the Barents Sea on-board a Volna rocket, a converted inter-continental ballistic missile (ICBM). The Planetary Society's COSMOS-1 solar sail is currently likely to be the first solar sail flight.[**]

1.3   Solar Radiation Pressure

The observation that light exerts a force 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 as noted by Kepler. 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.[1, 35] 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. In certain conditions the grains spiral close to the Sun causing them to melt; 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, thus ejecting dust into interstellar space. Note that this is similar to the Solar Photonic Assist (SPA) trajectory as will be briefly discussed in Section 1.5.2 and then in much more detail in Section 6. In the current section the physical description of the momentum transfer process associated with solar radiation pressure will be
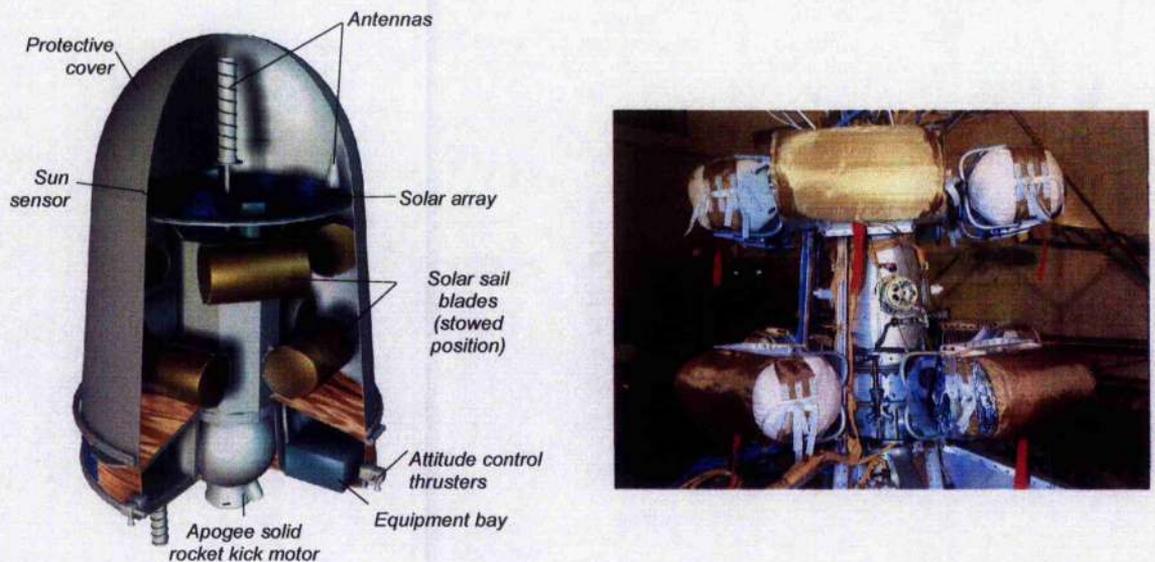
Figure 8 The Planetary Society's COSMOS-1 solar sail, annotated CAD model (left) and actual spacecraft (right). Image credit L. Friedman and the Planetary Society.

[**] More information can be found at the Planetary Society web-site; http://www.planetary.org/ and at the COSMOS-1 tracking web-site; http://www.solarsail.org/

discussed. Following this discussion, issues such as the standard inverse square law for solar radiation pressure with solar radius and its limitations at low radii will be examined. Furthermore, solar sail performance characteristics and different solar sail force models will be discussed.

1.3.1   Electromagnetic and Quantum Descriptions

In the electromagnetic description of light, momentum is transported by electromagnetic waves. The electric field component of the wave induces a current in the sail, as shown in Figure 9. The magnetic component of the incident wave generates a Lorentz force, $\mathbf{j} \times \mathbf{B}$, in the direction of propagation of the wave. The induced current results in the generation of another electromagnetic wave, which is essentially the reflection of the incident wave. A wave propagating along the x-axis exerts a force on a current element given as

$$df = j_z B_y \, dx \, dy \, dz \qquad\qquad\qquad \text{Equation 1}$$

where j_z is the current density induced in the surface of the reflector along the z-axis shown in Figure 9. The ensuing pressure on the current element can then be written as

$$dP = j_z B_y \, dx \qquad\qquad\qquad \text{Equation 2}$$

where dP is defined as the force per unit area. From Maxwell's equations of electrodynamics the current term in Equation 2 can be replaced by field terms. It can thus be demonstrated that the time average pressure is given as

$$\langle dP \rangle = -\frac{\partial}{\partial x}\left[\frac{1}{2}\varepsilon_o E_z^{\,2} + \frac{1}{2\mu_o}B_y^{\,2}\right]dx \qquad\qquad\qquad \text{Equation 3}$$

Allowing the energy density for the electric component and magnetic component of the incident wave to be defined as

$$U = \frac{1}{2}\varepsilon_o E^2 + \frac{1}{2\mu_o}B^2 \qquad\qquad\qquad \text{Equation 4}$$
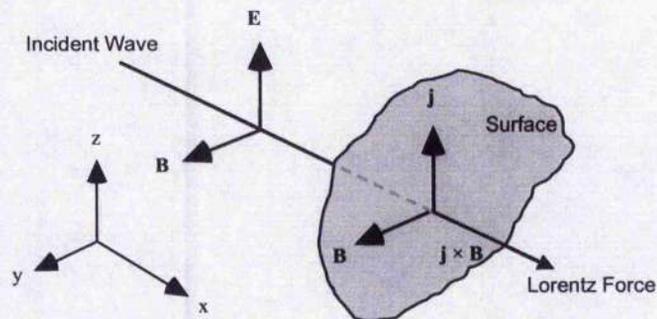
Figure 9 Electromagnetic radiation pressure.

The pressure exerted on a surface of thickness Δl is then obtained by integrating Equation 3

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

Equation 5

The pressure exerted on the surface is given by the total energy density of the electromagnetic wave for a perfectly absorbing medium, such that $\langle P \rangle = \langle U \rangle$.

Consider two plane waves separated by Δx and incident on a surface of area A, as shown in Figure 10. The volume between the two waves impinging on the surface is then $A\Delta x$. Further, the spacing 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 wave is thus

$$U = \frac{\Delta E}{A(c\Delta t)}$$

Equation 6

where, ΔE is the energy contained within the volume element. Additionally, the energy flux across the surface can be written

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

Equation 7
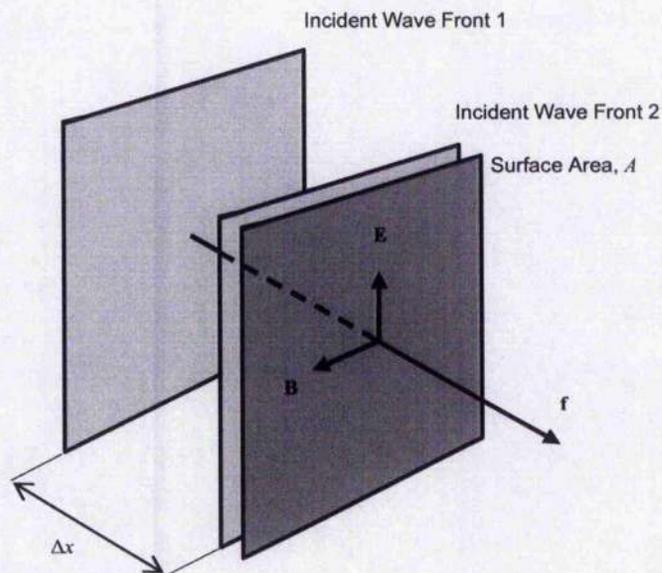
Figure 10 Energy density of an electromagnetic wave.

Therefore,

$$U = \frac{W}{c}$$

Equation 8

Thus, for the electromagnetic description of light the radiation pressure is simply the energy density of the electromagnetic wave.

Using quantum mechanics radiation pressure can be visualised as momentum transported by photons impacting and then reflecting a surface. The term "*photon*" was coined by Gilbert N. Lewis in a letter to Nature magazine, in 1926.[67, 68] From Plank's Law, a photon of frequency v will transport the energy given by

$$E = h v$$

Equation 9

Using special relativity the total energy of a moving body may be written as

$$E^2 = m_o^2 c^4 + p^2 c^2$$

Equation 10

Since a photon has zero rest mass its energy may be written as

$$E = pc$$

Equation 11

Using the photon energy defined by Equation 9 and Equation 11, the momentum transported by a single photon is

$$p = \frac{h v}{c}$$

Equation 12

The pressure on a body is found through consideration of the momentum transported by a flux of photons. At distance r from the Sun the energy flux may be written in terms of the solar luminosity, L_S, and scaled by the Sun–Earth distance, giving

$$W = \frac{L_S}{4\pi R_e^2} \left[\frac{R_e}{r} \right]^2 = W_e \left[\frac{R_e}{r} \right]^2$$

Equation 13

From Equation 13, the energy ΔE transported across a surface of area A, normal to the incident radiation, in time Δt is given by

$$\Delta E = W A \Delta t$$

Equation 14

From Equation 11, the energy then transports momentum Δp,

$$\Delta p = \frac{\Delta E}{c}$$

Equation 15

The pressure on the surface is thus defined as the momentum transported per unit time, per unit area, such that

$$P = \frac{1}{A}\left[\frac{\Delta p}{\Delta t}\right]$$

Equation 16

Accordingly, using Equation 14 the pressure exerted on the surface due to momentum transport by photons is

$$P = \frac{W}{c}$$

Equation 17

This is the same expression as Equation 8. For a perfectly reflecting surface the actual pressure is twice the value given by Equation 17, as momentum is transferred by incident photons and by reflected photons, following Newton's second law. The solar radiation pressure exerted on a solar sail at the Earth's mean distance from the Sun, 1 AU, may now be calculated using Equation 17 and following Reference 1. As the orbit of the Earth about the Sun is slightly elliptical, solar irradiance at the Earth varies by approximately 3.5 % over the year. An accepted mean value of the solar constant is 1367.6 J s^{-1} m^{-2}.[1] Thus, the pressure exerted on a perfectly reflecting solar sail at 1 AU is taken to be 9.12×10^{-6} N m^{-2}.

1.3.2   Force on a Perfectly Reflecting Solar Sail

The acceleration experienced by a solar sail is a function of the attitude of the sail reflective surface with respect to the Sun. For a solar sail, as shown in Figure 11, the force exerted on the surface due to incident photons is given by

$$\mathbf{f}_i = PA(\mathbf{u}_i.\mathbf{n})\mathbf{u}_i$$

Equation 18

The reflected photons will exert a force of equal magnitude on the surface, but in the specular reflected direction, $-\mathbf{u}_r$.

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

Equation 19

Noting, $\mathbf{u}_i\text{-}\mathbf{u}_r = 2(\mathbf{u}_i.\mathbf{n})\mathbf{n}$, the total force exerted on the solar sail is given by

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

Equation 20

Using Equation 13 and Equation 17 the total force may be written as

$$\mathbf{f} = \frac{2AW_e}{c}\left[\frac{R_e}{r}\right]^2 (\mathbf{u_i} . \mathbf{n})^2 \mathbf{n}$$

Equation 21

The solar sail performance is quantified by the total spacecraft mass per unit area ($^m/_A$) and is called the sail loading.[1] The sail loading is an important solar sail design parameter. The sail pitch angle is defined as the angle between the sail normal and the incident radiation, as shown in Figure 11 and Figure 12. Using these
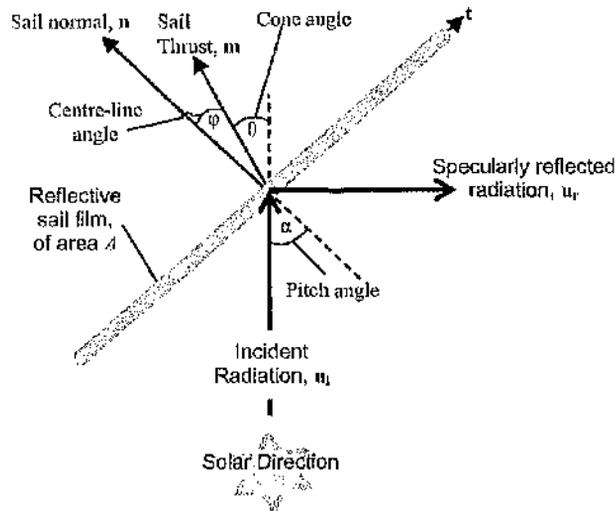
Figure 11 Incident and reflected components on ideal solar sail.
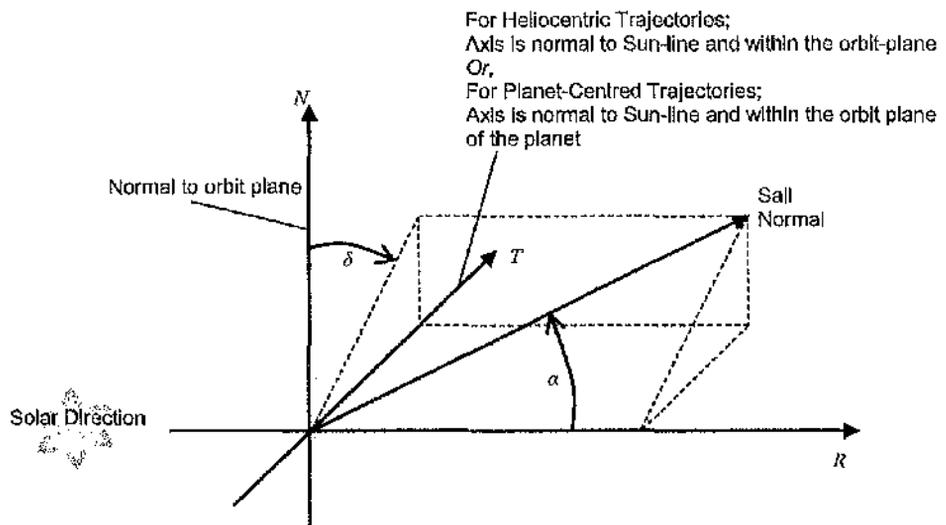
Figure 12 Orientation of the sail pitch and clock angles in Sun-sail line reference frame. Note, for heliocentric trajectories the Sun-sail line reference frame is coincident with the radial, transverse and normal reference frame of the solar sail.

definitions the solar sail acceleration may now be written as

$$\mathbf{a}_s = \frac{2W_e}{c} \frac{1}{\sigma} \left[\frac{R_e}{r} \right]^2 \cos^2 \alpha \, \mathbf{n}$$

Equation 22

The characteristic acceleration is defined as the actual acceleration experienced by the sail at a solar distance of 1 AU with the sail normal to the Sun, such that $\alpha = 0$.[1] The characteristic acceleration is a parallel design parameter to sail loading and may be written as

$$a_{s_c} = \frac{9.12\eta}{\sigma \left[g \ m^{-2} \right]} \quad \left[mm \ s^{-2} \right]$$

Equation 23

where, an overall efficiency factor is used to account for the finite reflectivity of the sail film. Typically the total solar sail efficiency is of order 0.85 - 0.9.[1, 69] It is important to note that the efficiency of a sail does not alter its characteristic acceleration; however it will alter the physical dimensions of the sail. The solar sail acceleration may also be written in terms of the solar gravitational acceleration as

$$\mathbf{a} = \beta \frac{GM_S}{r^2} \left(\hat{\mathbf{r}}.\mathbf{n} \right)^2 \mathbf{n}$$

Equation 24

The dimensionless sail parameter β is defined as the ratio of the solar radiation pressure acceleration to the solar gravitational acceleration.[1] This parameter is called the sail lightness number. The solar radiation pressure acceleration and the solar gravitational acceleration are both assumed to have an inverse square variation, thus the lightness number is independent of the Sun-sail distance. Using Equation 13, Equation 22 and Equation 24 the solar sail lightness number may be written as

$$\beta = \frac{\sigma^*}{\sigma}$$

Equation 25

$$\sigma^* = \frac{L_S}{2\pi GM_S c}$$

Equation 26

The critical solar sail loading parameter, σ^*, is 1.53 gm^{-2}.[1] This constant is unique within our solar system and is a function of the solar mass and the solar luminosity, which have assumed constant values. With a mass per unit area equal to the critical loading parameter the lightness number is one, however such a sail loading is an extremely challenging requirement. 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 McInnes.[1]

1.3.3   Force on a Non-Perfectly Reflecting Solar Sail

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 represented more realistically through consideration of optics theory; considering sail reflection, absorption and re-radiation. Several assumptions are made within the optical force model which compromises its accuracy, including that the sail is perfectly flat. Note however that this assumption has recently been removed through development of new methodology for the analytical description of the force and moment generated by a solar sail of arbitrary shape, while continuing to describe the surface optical properties through the optical force model.[70] To account for the assumptions within the optical force model a numerical parametric force model was developed by JPL for the Halley rendezvous mission.[1, 69] Note however that both models, theoretical and measured, apply standard optical theory where non-specular reflections at greater than 6 – 10 deg are assumed to be of no use. A solar sail can however utilise all the reflected photons no matter what angle they are reflect at, as all reflected photons will exert a force of some magnitude on the sail surface. Recent work has been performed which shows the assumptions made within both the traditional models significantly compromise the results, as reflection is highly symmetric about the specular line with *"reflected force components normal to the specular reflection line cancelling".*[71] Note further that the same work concluded an aluminium coated sail under zero tension would remain 88 % reflective, a reduction in only 2 % from the nominal ideal reflectivity, even if heavily creased.

Sail optical properties and hence the sail force model will vary through the duration of a mission due to thermal cycles, radiation degradation and so forth.[72 - 74] The effect of optical solar sail degradation on solar sail trajectories will be considered in Section 6. Despite the concerns discussed above, the optical force model will on occasion be used within this dissertation as it is currently as good as any other non-ideal sail force model in the public domain. Note that use of a non-ideal sail force model within this dissertation is not intended to provide a highly accurate model of a real sail trajectory, rather to demonstrate the ability of the orbit design methods generated in later sections to manage a more complex sail force model. Thus, through successful utilisation of the optical force model one can demonstrate the ability of the proposed methods to adapt to any other complex sail force model, while providing an initial assessment of the effect of a non-ideal sail. Similarly, the optical solar sail degradation model, which is based on the optical force model, is used to demonstrate the capabilities of the proposed method while providing an initial assessment of the effect of optical surface degradation on trajectory design.

1.3.3.1   Optical Solar Sail Force Model
The total force due to solar radiation pressure may be divided into its component parts and written as

$$\mathbf{f} = \mathbf{f}_r + \mathbf{f}_n + \mathbf{f}_c$$

Equation 27

The main optical properties of the sail film can be defined within the constraint $\tilde{r} + \tilde{a} + \tau = 1$. However, since the transmission coefficient, τ, is zero on the reflecting side of the solar sail it follows that

$$\tilde{a} = 1 - \tilde{r}$$

Equation 28

The unit vectors shown in Figure 11 can be related using,

$$\mathbf{u_i} = \cos\alpha \, \mathbf{n} + \sin\alpha \, \mathbf{t} \qquad\qquad \text{Equation 29}$$

$$\mathbf{u_r} = -\cos\alpha \, \mathbf{n} + \sin\alpha \, \mathbf{t} \qquad\qquad \text{Equation 30}$$

Combining Equation 29 and Equation 30 it can be seen that

$$\mathbf{u_r} = \mathbf{u} - 2\cos\alpha \, \mathbf{n} \qquad\qquad \text{Equation 31}$$

Initially assuming all incident photons are absorbed by the sail the force exerted on the solar sail is given by $PA\cos\alpha \, \mathbf{n}$. Resolving this force into normal and transverse components using Equation 29 it is found that

$$\mathbf{f_a} = PA\left(\cos^2\alpha \, \mathbf{n} + \cos\alpha\sin\alpha \, \mathbf{t}\right) \qquad\qquad \text{Equation 32}$$

A fraction $\tilde{r}$ of the incident photons are now reflected. Of this fraction, another fraction will be specularly reflected in direction $\mathbf{u_r}$, so providing a force $\mathbf{f_{rs}}$ in direction $-\mathbf{u_r}$ given by

$$\mathbf{f_{rs}} = -(\tilde{r}s)PA\cos\alpha \, \mathbf{u_r} \qquad\qquad \text{Equation 33}$$

A further fraction of photons are uniformly scattered from the reflecting surface of the sail due to non-specular reflection. This component generates a force $\mathbf{f_{ru}}$ in direction $\mathbf{n}$ given by

$$\mathbf{f_{ru}} = B_f\tilde{r}(1-s)PA\cos\alpha \, \mathbf{n} \qquad\qquad \text{Equation 34}$$

The total force due to reflected photons is thus $(\mathbf{f_{rs}} + \mathbf{f_{ru}})$. Using Equation 30 to write the total force in terms of the normal and transverse directions yields

$$\mathbf{f_r} = PA\left\{\left[\tilde{r}s\cos^2\alpha + B_f(1-s)\tilde{r}\cos\alpha\right]\mathbf{n} - \tilde{r}s\cos\alpha\sin\alpha \, \mathbf{t}\right\} \qquad\qquad \text{Equation 35}$$

Finally, one must consider photons which have been absorbed and then re-emitted as thermal radiation from both the front and back surfaces of the sail. The emitted power from a unit area of the sail at temperature T is $\varepsilon\tilde{\sigma}T^4$. Assuming the sail has uniform temperature and allowing for the non-Lambertian nature of the front and back sail surfaces the force due to emission by re-radiation is found to be

$$\mathbf{f_e} = \frac{\tilde{\sigma}T^4}{c}\left(\varepsilon_f B_f - \varepsilon_b B_b\right)\mathbf{n} \qquad\qquad \text{Equation 36}$$

The sail temperature is obtained from the balance between the thermal input and the thermal output of $(\varepsilon_f + \varepsilon_b)\tilde{\sigma}T^4$, as shown in Figure 13. The thermal input is given by $(1-\tilde{r})W\cos\alpha$ and since the pressure is $W/_c$ the sail temperature may be written as

$$T = \left[\frac{(1-\tilde{r})cP\cos\alpha}{\tilde{\sigma}(\varepsilon_f + \varepsilon_b)}\right]^{1/4}$$

Equation 37

Substituting for the sail equilibrium temperature,

$$\mathbf{f_e} = PA(1-\tilde{r})\frac{\varepsilon_f B_f - \varepsilon_b B_b}{\varepsilon_f + \varepsilon_b}\cos\alpha\,\mathbf{n}$$

Equation 38

Equation 38 shows that a low emissivity sail backing is required to maximise the normal force, however the emissivity of the backing is the only coefficient which can be used to control the sail temperature. The total force, due to all components, may thus be written in terms of normal and transverse components as

$$\mathbf{f_n} = PA\left\{(1+\tilde{r}s)\cos^2\alpha + B_f(1-s)\tilde{r}\cos\alpha + (1-\tilde{r})\frac{\varepsilon_f B_f - \varepsilon_b B_b}{\varepsilon_f + \varepsilon_b}\cos\alpha\right\}\mathbf{n}$$

Equation 39

$$\mathbf{f_t} = PA(1-\tilde{r}s)\cos\alpha\sin\alpha\,\mathbf{t}$$

Equation 40

The total force vector and magnitude are consequently written as

$$\mathbf{f} = f\mathbf{m}$$

Equation 41

$$f = \left(f_n^{\,2} + f_t^{\,2}\right)^{1/2}$$

Equation 42
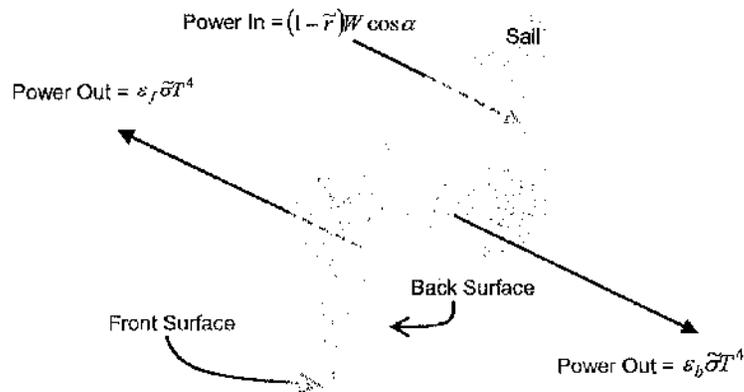
Figure 13 Solar sail thermal balance.

The direction of the force, or thrust vector will not be normal to the sail surface as in the ideal force model discussed in Section 1.3.2. The thrust due to absorbed photons is somewhat greater than that due to reflected photons; therefore the thrust vector is biased towards the direction of incident radiation as shown in Figure 11. The angle of the thrust vector relative to the incident radiation is defined by the cone angle, θ.[1] The angle between the thrust vector and the sail normal is the centre-line angle given by

$$\tan \varphi = \frac{f_t}{f_n}$$ Equation 43

An important effect of the centre-line angle for the optical solar sail model is that it limits the thrust vector to a maximum cone angle of 55.5 deg corresponding to a sail pitch angle of 72.6 deg, for the optical coefficients given in Table 1. This is in contrast to the ideal solar sail where the thrust vector is always directed normal to the sail surface and can in principle be directed up to 90 deg from the Sun-line. This limitation on the direction of the thrust vector poses constraints for some solar sail applications.[1] Furthermore, application of the optical surface degradation model in Section 1.3.4 results in a reduction of the maximum cone angle as will be shown in Section 6.2.5.[72 - 74] The sail optical parameters derived from the NASA – JPL comet Halley rendezvous studies are listed in Table 1.[1, 69] These are the coefficients which are used within this dissertation when applying a non-ideal sail force model and are the initial values used within the optical degradation model.

1.3.4   Optical Solar Sail Degradation

Spacecraft in the interplanetary environment experience a number of effects due to the Sun: surface charging (either positive or negative) and/or deep dielectric charging that can result in damaging arcs, UV and radiation effects on materials, and plasma wake or sheath effects that could impact experiments or instruments on the spacecraft (see Section 1.4.3). All surfaces which are exposed to the space environment are subject to continual degradation from a number of sources. For spar supported solar sails, micrometeorite impacts will cause only local damage due to the high impact velocities. As the micrometeorite passes though the sail it creates effective rip-stops; singeing the edges of the hole it is creating. However, for spinning solar sails the centrifugal tension in the sail film may require the use of rip-stops to prevent tear propagation, due to the additional loads on the sail film. The integration of rip-stops during the manufacture process could be achieved through many different means.[1] In addition to particulate damage, solar radiation can degrade plastics through weakening and reduction of tensile strength. While some work has been performed in this field much more work remains to be conducted to allow accurate and optimal mission design.[75 - 79] Furthermore, radiation and particle damage will inevitably alter the optical characteristics of the sail reflective surface, this has to date been studied very little though some effort is now being conducted into this area of research.[72 - 74]

The initial optical solar sail degradation (OSSD) model assumes that the optical coefficients do not depend on the sail temperature; however this is a planned extension to current work.[72] The OSSD model also assumes that the optical coefficients do not depend on the light incidence angle and that no self-healing

Parameter	Symbol	Value
Reflectivity Coefficient	$\tilde{r}$	0.88
Specularly Reflected Coefficient	s	0.94
Emissivity of Front Surface	ε_f	0.05
Emissivity of Back Surface	ε_b	0.55
Non-Lambertian Coefficient of Front Surface	B_f	0.79
Non-Lambertian Coefficient of Back Surface	B_b	0.55

Table 1   Initial optical coefficients for a non-ideal sail.

effects occur in the sail film.[72] Letting p be an arbitrary optical coefficient from P, where P represents the set of optical coefficients defined in Table 1. With OSSD p becomes $p(t)$, which is a function of the solar radiation dose (SRD). SRD is defied as

$$\Sigma(t) = \int_{t_0}^{t} \dot{W} \cos\alpha \, dt = W_e R_e^2 \int_{t_0}^{t} \frac{\cos\alpha}{r^2} \, dt$$

Equation 44

The OSSD model assumes that $p(t)$ has an exponential form, varying from $p(t_0) = p_0$ to $p(t_\infty) = p_\infty$, thus

$$p(t) = p_\infty + (p_0 - p_\infty)\exp(-\lambda_p \Sigma(t))$$

Equation 45

where the degradation constant is related to the Half-Life SRD by

$$\lambda_p = \frac{\ln 2}{\hat{\Sigma}_p}$$

Equation 46

In order to reduce the number of free parameters the initial OSSD model introduces a degradation limit, which together with a single half-life SRD for all p allows the free parameters to be suitably reduced; explicitly $\hat{\Sigma}_p = \hat{\Sigma} \forall p \in P$. Furthermore, the reflectivity of the sail will decrease with time, thus becoming more matt, which in-turn will mean front emissivity increasing with time. The problem thus reduces to

$$\tilde{r}_\infty = \frac{\tilde{r}_0}{1+d}$$

Equation 47

$$s_\infty = \frac{s_0}{1+d}$$

Equation 48

$$\varepsilon_{f_\infty} = (1+d)\varepsilon_{f_0}$$

Equation 49

It can also be shown (see Reference 72) that the optical force model is least sensitive to the three remaining optical coefficients in Table 1, thus $\varepsilon_{b_\infty} = \varepsilon_{b_0}$, $B_{f_\infty} = B_{f_0}$, $B_{b_\infty} = B_{b_0}$. The variation of the optical coefficients with solar radiation dose is shown in Figure 14 for a degradation limit of 20 % (i.e. $d = 0.2$) and Half-Life SRD of 0.5 W_e yr. The effect of OSSD on solar sail trajectory design will be considered in Section 6.

1.3.5   Radiation Pressure with a Finite Solar Disc

The variation of the solar radiation pressure with distance from the Sun can be approximated by an inverse square relationship, see Equation 21. However, this assumption breaks down close to the Sun when the finite angular size of the solar disc must be considered.[1, 80]

The modelling of the source of radiation pressure is distinct and independent from the modelling of solar radiation pressure force on the sail. Initially assume the solar disc to be a Lambertian surface. 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 can be written as

$$P(r) = \frac{2}{c} \int_0^\infty \int_0^{2\pi} \int_0^{\theta_0} I \cos^2 \theta \, d\Omega \, d\nu \, , \quad d\Omega = \sin\theta \, d\theta \, d\phi \qquad \text{Equation 50}$$

Noting geometric symmetry about the azimuth, see Figure 15, and that the specific intensity is independent of solar distance, Equation 50 reduces to

$$P(r) = \frac{4\pi}{c} I_o \int_{\xi_0}^1 \xi^2 d\xi \, , \quad \xi = \cos\theta \, , \quad \xi_o = \cos\theta_0 \qquad \text{Equation 51}$$
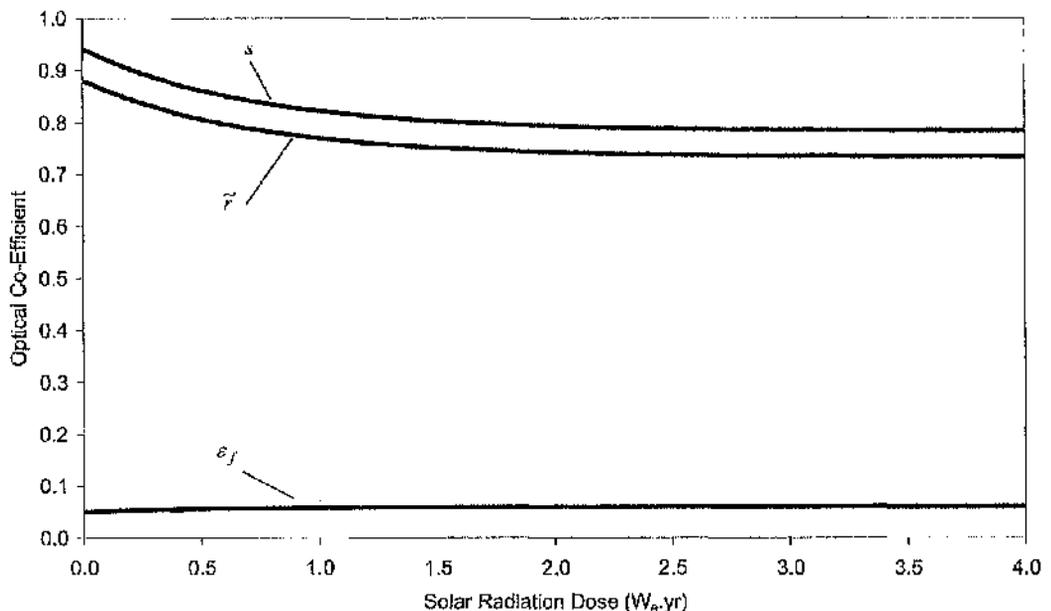
Figure 14 Variation of optical coefficients with solar radiation dose.
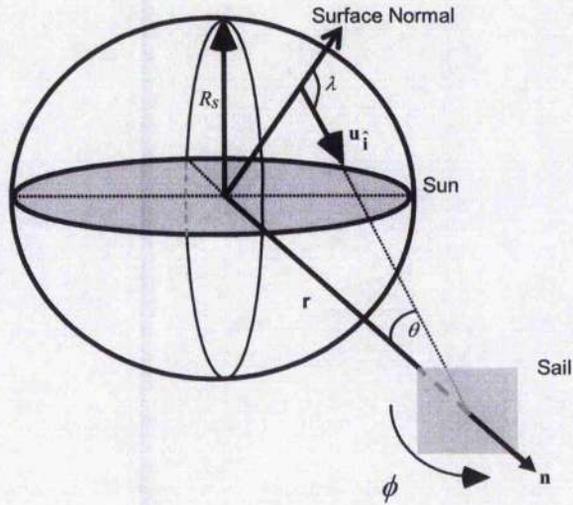
Figure 15 Solar radiation pressure due to a finite solar disc.

Performing this integration and substituting for ξ_0,

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

Equation 52

Which expanding in powers of $(R_S / r)^2$ and for $r >> R_S$, to the first order gives,

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

Equation 53

At large solar distances this expansion must match asymptotically with the expression for the radiation pressure from a distant point source.

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

Equation 54

Comparing Equation 53 and Equation 54 allows the frequency integrated specific intensity to be identified as

$$I_0 = \frac{L_S}{4\pi^2 R_S^2}$$

Equation 55

Substituting for I_0 in Equation 52,

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

Equation 56

A more useful way of expressing this is in terms of the usual inverse square law, $P^*(r)$.[1]

$$P(r) = P(r)^* F(r)$$

Equation 57

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

Equation 58

The function $F(r)$ describes the deviation of solar radiation pressure from an inverse square law. $F(r)$ attains a minimum value at $r = R_S$, where $F(R_S) = \frac{2}{3}$. As $r \to \infty$, $F(r) \to 1$ as the solar disc becomes more like a point-source. From Figure 16 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. The deviation from the inverse square approximation is due to photons from the solar limb striking the sail at an oblique angle, while photons from the centre of the disc strike along the sail normal. At large solar distances photons from all parts of the solar disc are incident along near-parallel rays.[1]
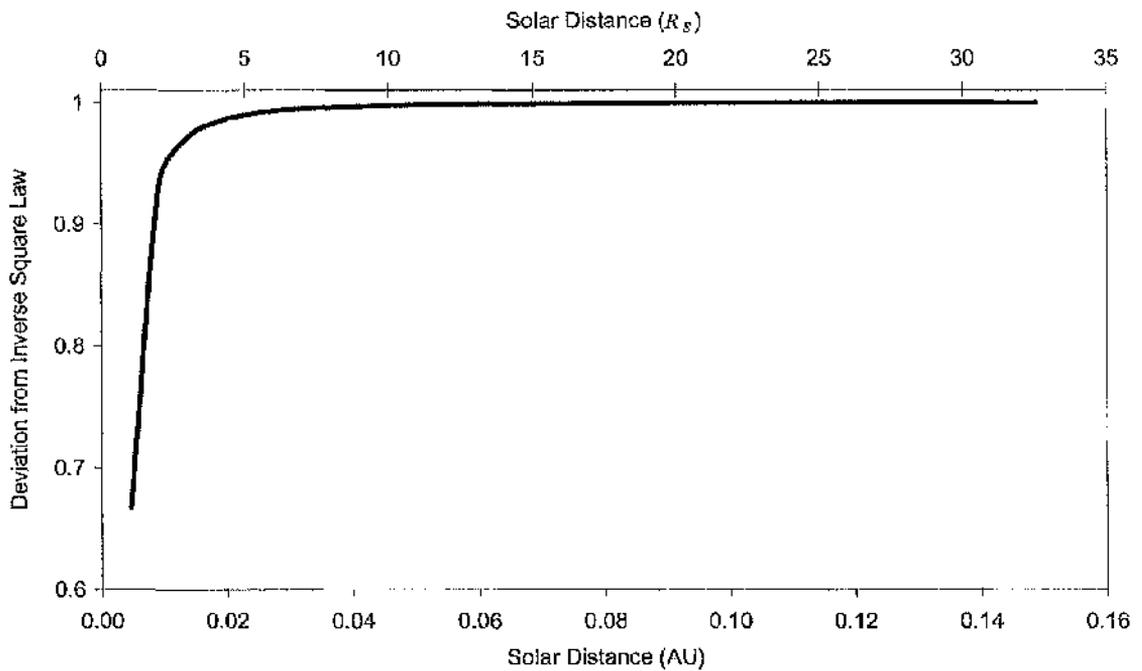
Figure 16 Deviation of the uniformly bright finite disc model from an inverse square model.

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.[1] 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.[1] For the limb darkened specific intensity it is found that $F(R_S) = 0.708$, such that the limb darkened solar radiation pressure deviates less from an inverse square form than the pressure from a uniformly bright disc.

1.4   Solar Sail Design

Note from Section 1.6 that this dissertation does not address solar sail hardware design issues; as such the current section is intended only to provide a context for the work contained within later sections. One of the key problems in solar sail design is the packing and subsequent deployment of large areas of thin film. The dimensional expansion ratio between a deployed and stowed solar sail can be over 100, thus innovative structural engineering solutions are required. The packing and deployment problem has perhaps been one of the greatest impediments to practical solar sail utilisation.[1] In addition, since the sail is folded for packing, the reflecting medium of the sail must be mounted on a thin substrate. The presence of a substrate leads to a fundamental limitation on solar sail performance due to the parasitic mass, defined as the total non-reflective mass of the solar sail and attached spacecraft, which this substrate represents. The conventional belief is that the solar sail film must be as flat as possible to maintain as high a reflectivity as possible. If this is true then tension must be applied to the deployed sail, either by a deployable structure, by spin-induced tension or a combination of both. It is recalled however, from Section 1.3.3 and Reference 71, that this may be a misconception and that a solar sail could conceivably be heavily wrinkled and remain suitably reflective. If this is true then the deployable structure or spin-induced tension is simply required to support the reflective surface, allowing it to be oriented such as to direct the resultant thrust vector. It is immediately clear that this distinction is of critical importance, if a sail film can be wrinkled then the structure need only support and not tension the film, if the film cannot be wrinkled a much heavier structure is required to apply a tension. Once deployed the sail film must be oriented to direct the solar radiation pressure force for orbit manoeuvring. Due to the large moments of inertia of solar sails innovative engineering solutions are again required.

1.4.1   Design Parameters

The primary design parameter for a solar sail is its characteristic acceleration, or its sail loading which is an equivalent parameter. These parameters determine the transfer time to a particular target object or even whether a particular class of orbits are possible.[1] The characteristic acceleration and sail loading are ideal parameters for discussion of solar sail astrodynamics. However, they are a function of both the efficiency of the solar sail design and the mass attached to the sail assembly (i.e. the spacecraft on-board). As such these parameters are less suitable for discussion of solar sail hardware design.

Recall from Section 1.3.2, the characteristic acceleration is defined as the actual acceleration experienced by the sail at a solar distance of 1 AU with the sail normal to the Sun, such that $\alpha = 0$. At this distance from the Sun the magnitude of solar radiation pressure exerted on a perfectly absorbing surface is 4.56×10^{-6} N m^{-2}.

Allowing for finite sail efficiency the characteristic acceleration is given by

$$a_{s_c} = \frac{2\eta P}{\sigma}, \quad \sigma = \frac{m}{A}$$

<div align="right">Equation 59</div>

Recall, the efficiency of a sail does not alter its characteristic acceleration; however it will alter the physical dimensions of the sail. It is for this reason that characteristic acceleration is less suitable for discussion of solar sail hardware design. Equation 59 is similar to Equation 23, from Section 1.3.2. The sail efficiency is a function of the optical properties of the sail film and the sail shape. The total mass of the solar sail can be split into two components, the sail assembly (composed of the sail film, booms, housing, et cetera) and the mass attached to the sail assembly (i.e. the spacecraft on-board). Note that often within solar sailing this spacecraft is referred to as the "*payload*",[1] thus using payload in the launch vehicle perspective, however such use can be confusing as typically the payload on a space vehicle is the science instruments. Consequently, within this dissertation the term payload will not be used to refer to the mass attached to the sail assembly in the form conventionally used, rather one will refer to this as the "*spacecraft*". The term spacecraft is more suitable as it has become clear recently that the optimal form of solar sail system design would be to make the sail a jettisonable sub-system of the spacecraft, thus the sail will normally only be controllable while attached to the spacecraft.[81] The term "*payload*" is thus an anomaly. The characteristic acceleration of the solar sail may now be written as

$$a_{s_c} = \frac{2\eta P}{\sigma_s + \left(m_p/A\right)}, \quad \sigma_s = \frac{m_s}{A}$$

<div align="right">Equation 60</div>

The sail assembly loading, defined in Equation 60, is the primary hardware design parameter for a solar sail; allowing a measure of the performance of the sail film and the efficiency of the solar sail structural design. Using Equation 60 the influence of various design parameters on the solar sail characteristic acceleration can be found. For a fixed sail area and efficiency Equation 60 becomes a function of the sail assembly loading and the payload mass. It can be shown that for large values of σ_s the solar sail characteristic acceleration is relatively insensitive to variations in the spacecraft mass.[1] This is due to the sail film and structural mass dominating the total mass of the solar sail. Conversely, for a large payload mass, the characteristic acceleration becomes insensitive to variations in the sail assembly loading.[1] Hence, a high characteristic acceleration is only gained if the sail assembly mass and the payload mass are small. In other words, the parasitic mass must be minimised in order to maximise sail performance. This is of course a logical conclusion, the important trade is thus whether to invest energy in designing an efficient high performance solar sail or to invest energy in developing a low mass miniaturised spacecraft to then be transported by solar sail. It appears that in-fact each is as important as the other and thus equal emphasis is required to develop solar sail technology towards flight status.[82]

1.4.2   Solar Sail Films

The sail film reflective layer is supported by a substrate. The substrate is required principally to allow handling, folding, packing and deployment. The substrate must be coated with a suitable reflecting material

for efficient photon reflection, typically aluminium is currently favoured.[1, 50, 55, 57 - 59, 69, 71]   A further front coating, such as Silicon Oxide, may also be required to reduce pre-launch oxidation of the reflecting surface with a resultant loss of reflectivity.[1, 69]   Alternatively a UV induced sublimation layer could be added to prevent pre-launch oxidation; such a layer would thus add no mass to the actual solar sail flight mass. The sail substrate must have sufficient strength so that it does not fail and create tears which may propagate during deployment or when fully deployed and under tension, if tension is required following deployment. Furthermore, since the reflective coating on the sail film will not have perfect reflectivity, a fraction of the incident solar radiation will be absorbed by the substrate. This absorbed energy must be dissipated; this can be achieved through a thermally emitting rear surface coating. The choice of a suitable, high emissivity coating is yet another design decision. However, it is recalled from Equation 38 that potentially a sail backing of low emissivity is required to maximise the sail thrust normal to the sail surface. Kapton® with a front aluminium coating has a rear emissivity of order 0.34 which is too low to provide passive thermal control for inner solar system missions.[1] However thin-film chromium, with an emissivity of order 0.64 appears to be a suitable candidate for a rear surface coating.[1, 69]

The sail substrate contributes a significant proportion of the total sail assembly mass, particularly for a large sail where the sail substrate mass dominates the sail's parasitic mass breakdown. The production of very thin films with good mechanical and thermal properties is thus central to solar sail realisation. There is extensive industrial experience of the manufacture, coating and handling of thin films for a number of ground and space applications. For example, primary spacecraft insulation is typically provided by multi-layer insulation (MLI) blankets which are constructed of alternative layers of aluminium coated Mylar® or Kapton® and a thin net of material such as nylon, Dacron®, Nomex®[††] or bridal veil.[83] Note however that currently the typical thickness of commercially available thin films is excessive for moderate performance solar sails. Mylar® however is commercially available down to a thickness of only 0.9 μm, but has low resistance to solar UV radiation and so is unsuitable for long duration exposure without double-sided coatings. Several thin film materials have been considered as potential sail substrates. The optimal sail film until recently was generally considered to be Kapton®.[1, 50, 69] Kapton® does not have a melting point as such, however it does suffer a phase transition (glass transition temperature) above approximately 680 K.[1] A safe, long term maximum operating temperature for solar sail applications is generally considered to be between 520 K and 570 K.[1] It is this thermal limit which gives rise to the widely accepted minimum solar radius of 0.25 AU for solar sailing, although of course this does not take into account the thermal limit of the sail booms and other structural components, nor does it account for the thermal limits of the attached spacecraft. An all aluminium sail film, that is to say one with no substrate, actually has a very similar minimum solar distance even though bulk aluminium has a much higher melting point.[84] The production of sail film of order 2 μm has been recently identified as a key technology requirement of mid-term solar sailing.[85] Such thin films are not routinely used for large volume commercial purposes, mainly due to the difficulty in handling during manufacture. In addition to solar sails, other space applications such as solar concentrators and space

---

[††] Mylar®, Kapton®, Nomex® and Dacron™ are all trade names of E.I. DuPont de Nemours & Co. Although one notes that "*dacron*" (with a small *d*) is often used as a generic form.

telescope Sun shades also require films thinner than commercially available Kapton®. To this end NASA and SRS Technologies have produced Clear Plastic-1 (CP-1) film down to a gauge of order 2 μm. CP-1 film has very similar properties to Kapton® film and as such is highly suitable for solar sail applications. CP-1 film is now generally accepted to be the sail film of choice and was used at 3 μm and 2 μm gauge in the 2004 Able Engineering and ATK sail deployment tests respectively, as shown Figure 6.[59]

1.4.3   Space Environmental Effects

The sail may acquire a differential electric charge between the front and rear surfaces due to exposure to UV radiation and the solar wind.[1] Surface charging and/or dielectric charging typically result in arcing which is a potential source of sail film structural failure. To prevent electrical discharges from the front surface to the rear of the sail both surfaces of the sail must be in electrical contact. The sail surfaces must also be grounded to the spars, stay lines and any other structural components. A further consequence of electrical charging is that the sail may form a "*bubble*" within the solar wind plasma.[1] Within this region, field and particle instruments would not be able to obtain accurate measurements.

Prior analysis of the interaction between the sail and local environment is limited, so definitive statements are unwise. Analysis of prior studies on the interactions between the sail and the local environment at 1 AU do however allow some limited extrapolation. Note however that this work is preliminary and forms only the first step of a much larger research program that is currently underway within NASA-JPL; thus many considerable assumptions are made.[86] Garrett and Wang (Reference 86) found the sail to be surrounded by a plasma sheath within which the potential is positive compared with the ambient plasma and followed by a separate plasma wake, which is negative relative to the plasma. This structure departs dramatically from a negatively charged plate such as might be found in the Earth's ionosphere on the night side where both the plate and its negative wake are contiguous. Furthermore, at 1 AU the plasma sheath in the ram side starts at a distance of ~2 λ_D. Notably, the sail size appears to have minimal impact on the plasma sheath, although the potential in the wake region is significantly different. Garrett and Wang concluded that although the plasma sheath at 1 AU extends to a distance of ~50 m in front of the sail, its effects on the solar wind electron measurement made near the sail surface should be minimal.[86] However, the sheath may have some adverse effects on solar wind proton measurements made within the sheath. An additional problem may be the outgassing/contamination cloud created as the sail approaches the Sun, potentially generating a "*pre-sheath*" due to the particles and material around the sail that extends out to at least one characteristic body length independent of the plasma models. No evidence of this has however been found, except for observations from the Shuttle and other similar cases. Potentially, there is more complex physics than just the solar wind plasma interactions.[87]

Consideration of the effect of the sail on the magnetometer environmental conditions is even more difficult to determine. The NASA-JPL study on plasma effects also intends to analyse magnetic field "*pile-up*" in front of a large sail but as yet this work has not commenced.[86] It is thought however that the potential exists for the creation of a self-generated contamination cloud entrapped around the sail, with perhaps a small

"*magnetosheath*" like a comet on the size of the largest dimensions of the sail.[87]   Clearly, this is a major source of concern for future space science missions, for example the Solar Polar Orbiter[‡‡] (SPO) and Interstellar Heliopause Probe[‡‡] (IHP) missions would be significantly enhanced if science data could be collected while attached to the solar sail. Furthermore, missions such as GeoSail,[‡‡] Geostorm[‡‡] and other Non-Keplerian orbit applications require the science data be collected while the sail is attached to the spacecraft, therefore if this is not possible such missions become unworkable with solar sail propulsion. This issue is thus critical to the realisation of solar sail technology for use in real-world science missions.

1.4.4   Solar Sail Structures

During launch the solar sail must be stowed in a small volume consistent with the launch vehicle payload fairing or fairing volume allocation, for example the DLR ground test in December 1999 sized the deployment module to fit the Ariane Structure for Auxiliary Payloads, on Ariane 5 (ASAP-5).[55, 88] Note however that most future sail missions would not have such stringent launch volume constraints as the ASAP-5 systems, instead being constrained to a volume similar to that of, say, a Soyuz-ST launch fairing.[89] Following a successful launch the solar sail must be deployed from its stowed configuration, this process must be reliable, controllable and predictable. Knowledge of deployable structures in space for applications such as experimental booms and solar arrays can be adapted and extended for solar sailing.[1] One such system is the CoilABLE booms developed by Able Engineering of California, now ATK, which has significant flight heritage. Triangular elements are joined to longerons to form the truss. Pre-tensioned diagonal elements store enough potential energy to allow self-deployment, although a lanyard cable attached to the end of the truss can be used to control the deployment rate via a damper or motor, with the boom tips rotating during deployment. An alternative form of deployment is from an internally-threaded canister shell which extrudes the boom.   This method allows the boom to be at near-full strength throughout the deployment and also has a retraction ability, which would aid sail deployment in the event of an anomaly. These benefits come at the expense of a larger stowage volume and heavier booms; however there is no rotation of the boom during deployment which would enable sail film and booms to be simultaneously deployed. The CoilABLE booms typically have a stowed length of order 1 – 2 % of the final deployed length and can be scaled to over 100 m in length. Able Engineering / ATK used the CoilABLE booms during their sail deployment test in mid-2004 and April 2005, Figure 6.[58, 59] Recall from Figure 5 that DLR has developed carbon fibre reinforced plastic (CFRP) solar sail booms. The structure is deployed using the potential energy stored in the pre-stressed flattened tube, or using a small drive motor for a more controlled deployment. By using carbon fibre with layers built-up in alternate directions deployable booms can be manufactured with essentially zero coefficient of thermal expansion.[56] Once the sail is deployed the drive motor, housings and associated hardware can be jettisoned in order to reduce the total mass of the system.

In addition to deployable mechanical structures, inflatable structures are an attractive means of reliable deployment. Inflatable structures have long been considered for solar concentrators, antenna reflectors and truss structures.[53, 90] The main benefit of inflatables is the ease and reliability of deployment with few failure

---

[‡‡] see Section 1.5 for further description of mission.

modes. The structure consists of a thin membrane which is deployed solely by internal gas pressure. Once the structure begins to deploy the internal gas pressure ensures full deployment and rigidisation. A space curing resin enclosed between two films in the inflatable membrane can be used to ensure rigidisation throughout the mission duration. This is the approach taken by L'Garde for the sail deployment demo in mid-2004, Figure 6.[57] The L'Garde booms build on flight heritage gained from Spartan mission 207, the IAE, Figure 3.[53]

1.4.5   Solar Sail Configurations

The essential requirement of any solar sail design is to provide a large, reflective surface with minimal parasitic mass that can be easily and reliably deployable and manufactured. The choice of sail configuration is dependent on the mission requirements with many potential configurations available, see Reference 69. The three primary configurations are the square, disc and heliogyro sails. Initially the heliogyro appears a very efficient configuration and it was this that led to the square sail being dropped by NASA-JPL during the comet Halley studies in favour of the heliogyro. However, the square sail had been extensively analysed by this stage and when later comparison was made on an equal basis the square sail was found to perform significantly better than the heliogyro.[69] It was found that very efficient structures could be developed for the square sail, while the heliogyro needed far more load-carrying members than previously anticipated. The square sail has also been favoured to-date due to its relative simplicity and ability to provide rapid turn rates, which are required for efficient planetary escape and capture spirals. It is noted however that the square sail architecture critically limits the sail performance due to the large parasitic mass required and as such may not prove suitable for far term mission applications, such as IHP. A schematic diagram of the three identified primary configurations is shown in Figure 17. It is interesting to note that the solar sail design chosen by The Planetary Society, Cosmos 1 Project, uses none of these designs. Rather the Cosmos 1 sail is a combination of all three configurations shown in Figure 17. The Cosmos 1 configuration has eight triangular sails, see Figure 18, rather than the four in a square sail. This large number of sails gives the appearance of a compact heliogyro and allows for attitude control through rotation of each sail, or blade, independent of the next, exactly like a heliogyro. Furthermore, the sail is spin stabilised, like a disc sail, so as to tension the film in each blade and further reduce parasitic mass. The Cosmos 1 sail is an interesting compromise of the benefits offered by each design configuration.

Square Sail   Heliogyro   Disc Sail
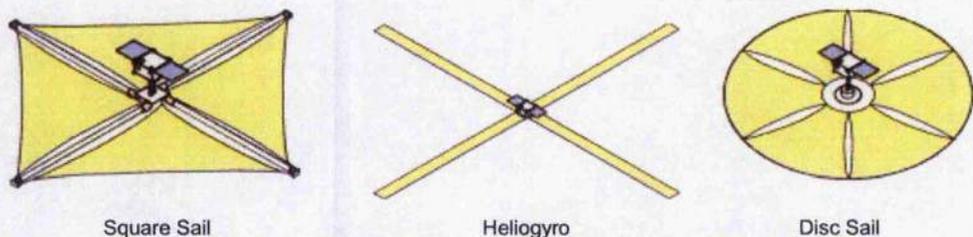
Figure 17 Solar sail design concepts, not to scale. Image credit NASA.
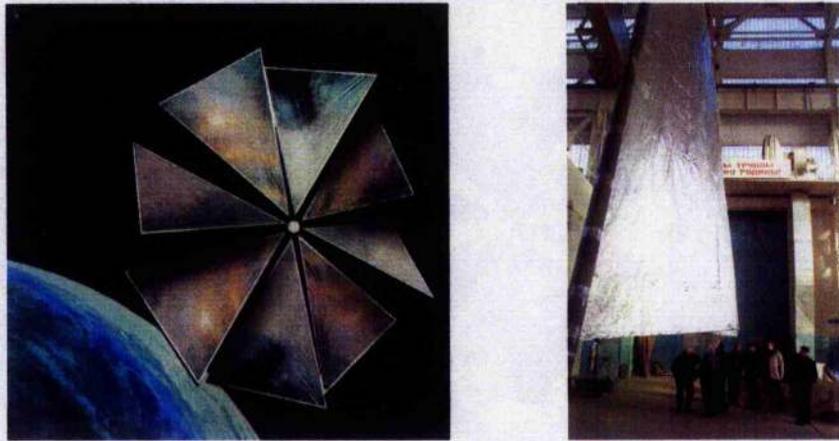
Figure 18   Cosmos 1 Sail (left) and zero-g deployment test of sail (right).   Image credit L.Friedman and the
Planetary Society.

1.5   Solar Sail Mission Applications

It is clear that solar sailing is currently undergoing a revival in interest and a renewed drive towards flight
status which could, with the correct will, lead to the realisation of a true solar sail science driven mission
within the decade. In addition to the technology developments of the last 15 years there has been an increase
in the number of potential solar sail mission applications.   To give a context of where each mission
application addressed within the dissertation may lie with respect to another, a brief review of potential solar
sail mission applications follows.

1.5.1   Solar Sail Demonstration and Near-Term Mission Applications

The ESA SMART (Small Missions for Advanced Research in Technology) program and the NASA New
Millennium program are specifically designed to flight test new technologies, allowing heritage and
confidence in design for future larger and more expensive missions such as ESA's Cornerstone Missions.
The New Millennium program has long been established as the most likely development route of solar
sailing within NASA, with solar sailing tendering proposals for ST-5, ST-7 and now ST-9.[91, 92] ST-5 and ST-
7 proposals envisaged Geostorm as the first solar sail mission.[93 - 95] The Geostorm mission concept provides
real-time monitoring of solar activity. It would operate sunward of the Earth's L_1 point, thus increasing the
warning time for geomagnetic storms.   The Geostorm mission is envisaged as an operational spacecraft
providing dedicated solar monitoring and as such is an excellent candidate mission as the required science
instrument mass is low. By imparting a radial outward force from the Sun the solar radiation pressure in
effect reduces solar gravity and allows the L_1 point to be moved sunward. As sail performance is increased
one can further "*reduce*" solar gravity and thus provide enhanced solar storm warning.[1] The conceptually
simple nature of the Geostorm mission is complicated by mission requirements, risk and budget factors and
by the unstable nature of sub-L_1 orbits. Some recent work on station keeping at sub-L_1 locations has shown
promise in this area[96,97] and it appears that station-keeping should be possible, although currently the required
sail pointing accuracy is excessive. Due to the newly realised complexity of Geostorm, ST-9 work has
focused on a GEO Disposal Orbit, at GEO plus 300 km, for a sail demonstration mission which would have

no scientific goals and would instead be designed purely for engineering purposes.[98] Geostorm is now seen within NASA as a follow on to this initial demonstration mission, which will have validated the ground generated results and model predictions. Thus, providing confidence in the sail systems for the Geostorm mission, nominally the first true solar sail mission as envisaged by NASA in cooperation NOAA and the United States Department of Defence.

In addition to the development of boom and other sail technologies DLR generated the Orbital Demonstration of an Innovative, Solar Sail driven Expandable structure Experiment (ODISSEE) proposal.[99] ODISSEE is intended to fly a 40-m square solar sail in geostationary transfer orbit (GTO) as an engineering experiment, performing little or no science. However, the mission goal of raising orbit energy, perhaps for a lunar flyby, is significantly compromised due to the existence of air-drag around perigee (see Section 3.4.5). DLR have also proposed a series of follow-on solar sail missions, such as Exploration of Near-Earth Asteroids with a Sailcraft (ENEAS) and ENEAS plus Sample Return (ENEAS-SR).[55, 100] However, the problem with such low-energy missions is that they can be performed easily by conventional propulsion for similar or less cost. It is thus unlikely that such missions will ever feature on a realistic solar sail mission roadmap. The ESA funded in-orbit deployment demonstration, visualised in Figure 4 and scheduled for early 2006, is an excellent opportunity to demonstrate the deployment capabilities of a ground matured and scaleable solar sail concept. An in-orbit deployment demonstration should be a demonstration, not an experiment, thus it can provide suitable heritage towards future science driven and solar sail enabled missions by validating ground generated model results. The DLR solar sail development program has stalled recently due to a lack of funding from the German Federal Government and the ODISSEE proposal now seems more distant than ever. Similarly, the in-orbit deployment demo increasingly appears an end in itself rather than the beginning of a process due to the lack of ground generated data and models, making the in-orbit deployment an experiment rather than a demonstration. This, coupled with a potential inability to confidently scale the design up to the very large area and low mass sails required to enable future science driven missions means that the current planned in-orbit demo may not provide suitable heritage to reduce the risk of future missions.

Solar sailing is an elegant concept, however it must be pulled forward by mission applications at the same time as it is pushed by technology development. A technology is rarely adopted within engineering simply because it is an elegant one. This also holds true for initial flight tests of solar sailing. Unless such flight tests provide confidence in the technology and a clear path towards some enabling capability, they will not perform a useful function. Thus, the use of low cost sounding rockets, as used by ISAS, to test multiple sail deployment mechanisms during the short period of free-fall allow for several tests of scaled prototypes at the same cost as a single launch to orbit. By spreading the risk over several tests the inevitable unforeseen single point failures of deployment can be identified prior to flying a full-scale demonstration mission. The demonstration mission could take many forms, however it was noted by McInnes[1] "...*it seems that what is required is a small, low cost and low risk solar sail mission for which there is either no feasible alternative form of propulsion or alternative option of comparable cost.*" It was thought until recently that Geostorm provided this mission, however on analysis the concept has proven to be of higher risk than originally believed. A mission concept however has been proposed recently which meets the requirements identified

by McInnes for a demonstration mission; this mission is called GeoSail and was first proposed in October 2000 by Macdonald and McInnes.[82, 101 - 106] GeoSail is motivated by the desire to achieve long residence times in the Earth's magnetotail, enabling high resolution statistical characterisation of the plasma in a region subject to a variety of external solar wind conditions. This is accomplished by the novel application of a solar sail propulsion system to precess an elliptical Earth-centred orbit at a rate designed to match the rotation of the geomagnetic tail, the orientation of which is along the Sun-Earth line above the night side of the Earth. Conventional, inertially fixed orbits with an apogee inside the geomagnetic tail will provide less than three months of science data due to the rotation of the geomagnetic tail with the Sun-Earth line in an inertial reference frame. It has been shown that the requirements to precess such orbits by chemical propulsion are prohibitively large, while electric propulsion significantly curtails the potential mission duration. The GeoSail orbit designed to achieve these science goals has a perigee located above the planetary dayside at approximately 11 R_E, corresponding to alignment with the magnetopause. Apogee is aligned within the geomagnetic tail reconnection region on the night-side of Earth, at 23 R_E. A key feature of the GeoSail orbit is the ability to investigate the near-downstream region over an extended period. Conventional missions have achieved extended observation times only in the deep tail by executing double-Lunar flybys to precess the orbit apse-line. The utilisation of a small solar sail allows orbit apse-line precession without the requirement of going as far as the Moon, at approximately 60 R_E, thus enabling extended study of this key region of the near-tail. The level of required sail performance to match the apse-line precession with the Sun-Earth line rotation is found to be very similar to the level of performance anticipated for a solar sail demonstration mission and is less than required for the Geostorm mission. Furthermore, the sail required for GeoSail can be significantly less complex than the Geostorm sail, as no active station keeping is required. For this reason GeoSail has recently been identified by ESA's Payload and Advanced Concepts Office (SCI-A) as a potential solar sailing SMART mission, as it allows technology demonstration while also enabling new and novel science.[106] Trajectory design for GeoSail will be discussed some more in Section 5 of this dissertation.

The interest of SCI-A represents the technology being pulled forward by mission applications, with interest stemming from a wish to study mission concepts which cannot be performed without solar sailing, that is, missions which are enabled by solar sailing. Following an extensive two-year study of solar sail mission applications ranging from inner-planet sample return to a heliopause probe mission and even beyond, the missions which are enabled or significantly enhanced by solar sailing were identified. These missions are all high-energy missions and typically require either very close solar passes, or spend the majority of the mission within the inner solar system. Thus, for example, while a Jupiter mission is high-energy it is also outer solar system and it was thus found that chemical propulsion was a better option for Jupiter missions than solar sailing.[107] Such results contradict traditional thinking, which believed solar sailing was well suited to Jupiter exploration and even for Europa exploration. SCI-A have recently introduced Technology Reference Studies (TRS) to focus the development of strategically important technologies of likely relevance to future science missions. This is accomplished through the study of technologically demanding and scientifically interesting missions, which are not part of the current ESA science programme. The TRS cover a wide range of mission profiles with an even wider range of strategically important technologies. All TRS mission profiles are based on small satellites, with miniaturised highly integrated payload suites, launched on a Soyuz Fregat 2-1b.[108]

Science missions are technologically very challenging. It is thus important to define and prepare critical technologies far in advance, ensuring they are developed in a timely manner and that associated cost, risk and feasibility of potential future mission concepts can be properly estimated. The TRS are set up to provide a set of realistic requirements for these technology developments far before specific science missions get proposed by the scientific community. Through their study a set of detailed requirements for technology development activities can be determined for missions in the mid to far-term. Currently two TRS require the development of solar sail propulsion, the Solar Polar Orbiter, (SPO)[81, 109] and the Interstellar Heliopause Probe (IHP).[110, 111] These missions will be discussed in Section 1.5.2. It appears that it is these studies and activities which will now drive forward the development of solar sail technology within Europe, with GeoSail perhaps providing the first step on the roadmap.

A recently proposed near-term roadmap for solar sailing suggested a potential path from on-orbit deployment demonstration to the mid-term solar sail missions, such as the Solar Polar Orbiter.[82] The near-term road map is illustrated in Figure 19. Identifying the technology progression from in-flight demonstration(s) through each mission in turn, the roadmap allows the prior mission to act as a step towards the next, leading to the realisation of much more technologically complex missions in the future. Note, further information on the Polar Observer mission can be found in References 1 and 82.

1.5.2 Mid to Far Term Solar Sail Mission Applications

Many of the truly exciting mission concepts which are enabled by solar sail propulsion require the technology to be in a mature state, with heritage gained from earlier demonstration and low-cost scientific missions such as GeoSail and Geostorm. The mid to far term missions generating the most current interest within Europe are SPO and IHP respectively, due to the TRS within SCI-A. The primary objective of the SPO mission is the delivery of a spacecraft into an orbit with inclination close to 90 deg with respect to the solar equator. The spacecraft orbit should be phased such that once on-station it will remain near to the solar limb from a terrestrial perspective. The spacecraft should also be positioned on an orbit interior to that of the Earth's.[81, 109] Note the spacecraft will jettison the sail prior to the beginning of science operations. A 1998 study from NASA-JPL, Solar Polar Sail Mission, also considered the use of solar sail technology to place a science payload into a true solar polar orbit.[112, 113] This mission has been studied by many authors and presents an excellent potential mission concept for solar sail propulsion.[35, 114] The SPO mission and its trajectory design will be discussed some more in Section 6 of this dissertation.

Several missions to the heliopause and beyond have previously been studied using many different propulsion systems. [110, 111, 115 - 124] As in all propulsion trades the optimal propulsion system depends on the technology level assumed and the mission constraints imposed. The use of SEP is limited by power and propellant availability and as such requires multiple revolutions about the Sun to minimise the gravity losses.[125] The use of nuclear electric propulsion (NEP) necessitates the use of a large launch vehicle as it is difficult to scale down reactors beyond a minimum mass and volume.[110, 117] Radioisotope electric propulsion (REP) requires high-efficiency radioisotope power sources of greater than 10 W kg^{-1} and extensive thrusting into the outer solar system[110, 116] which would likely dilute science returns from the mission due to the interaction of the
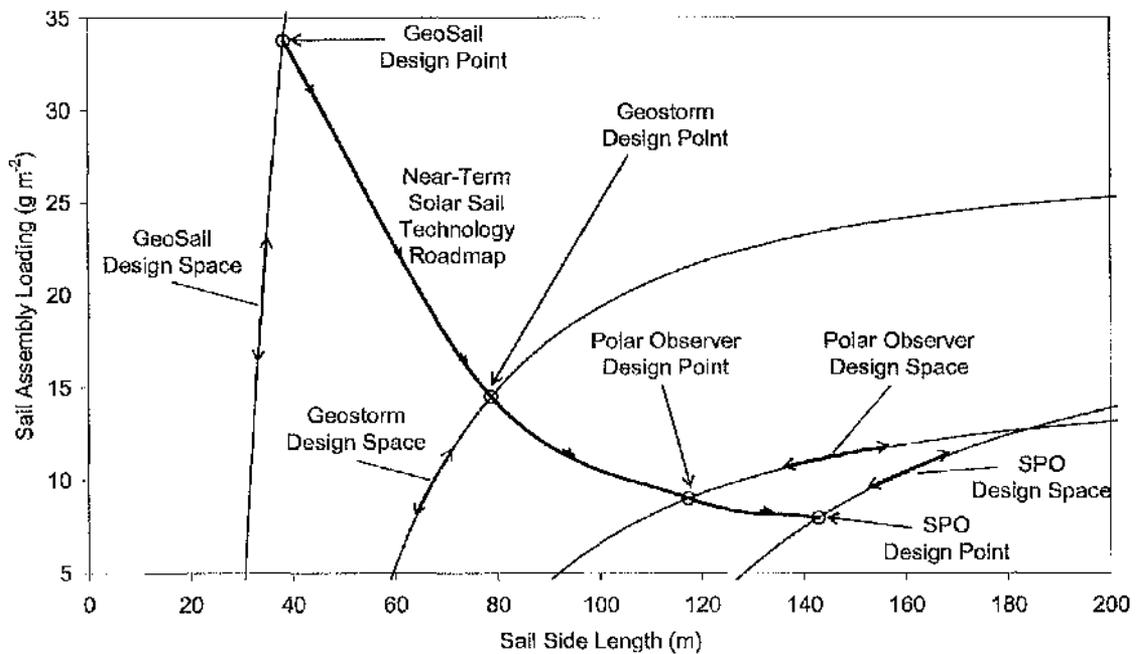
Figure 19 Sail design space for near-term roadmap leading into the mid-term missions.[82]

engine plume with the local environment. Following the TRS requirements for the IHP mission analysis, it has been demonstrated that solar sailing is the optimal propulsion system.[110] The TRS states that the spacecraft should reach 200 AU in 25 years. Furthermore, the sail should be jettisoned at 5 AU to eliminate any potential interference caused by the solar sail on the local space environment (recall Section 1.4.3) and the spacecraft should be delivered to the nose of the heliosphere at latitude 7.5 deg and longitude 254.5 deg at 200 AU from the Sun, in the ecliptic coordinate frame. As briefly mentioned in Section 1.3 the high-velocity solar system escape is attained through a very close solar pass often called a Solar Photonic Assist (SPA), this enables the solar sail to utilise the $1/r^2$ variation of solar irradiance to gain energy from the Sun during the close solar pass.[35, 69, 126, 127] The IHP mission and the design of suitable trajectories will be discussed in Section 6 of this dissertation.

Other mid to far term solar sail missions of interest include missions to Mercury,[35, 128 - 135] outer planet fast fly-by missions[1, 35, 136, 137] and Non-Keplerian orbit applications.[1, 138 - 144] Mercury is an attractive environment for solar sailing due to the abundance of solar energy so close to the Sun, however the thermal environment is challenging. Mercury applications of solar sailing will be discussed in Sections 4, 5 and 6 of this dissertation. Outer planet fly-by missions using solar sailing are attractive, as the fly-by can be performed very rapidly, however if the fly-by is attained quickly it is likely due to the spacecraft having a very high encounter velocity, which degrades the mission science returns. Such missions are not directly addressed within this dissertation. Non-Keplerian orbit applications are some of the most exotic and exiting proposed for solar sailing due to the requirement for continuous thrust. However, it is also this requirement which may pose the primary problem for such missions due to the potentially limited sail pointing stability,[92, 145, 146] coupled with the potential contamination of the local space environment by the sail as discussed in Section 1.4.3.[86, 87] Such issues may mean that solar sailing can only be used to provide a high-

energy orbit which can be maintained with zero thrust; allowing the sail to be jettisoned and the science phase of the mission to begin, for example SPO and IHP. Non-Keplerian orbits are not directly addressed within this dissertation.

Solar sail propulsion opens up a diverse range of new and exciting mission opportunities. The most attractive are the high energy missions which are truly enabled by solar sailing. While some are clearly high energy missions, such as the SPO mission[81, 109, 112, 113] and Sun-synchronous Mercury orbiter,[129 - 131] others do not so obviously belong to this class. For example, GeoSail requires only a moderate performance solar sail. However, the time integrated acceleration required to precess the orbit apse-line over an extended duration also places this mission within the high energy class. It should therefore be remembered that the use of a low or moderate performance solar sail does not constitute a low energy mission if the solar sail is used for an extended duration. A solar sail development route is shown in Figure 20 (from Reference 147) where one sees the required sail area plotted against sail assembly loading for the range of solar sail missions discussed above. A generalised trend can be seen in Figure 20, where the near-term sail missions appear at the top-left of the plot, such as GeoSail and Geostorm, with more advanced missions appearing at the bottom right, such as the IHP mission. Missions are denoted as near, mid and far-term missions. Near-term missions are defined as having an assembly loading of greater than 10 gm^{-2} and sail area of < 15,000 m^2 as sail assembly loading increases towards > 30 gm^{-2}. Mid-term missions are defined as sail assembly loadings of > 5 gm^{-2} and sail areas of < 40,000 m^2 at 12 gm^{-2}, while far-term missions have a lower assembly loading bound of 0.5 gm^{-2}, rising to 1 gm^{-2} as sail area approaches 140,000 m^2. All points under the far-term curve, such as the sail required for an Oort Cloud Fly-Through Mission,[148] are defined as Beyond Far-Term. Figure 20 considers only the sail size and mass requirements without considering mission complexity, such as high slew rates or thermal loads. It is therefore required that one takes these factors into account. Table 2
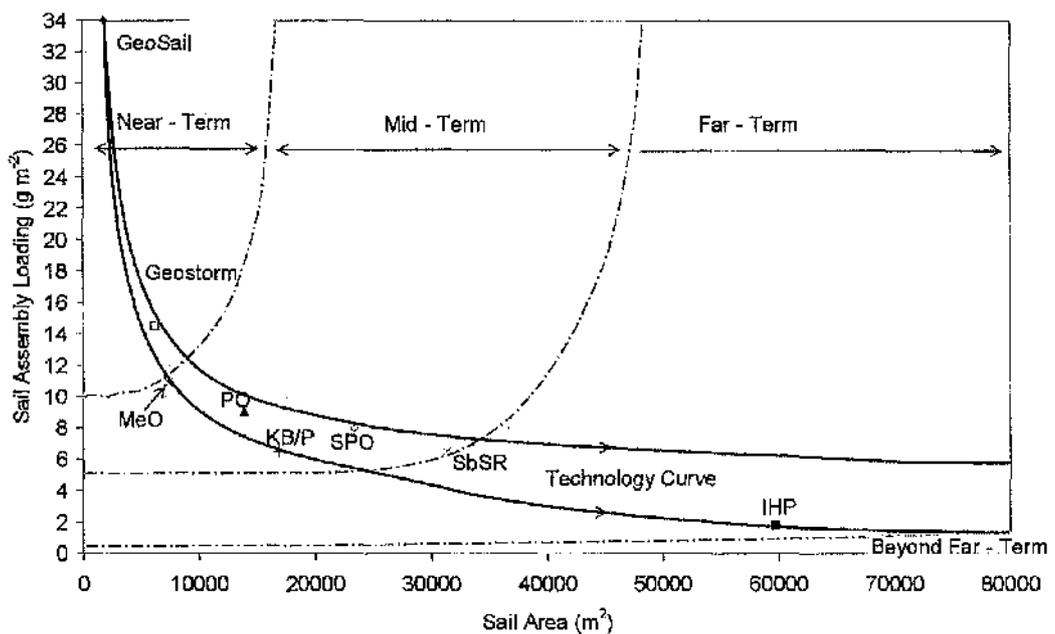
Figure 20 Solar sail technology development route.[147]

Mission	a_{s_0}		σ_s		A	
GeoSail	0.10	mm s^{-2}	35	g m^{-2}	1850	m^2
Geostorm	0.31	mm s^{-2}	15	g m^{-2}	6200	m^2
Polar Observer (PO)	0.55	mm s^{-2}	9	g m^{-2}	14000	m^2
Solar Polar Observer (SPO)	0.42	mm s^{-2}	8	g m^{-2}	23400	m^2
Kuiper Belt / Pluto Mission (KB/P)	0.50	mm s^{-2}	6	g m^{-2}	16900	m^2
Mercury Orbiter (McO)	0.25	mm s^{-2}	10	g m^{-2}	7500	m^2
High-Energy Small Body Sample Return (SbSR)	0.50	mm s^{-2}	6.5	g m^{-2}	31000	m^2
Interstellar Heliopause Probe (IHP)	1.50	mm s^{-2}	< 2	g m^{-2}	60000	m^2

Table 2 Solar sail mission applications and summaries; sorted in near to far term order with nearest first.[147]

summaries potential solar sail mission applications which are significantly enhanced or enabled by sail propulsion and sorts them in near to far term order, with nearest first. Note that the missions are not simply sorted in order of increasing sail performance or size. The GeoSail, Solar Polar Orbiter and Interstellar Heliopause Probe missions are denoted key missions in the near, mid and far-term respectively. While it would be possible to go directly from one key mission to the next intermediate missions offer risk reduction assistance, for example the Kuiper Belt / Pluto Mission could be an Interstellar Heliopause Probe precursor mission.[147]

1.6   Work Objectives and the Context of this Dissertation

This dissertation will address the following questions and issues:

1. Perform an analytical investigation of planetary escape using solar sail propulsion such as to explain previously noted but unexplained anomalies; for example, the variation in Earth escape time as a function of the Earth's position about the Sun.

2. Investigate, for the first time, the variation in solar sail escape time from Mercury as a function of Mercury's position about the Sun.

3. Generate an analytical solar sail trajectory design method which allows the generation of near-optimal realistic planetary escape trajectories for the first time (trajectories which do not, for example, have negative altitude phases).

4. Perform a thorough investigation of solar sail Earth escape trajectories which do not pass through the Earth's shadow cone, while using the methods developed for point 3.

5. Develop an analytical solar sail trajectory design method which can rapidly produce complex planet-centred orbit transfers and station-keeping algorithms for the first time.

6. Develop an analytical heliocentric solar sail trajectory design method which can rapidly produce near-optimal solutions (trajectories within 1 – 3 % of optimal).

One of the most time consuming phases of high-level mission analysis (where technology drivers, mission time-scale, cost, et cetera, are the issue rather than mission specifics) is the trajectory generation and optimisation phase. To obtain true-optimal trajectories, which fully match the two-point boundary value problem for transfer and rendezvous, one must use numerical methods and optimal control theory. Optimal trajectory generation is a complex field and many schemes exist, however these are typically characterised as being computationally intensive and requiring a good degree of engineering judgement. One such method is calculus of variations, which forms the basis of the NASA-JPL VARITOP low-thrust trajectory optimisation tool, where the control Hamiltonian is maximised while also satisfying the transversality condition.[149] 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.[150, 151] 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 inexperienced user and the problem is often highly sensitive to them. For this reason, the calculus of variations based method is classed as an indirect method.

Alternatively, gradient-based, non-linear programming methods can be used to solve the constrained parameter optimisation, two-point boundary value problem by iteratively selecting a discretised thrust cone and clock angle control history that satisfies the boundary conditions and orbit constraints, while minimising the transfer time. Such methods are termed direct methods. The sail control 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. A significant number of different direct optimisation methods exist. Multiple-shooting methods propagate adjacent trajectory segments backwards and forwards in time through an iterative process, attempting to match each of the segment boundary states. Another direct parameter optimisation method uses non-linear programming algorithms such as sequential quadratic programming (SQP) to optimise the parameters. Direct parameter optimisation methods are reasonably robust but suffer due to the deterministic, gradient-based, local-search methods employed. All local search methods can converge to a local 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 to that optimal. If the initial guess is poor the solution will likely converge to a local rather than global optimal. Such scenarios can be difficult to identify, leading to locally optimal solutions wrongly being presented as globally or near-globally optimal. Alternatively, global search methods negate the requirement for an initial guess of any kind and can in principle converge on the optimal solution; removing the requirement for an experienced user. Most global methods employ stochastic processes which use analogies from the biological and physical world, such as genetic algorithms or an evolutionary neurocontroller.[137, 152 - 155] However, the computational cost of global methods can be high and hence prohibitive for assessing potential mission scenarios very rapidly. It is thus clear that conventional optimisation methods are powerful optimisation techniques, however as the number of orbit revolutions is increased such methods become increasingly time consuming. As such, for low-thrust planet-centred trajectories a different approach is required. Similarly, to enable the very rapid generation of

near-optimal heliocentric solar sail trajectories a new approach is required. Within this dissertation the use of locally optimal control laws (sometimes called "*heuristic control laws*") will be considered.

1.7   Publication List

As support of the application for the degree of Doctor of Philosophy this section of the dissertation lists journal and conference papers published by the candidate. These papers relate to the main subject of the thesis but do not, in the majority, replicate work which is presented within the dissertation.

1.7.1   Journal Papers

The following journal papers support the application for the degree of Doctor of Philosophy:

- Macdonald, M., Hughes, G., McInnes, C. R., Lyngvi, A., Falkner, P., Atzei, A., "The Solar Polar Orbiter: A Technology Reference Study", Journal Spacecraft & Rockets, In Press.

- Hughes G., Macdonald M., McInnes C. R., Atzei, A., Falkner, P., "Sample Return from Mercury and other Terrestrial Planets Using Solar Sail Propulsion", Journal of Spacecraft and Rockets, In Press.

- Macdonald M., McInnes C. R., Dachwald, B., "Heliocentric Solar Sail Orbit Transfers with Locally Optimal Control Laws", Journal of Spacecraft and Rockets, In Press.

- Dachwald, B., Mengali, G., Quarta, A.A., Macdonald, M., "Parametric Model and Optimal Control of Solar Sails with Optical Degradation", Journal of Spacecraft and Rockets, In Press.

- Macdonald M., McInnes C. R., "Analytical Control Laws for Planet-Centred Solar Sailing", Journal of Guidance, Control, and Dynamics, Vol. 28, No. 5, pp. 1038-1048, 2005.

- Macdonald M., McInnes C. R., "Realistic Earth Escape Strategies for Solar Sailing", Journal of Guidance, Control, and Dynamics, Vol. 28, No. 2, pp. 315-323, 2005.

- McInnes C. R., Hughes G., Macdonald M., "Low Cost Mercury Orbiter and Mercury Sample Return Missions Using Solar Sail Propulsion", The Aeronautical Journal, Paper No. 2790, pp. 469-478, August 2003.

- McInnes C. R., Hughes G., Macdonald M., "Payload Mass Fraction Optimisation for Solar Sail Cargo Missions", Journal of Spacecraft and Rockets, Vol. 39, No. 6, pp. 933-935, November 2002.

- 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, pp. 622-629, July-August 2001.

1.7.2 Supporting Conference Papers

The following conference papers support the application for the degree of Doctor of Philosophy:

- McInnes, C.R., Hughes, G., Macdonald, M., "Small Bodies Missions and Technologies (Part 2)", IAC-05-A3.5.B.05, Proceedings of 56th International Astronautical Congress, Fukuoka, Japan, October 2005.

- Dachwald, B., Seboldt, W., Macdonald, M., Mengali, G., Quarta, A.A., McInnes, C.R., Rios-Reyes, L., Scheeres, D.J., Wie, B., Görlich, M., Lura, F., Diedrich, B., Baturkin, V., Coverstone, V.L., Leipold, M., Garbe, G.P., "Potential Solar Sail Degradation Effects on Trajectory and Attitude Control", AIAA Paper 05-6172, Proceedings of AIAA/AAS Guidance, Navigation, and Control Conference, San Francisco, August 2005.

- Dachwald, B., Baturkin, V., Coverstone, V. L., Diedrich, B., Garbe, G. P., Görlich, M., Leipold, M., Lura, F., Macdonald, M., McInnes, C. R., Mengali, G., Quarta, A.A., Rios-Reyes, L., Scheeres, D.J., Seboldt, W., Wie, B., "Potential Effects of Optical Solar Sail Degradation on Interplanetary Trajectory Design", AIAA Paper 05-413, Proceedings of AAS Astrodynamics Specialist Conference, Lake Tahoe, California, August 2005.

- Macdonald M., McInnes C. R., "A Near-Term Roadmap for Solar Sailing", IAC-04-U.1.09, Proceedings of 55th International Astronautical Congress, Vancouver, Canada, October 2004.

- Hughes, G., Macdonald M., McInnes C. R., "Analysis of a Solar Sail Mercury Sample Return Mission", IAC-04-Q.2.B.08, Proceedings of 55th International Astronautical Congress, Vancouver, Canada, October 2004.

- Macdonald, M., Hughes, G., "Solar Sailing", Proceedings of Summer Workshop on Advanced Topics in Astrodynamics, Barcelona, July 2004.

- Macdonald M., McInnes C. R., Alexander D., Sandman A., "GeoSail: Exploring the Magnetosphere Using a Low-Cost Solar Sail", Proceedings of 5th IAA International Conference on Low-Cost Planetary Missions, ESA Special Publication SP-542, pp. 341-349, September 2003.

- Hughes G., Macdonald M., McInnes C. R., Atzei, A., "Terrestrial Planet Sample Return Using Solar Sail Propulsion", Proceedings of 5th IAA International Conference on Low-Cost Planetary Missions, ESA Special Publication SP-542, pp. 377-384, September 2003.

- Macdonald M., McInnes C. R., "Seasonal Efficiencies of Solar Sailing Planetary Orbit", IAC-02-S.6.01, Proceedings of 53rd International Astronautical Congress, Houston, October 2002.

- 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, Proceedings of 53rd International Astronautical Congress, Houston, October 2002.

- Macdonald M., McInnes C. R., "Solar Sail Capture Trajectories at Mercury", AIAA-2002-4990, Proceedings of AIAA/AAS Astrodynamics Specialist Conference, Monterey, August 2002.

- Alexander D., McInnes C. R., Angelopoulos V., Sandman A. W., Macdonald M., "GeoSail: A Novel Solar Sail Mission Concept for Geospace", Space Technology and Applications International Forum (STAIF 2002), Albuquerque, February 2002.

- 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.

- Macdonald M., McInnes, C. R. "GeoSail; An Enhanced Magnetosphere Mission, Using a Small Low Cost Solar Sail", IAF-00-W.1.06, 51st International Astronautical Congress, Rio de Janeiro, Brazil, 2-6 October 2000.

1.7.3   Other Papers Published by the Candidate, Non – Supporting

A list of other publications by the candidate is presented below. These papers are however not included as support for the application for the degree of Doctor of Philosophy.

- Macdonald M., McInnes C. R., "Spacecraft Planetary Capture Using Gravity Assist Manoeuvres", Journal of Guidance, Control, and Dynamics, Vol. 28, No. 2, pp. 365-369, 2005.

2   Locally Optimal Control Laws

The primary advantage of locally optimal control laws (sometimes called *"heuristic control laws"*) is the speed which they can be implemented in a trajectory calculation; giving results up to several orders of magnitude quicker than either direct or indirect methods. The primary disadvantage of locally optimal control laws is the non-optimal nature of the method and resulting solution. One method of *"heuristic"* trajectory generation is based on Lyapunov feedback control, where a suitable Lyapunov function must be defined by the mission designer.[156 - 158] Another method of locally optimal control involves the blending of locally optimal control laws where the trajectory is typically split into several phases, selected through engineering judgement. Blending control laws has previously been used for low-thrust orbit transfers where no constraint is placed upon the thrust vector orientation, such as orbit transfers by SEP.[159 - 162] The use of blended locally optimal control laws has also been established for solar sail trajectories by Macdonald and indeed is the subject of this dissertation.[163 - 167] Prior blending methods, that is to say those used for SEP transfers in Reference 159 and 160, have used optimisation techniques to set the weight function of each control law; giving the weightings as a function of time from start epoch. Thus, while the individual control laws are a function of only the orbit elements as will be discussed in Section 2.1, the final blended optimal force vector is a function of time due to the optimisation process. The approach adopted within this dissertation is that the weight functions should be independent of time; using the osculating orbit elements to set the weight functions of each control law prior to blending. Defining the weight functions as functions of only the orbital elements offers several potential benefits. As the sail control angles are a function of only the osculating elements the control system is able to adjust for small unforeseen orbit perturbations or perturbations which cannot currently be accurately modelled due to lack of real-world knowledge, such as sail wrinkles. Thus, the system should potentially be suitable as an on-board autonomous controller, significantly reducing the amount of data in the uplink telecom budget with the sail requiring only its current position rather than an entire new set of control angles. This concept was initially proposed by the candidate in Reference 163 and has since been widely used for solar sail trajectory generation by the candidate.[73, 74, 81, 109, 133 - 135, 163 - 167] Following this work, from 2001 to the present, Petropoulos in 2003 proposed a similar approach for SEP orbit transfers.[162] Petropoulos uses the control laws for SEP orbit transfers and blends them to generate relatively simple transfer trajectories.

The optimality of the blended system depends heavily on the weight functions applied in obtaining the blended locally-optimal thrust vector. The method used to generate the weights for planetary escape trajectories in Section 4 are relatively simple, as only two control laws are being blended. However, in Section 5 the transfer and station-keeping trajectories generated are much more complex and require the blending of more than two control laws. Thus, a new control method is developed in order to set the weighted importance of each of the orbit elements through consideration of multiple criteria. The algorithms developed in Section 5 allow complex planet-centred solar sail transfer trajectories to be generated for the first time. Within Section 6 the algorithms developed in Section 5 are evolved for use in heliocentric solar sail orbit transfers. It is found that the algorithms developed generate heliocentric trajectories which are very

close to the optimal trajectory. Discretisation of the locally optimal control law solution can thus be used to provide an initial guess in order to find the optimal and to help avoid derivation of sub-optimal solutions.

2.1   The Generation of Locally Optimal Control Laws

The rate of change of any orbit element can be calculated; hence a locally optimal control law can be generated for any orbital element. Such control laws maximise the instantaneous rate of change of the element and provide the required thrust orientation in analytical form. It is important to note that local optimality does not guarantee global optimality.

The variational equation of the element to be optimally altered is written in the form,

$$\frac{d\sigma}{dt} = \mathbf{f}.\boldsymbol{\lambda}_\sigma$$

Equation 61

where σ represents an arbitrary orbit element. The required relative perturbing force, $\mathbf{f}$, on each of the Radial, Transverse and Normal (*RTN*) axes to maximise the rate of change of σ is found as the orientation of $\boldsymbol{\lambda}_\sigma$. Maximising the thrust vector along $\boldsymbol{\lambda}_\sigma$ maximises the right-hand side of Equation 61 and thus the instantaneous rate of σ is maximised. In order to determine the sail control angles which maximise the sail thrust along $\boldsymbol{\lambda}_\sigma$ it is required to define $\boldsymbol{\lambda}_\sigma$ in the same reference frame as the sail control angles. Recall from Figure 12 that the sail control angles are defined within the *RTN* reference frame for a heliocentric solar sail trajectory, thus no conversion of $\boldsymbol{\lambda}_\sigma$ is required. However, a planet-centred solar sail trajectory does not define the sail control angles within the *RTN* reference frame, thus $\boldsymbol{\lambda}_\sigma$ must be transposed into the Sun-sail line reference frame defined within Figure 12. This transformation is performed using standard transformation matrix, and will be discussed in Section 3.4.1.[168, 169] With conversion of $\boldsymbol{\lambda}_\sigma$ into the Sun-sail line coordinate system, where required, the pitch angle of the ideal force vector is defined as the angle between the Sun-sail line and the ideal force vector, that is to say,

$$\widetilde{\alpha} = \arccos\left(\lambda_{\hat{x}}\right)$$

Equation 62

where $\hat{\boldsymbol{\lambda}}_\sigma = \left[\lambda_x/|\boldsymbol{\lambda}_\sigma| \quad \lambda_y/|\boldsymbol{\lambda}_\sigma| \quad \lambda_z/|\boldsymbol{\lambda}_\sigma|\right] = \left[\lambda_{\hat{x}} \quad \lambda_{\hat{y}} \quad \lambda_{\hat{z}}\right]$. Note, the derivation of locally optimal control laws within this dissertation implicitly assume an ideal sail force model. With the derivation of the pitch angle of the ideal force vector a standard optimisation derivative is used to find the sail orientation which will maximise the sail thrust vector along the ideal force vector, Equation 63.[170] The locally optimal sail pitch angle is thus found directly from the Equation 63 and $\widetilde{\alpha}$ as

$$\tan \alpha = \frac{-3\cos\widetilde{\alpha} + \sqrt{9\cos^2\widetilde{\alpha} + 8\sin^2\widetilde{\alpha}}}{4\sin\widetilde{\alpha}}$$

Equation 63

The locally optimal sail clock angle is found directly from the ideal force vector using Equation 64 and does

not require optimisation since the sail acceleration magnitude does not depend on the sail clock angle.

$$\delta = \arccos\left(\frac{\lambda_{\hat{z}}}{\sqrt{\lambda_{\hat{y}}^2 + \lambda_{\hat{z}}^2}}\right)$$

Equation 64

If a negative rate of change of σ is desired a negative of the ideal force vector should be used, thus inverting the ideal force vector orientation prior to application of Equation 61 – Equation 64. The sail thrust vector then induces a negative rate of change of element σ.

2.2   Orbit Elements and Gauss' Form of the Variational Equations

The classical orbit elements are illustrated in Figure 21, where the illustrated central body is the Earth. The definition however is independent of the central body, for example, the axis system simply changes to a Sun-centred inertial reference frame for heliocentric orbits. A further definition of the classical and other orbit elements can be found in Table 3.

Lagrange's variational equations are derived for the special case in which the disturbing acceleration is represented as the gradient of the disturbing function; an unnecessary constraint for the purposes of this dissertation. The variational equations can instead be derived appropriate to the various choices of component resolutions of the disturbing acceleration vector; as attributed to Gauss. The derivation of the variational equations in the Gaussian form can be found in References 168 and 169. The variational equations of the five classical orbit elements in the Gauss' form, and following Equation 61, are given in Equation 65 – Equation 69 as

$$\frac{da}{dt} = \frac{2a^2}{\sqrt{\mu p}}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} e\sin\nu \\ (1 + e\cos\nu) \\ 0 \end{bmatrix}$$

Equation 65
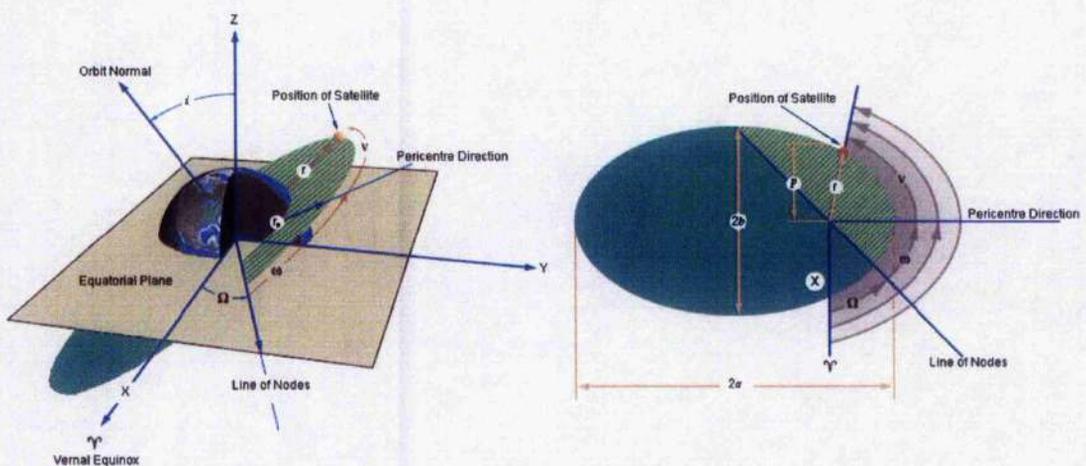
Figure 21 Orbit elements. Image adapted from a NASA original.

Parameter	Symbol	Definition
Semi-major axis	a	Half the major axis of an orbit's ellipse.
Semi-minor axis	b	Half the minor axis of an orbit's ellipse.
Eccentric Anomaly	E	$= \cos^{-1}\left[\frac{1}{e}\left(1 - \frac{r}{a}\right)\right]$
Eccentricity	e	$= \sqrt{1 - b^2/a^2}$
Inclination	i	Angle between the orbital plane and a reference plane, typically the Equatorial plane for Earth -- centred or the Ecliptic plane for heliocentric.
Eccentric longitude	K	$= \varpi + E$
True Longitude	L	$= \Omega + \omega + \nu = \varpi + \nu$, a broken angle, measured in the reference plane from the zero point to the ascending node and then around the orbit to the satellite.
Mean longitude	l	$= \varpi + M$, a broken angle, measured in the reference plane from the zero point to the ascending node and then around the orbit.
Mean Motion	n	$= \sqrt{\mu/a^3}$, the mean motion, or mean angular velocity
Semi-latus rectum	p	$= a\left(1 - e^2\right)$, half a chord through the focus and parallel to the conic section directrix.
Orbit radius	r	Distance from the coordinate system origin, typically coincident with the centre of the central body, to the satellite.
True Anomaly	ν	$= \cos^{-1}\left[\frac{1}{e}\left(\frac{p}{r} - 1\right)\right]$, angle from pericentre to the satellite, measured within the orbit plane.
Argument of Pericentre	ω	Angle from the ascending node to the satellite when at pericentre, measured within the orbit plane.
Longitude of Ascending Node	Ω	Angle between line of nodes and the zero point of longitude in the reference plane.
Longitude of pericentre	ϖ	$= \Omega + \omega$, a broken angle, measured in the reference plane from the zero point to the ascending node and then around the orbit to pericentre.

Table 3 Definition of orbit elements.

$$\frac{de}{dt} = \sqrt{\frac{p}{\mu}}[R \quad T \quad N]\begin{bmatrix} \sin v \\ \cos v + \cos E \\ 0 \end{bmatrix}$$

Equation 66

$$\frac{di}{dt} = \frac{r}{\sqrt{\mu p}}[R \quad T \quad N]\begin{bmatrix} 0 \\ 0 \\ \cos(v + \omega) \end{bmatrix}$$

Equation 67

$$\frac{d\Omega}{dt} = \frac{r}{\sqrt{\mu p}}[R \quad T \quad N]\begin{bmatrix} 0 \\ 0 \\ \dfrac{\sin(v + \omega)}{\sin i} \end{bmatrix}$$

Equation 68

$$\frac{d\omega}{dt} = \frac{1}{\sqrt{\mu}}[R \quad T \quad N]\begin{bmatrix} -\dfrac{\sqrt{p}}{e}\cos v \\ \left(1 + \dfrac{r}{p}\right)\dfrac{\sqrt{p}}{e}\sin v \\ -\dfrac{r}{\sqrt{p}}\cot i \sin(v + \omega) \end{bmatrix}$$

Equation 69

Equation 65 – Equation 69 can be used to propagate a trajectory with the inclusion of a sixth position fixing element. The sixth element could be the true anomaly, eccentric anomaly, mean anomaly or the true longitude. However, while a variational equation can be defined for these elements and hence a locally optimal control law derived, such position fixing locally optimal control laws are meaningless. Thus, rendezvous orbits cannot be analytically generated using locally optimal control laws; however they can be generated by simple trial and error as performed in Reference 160. Equation 65 – Equation 69 will now be used to generate control laws for the five classic orbit elements, plus control laws for the radius of pericentre and apocentre. The variational equations for radius of pericentre and apocentre are given in Equation 70 and Equation 71.

$$\frac{dr_p}{dt} = \frac{da}{dt}(1 - e) - a\frac{de}{dt}$$

$$= a\sqrt{\frac{p}{\mu}}[R \quad T \quad N]\begin{bmatrix} \dfrac{2ae(1 - e)\sin v}{p} - \sin v \\ \dfrac{2a(1 - e)(1 + e\cos v)}{p} - (\cos v + \cos E) \\ 0 \end{bmatrix}$$

Equation 70

$$\frac{dr_a}{dt} = \frac{da}{dt}(1+e) + a\frac{de}{dt}$$

$$= a\sqrt{\frac{p}{\mu}}[R \quad T \quad N]\begin{bmatrix} \dfrac{2ae(1+e)\sin v}{p} + \sin v \\[2mm] \dfrac{2a(1+e)(1+e\cos v)}{p} + (\cos v + \cos E) \\[2mm] 0 \end{bmatrix}$$ Equation 71

Note from Section 3 that the equations of motion are propagated in modified equinoctial elements; the control laws will thus be defined in classic orbit elements and modified equinoctial elements.

2.3   Locally Optimal Semi-Major Axis Control Law

The semi-major axis control law is also known as the energy gain control law as it provides for a locally optimal variation in the orbit energy. This control law is the most widely used locally optimal control law; it is often used to generate crude estimates of low-thrust planetary escape trajectories, as will be discussed in Section 4. Recall from Equation 61 that one must identify the vector λ_a from Equation 65, as seen in Equation 72. Note, λ_a is defined in classical elements and then converted to modified equinoctial elements, as these are the equations of motion selected in Section 3. The subsequent control laws are also defined in classical elements and then converted to modified equinoctial elements. The modified equinoctial elements will be discussed and derived in Section 3.2.

$$\lambda_a = \begin{bmatrix} e\sin v \\ (1+e\cos v) \\ 0 \end{bmatrix} = \begin{bmatrix} f\sin L - g\cos L \\ 1 + (f\cos L + g\sin L) \\ 0 \end{bmatrix}$$ Equation 72

Ensuring that the vector λ_a is defined within the Sun-sail line reference frame, and converting it to this frame if not, as will be discussed in Section 3.4.1, Equation 72 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64. Note that the heliocentric use of this control law, along with one for eccentricity, aphelion radius, inclination and ascending node are derived and illustrated in Reference 1. Figure 22 and Figure 23 show the use of the locally optimal semi-major axis control law in an Earth-centred orbit used to gain and to reduce orbit energy over a 3 day period starting approximately on the vernal equinox of the year 2000, to be exact Julian Day (JD) 2451624.5. The initial orbit is circular, with GEO radius and is placed, as close as possible, within the ecliptic plane. The sail characteristic acceleration is 1 mm s^{-2}; a value used within the remainder of this section of the dissertation. The trajectory model used is described in Section 3. No orbit perturbations are considered, other than sail thrust. The Sun is assumed to be a point source and the sail is assumed to be an ideal reflector, as discussed in Section 1.3.2. Furthermore, all periods of occulation are neglected. As defined in Section 3.5 this is denoted as Model 1. Note from Section 3 that the trajectory model defines the Earth's position as true-to-date, thus the eccentricity of the Earth's orbit is implicitly included within all trajectory calculations. It is thus critically important that the sail acceleration also be corrected for the true Sun – sail distance as failure to do so results in significant errors under certain

scenarios.[§§] Exclusion of all perturbations allows the behaviour of the orbit elements under the use of each control law to be clearly illustrated. Note from Figure 23 that the use of the locally optimal control law ensures that the rate of change, i.e. $^{da}/_{dt}$, is always in the desired direction. The results in Figure 23 correspond well with a simplified model used in Reference 1 (pp 156 – 159) to illustrate a locally optimal energy gain control law which is derived by maximising the sail thrust along the velocity vector. Recall, the semi-major axis control law is a locally optimal energy gain control law. This correlation of results gives an initial indication that the basic sail trajectory model is correct, however further model validation will be conducted in Section 3.

2.4   Locally Optimal Eccentricity Control Law

Identifying the vector λ_e from Equation 66 one obtains Equation 73. Ensuring that the vector λ_e is defined within the Sun-sail line reference frame and converting it to this frame if not, as will be discussed in Section 3.4.1, Equation 73 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64. Figure 24 and Figure 25 illustrate the locally optimal eccentricity control law from the same initial orbit as in Section 2.3 for the eccentricity increase case. The eccentricity decrease case maintains the same initial conditions except that the eccentricity is increased to 0.1. Note that the eccentricity increase case has an eccentricity of 0.1295 after 3 days, yet the reduction case is unable to reverse these gains in the same period. This apparent anomaly is due to the orbit orientation about the planet, with respect to the Sun, requiring the sail thrust to be directed further from the Sun-line for the reduction case than the increase case as illustrated by the sail pitch angles in Figure 24. This inconsistency is unique to the planet-centred solar sail application of locally optimal control laws and does not occur in heliocentric trajectories. Furthermore, the apparent anomaly occurs in the radius of pericentre and apocentre control laws as well as the argument of pericentre control law.

$$\lambda_e = \begin{bmatrix} \sin \nu \\ \cos \nu + \cos E \\ 0 \end{bmatrix} = \begin{bmatrix} \dfrac{(f \sin L - g \cos L)}{\sqrt{f^2 + g^2}} \\ \dfrac{(f \cos L + g \sin L)(1 + r/p)}{\sqrt{f^2 + g^2}} + \dfrac{r\sqrt{f^2 + g^2}}{p} \\ 0 \end{bmatrix} \qquad \text{Equation 73}$$

[§§] This error mode was identified during the GeoSail mission analysis when determining the effect of various perturbations on the ability of the sail to track the Sun – Earth line.
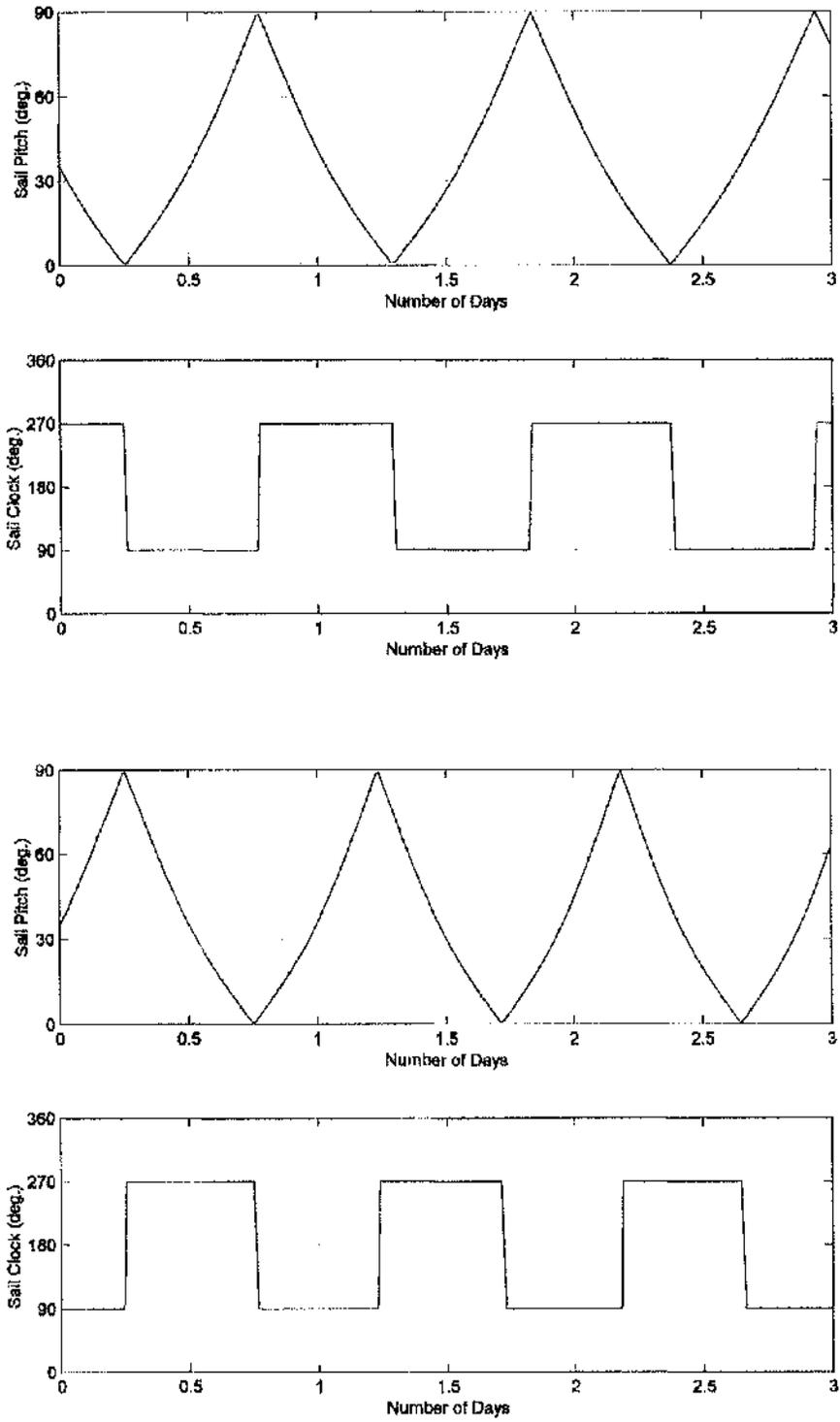
Figure 22 Locally optimal variation of semi-major axis control angles; gain (top) and decrease (bottom).
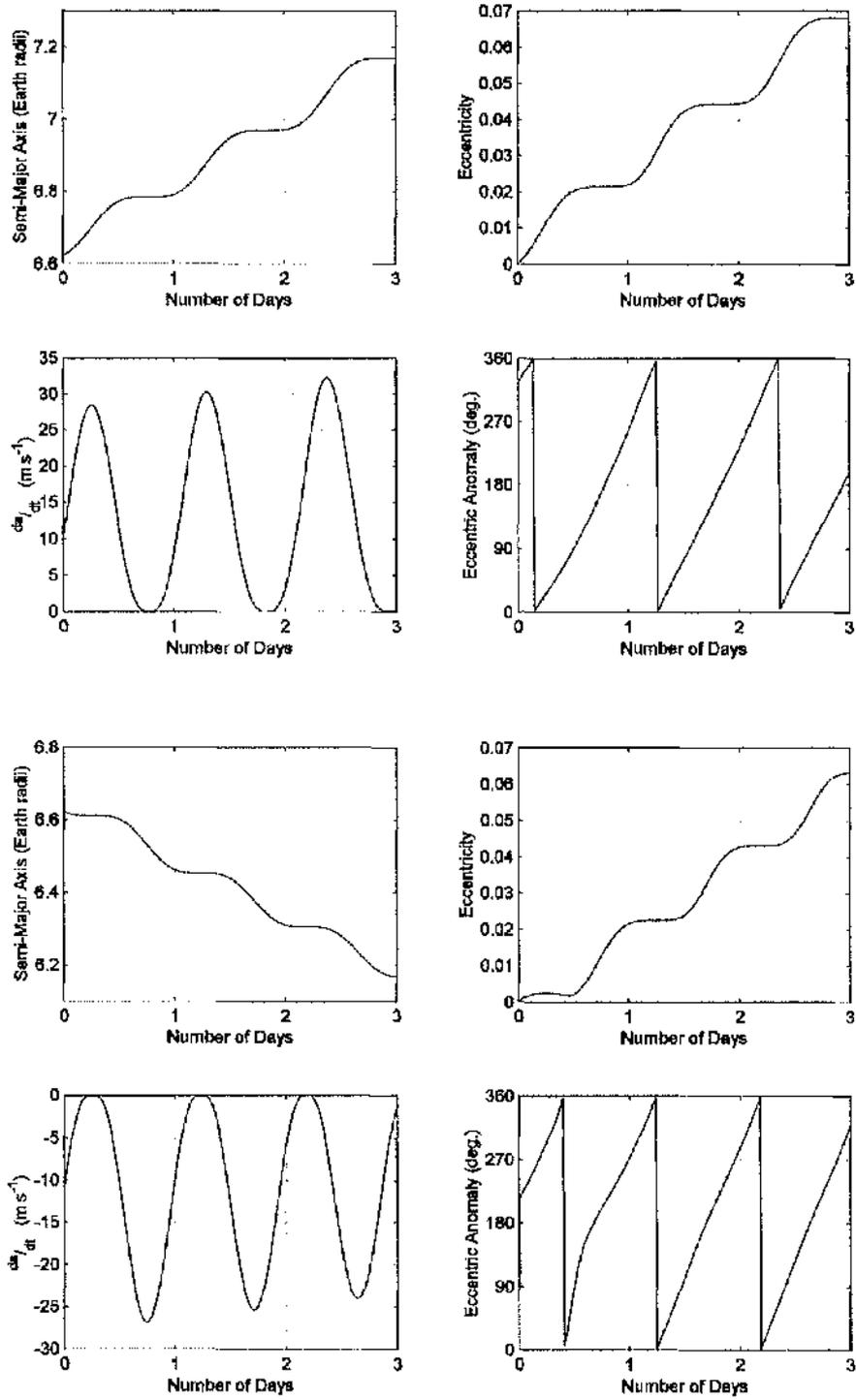
Figure 23 Orbit elements resulting from control angles in Figure 22; gain (top) and decrease (bottom).
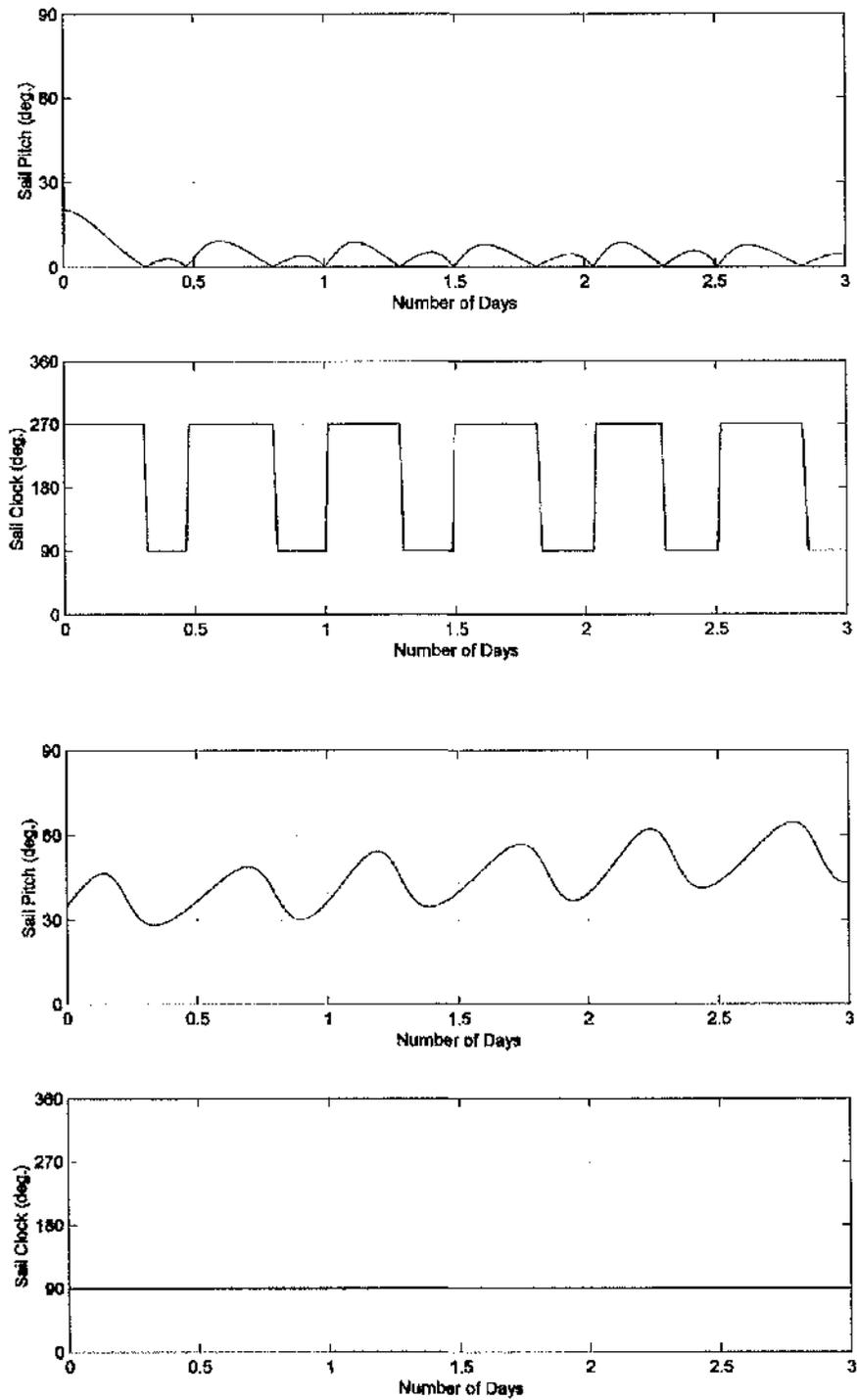
Figure 24 Locally optimal variation of eccentricity control angles; gain (top) and decrease (bottom).
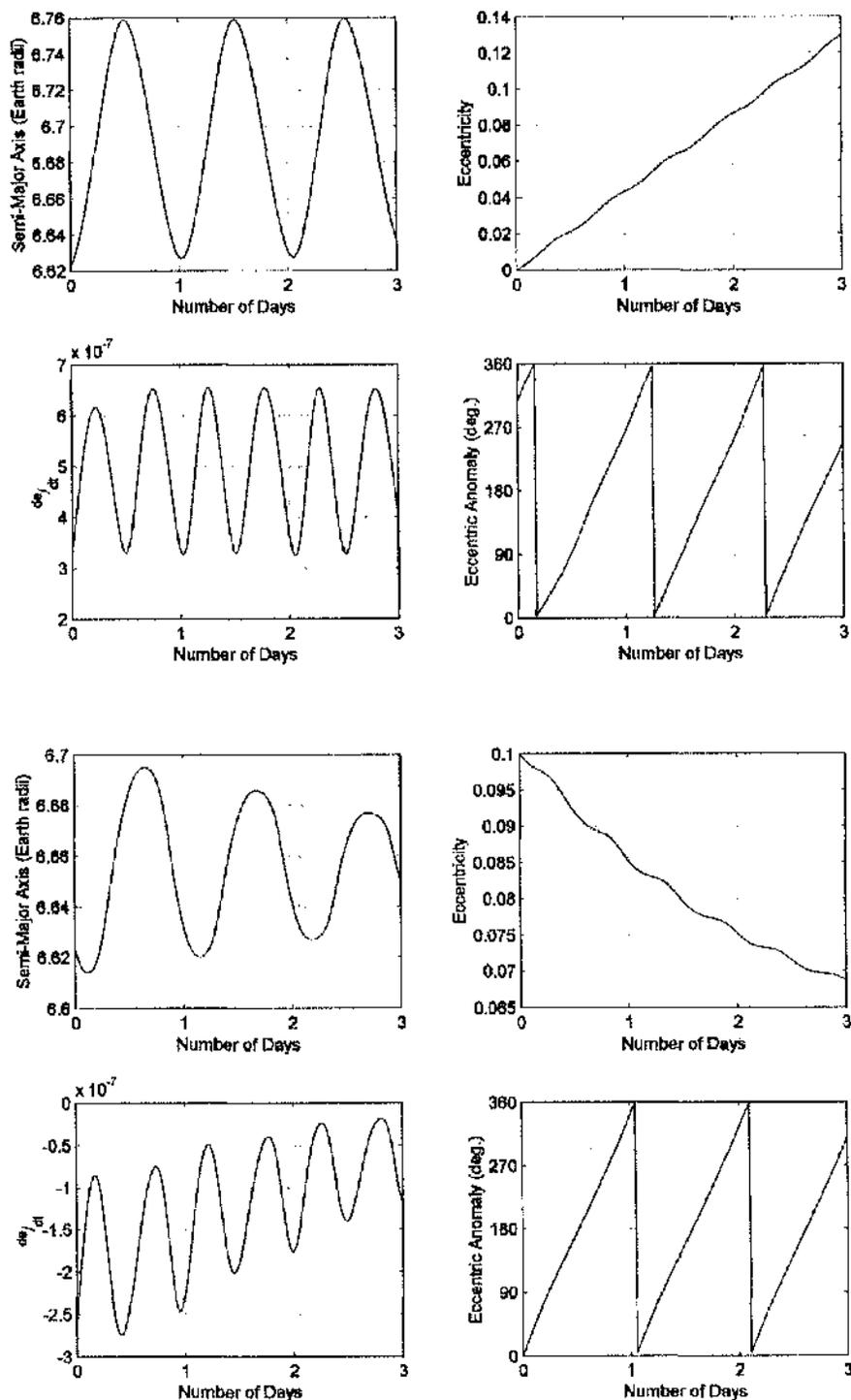
Figure 25 Orbit elements resulting from control angles in Figure 24; gain (top) and decrease (bottom).

2.5   Locally Optimal Inclination Control Law

Contrasting the semi-major axis and eccentricity control laws, the rate of change of inclination depends on only the one perturbing force as shown in Equation 67; the out-of-plane perturbation. Thus, the method used for locally optimal variation is somewhat different. From Equation 67 one sees that a switching relationship is required in order to maintain the desired sense of rate of change, positive or negative. For example, if $\cos(v + \omega)$ is negative one requires a negative out-of-plane sail force, hence generating a positive rate of change. One can thus identify the vector λ_i as,

$$\lambda_i = \begin{bmatrix} 0 \\ 0 \\ \text{sgn}[\cos(v + \omega)] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \text{sgn}\left(\dfrac{h \cos L + k \sin L}{\tau} \right) \end{bmatrix} \qquad \text{Equation 74}$$

Ensuring that λ_i is defined within the Sun-sail line reference frame, and converting it to this frame if not as will be discussed in Section 3.4.1, Equation 74 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64.

Figure 26 and Figure 27 illustrate the locally optimal inclination control law from the same initial orbit as in Section 2.3. Note that the sail pitch angle is fixed, while the sail clock angle displays a square wave with a phase difference of π between the inclination increase and decrease control laws. Note from Figure 27 that the use of this simple locally optimal control law ensures that the rate of change, i.e. $^{di}/_{dt}$, is always in the desired direction

2.6   Locally Optimal Longitude of Ascending Node Control Law

Similar to the locally optimal inclination control law in Section 2.5, it is recalled from Equation 68 that the longitude of ascending node is dependent only on the out-of-plane perturbation. Identifying the switching function one finds λ_Ω as,

$$\lambda_\Omega = \begin{bmatrix} 0 \\ 0 \\ \text{sgn}\left[\dfrac{\sin(v + \omega)}{\sin i} \right] \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \text{sgn}\left(\dfrac{h \sin L - k \cos L}{\tau} \right) \end{bmatrix} \qquad \text{Equation 75}$$

Note $\sin i \geq 0$, as $0 \leq i \leq {}^7/_2$, thus one can neglect this parameter when converting into modified equinoctial elements. Ensuring that λ_Ω is defined within the Sun-sail line reference frame, and converting it to this frame if not as will be discussed in Section 3.4.1, Equation 75 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64.

Figure 28 and Figure 29 illustrate the locally optimal ascending node control law from the same initial orbit as in Section 2.3. Note that the sail pitch angle is once again fixed, with the sail clock angle displaying a square wave with a phase difference of π between the increase and decrease control laws. Furthermore, a phase difference of $^{\pi}/_{2}$ is seen from the inclination control law discussed in Section 2.5. Note from Figure 29 that the use of this simple locally optimal control law ensures that the rate of change, i.e. $^{d\Omega}/_{dt}$, is always in the desired direction.

2.7   Locally Optimal Argument of Pericentre Control Law

It is noted from Equation 65 – Equation 71 that unlike most other orbit parameters the variation of the argument of pericentre depends on the perturbing acceleration along all three *RTN* axis. Despite this the locally optimal control law is derived in a similar manner to other orbit elements. Identifying the vector λ_ω from Equation 69 one gains Equation 76. Ensuring that the vector λ_ω is defined within the Sun-sail line reference frame, and converting it to this frame if not as will be discussed in Section 3.4.1, Equation 76 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64. Figure 30 and Figure 31 illustrate the locally optimal argument of pericentre control law from the same initial orbit as in Section 2.3 but with an initial eccentricity of 0.5 for both the increase and decrease scenarios. A starting epoch of the year 2000 summer solstice is used for the decrease case, to be precise a Julian date of 2451716.5. Note from Figure 30 that the control angles are now of a much more complex form than the previous control laws, however one also notes from Figure 31 that the use of this simple locally optimal control law ensures that the rate of change, i.e. $^{d\omega}/_{dt}$, is always in the desired direction

$$
\lambda_\omega = \begin{bmatrix} -\dfrac{\sqrt{p}}{e}\cos v \\[2ex] \left(1+\dfrac{r}{p}\right)\dfrac{\sqrt{p}}{e}\sin v \\[2ex] -\dfrac{r}{\sqrt{p}}\cot i\sin(v+\omega) \end{bmatrix}
$$

$$
= \begin{bmatrix} -\dfrac{f\cos L+g\sin L}{f^2+g^2} \\[2ex] \left(1+\dfrac{r}{p}\right)\dfrac{f\sin L-g\cos L}{f^2+g^2} \\[2ex] -\dfrac{r}{p}\left[\dfrac{\cot\left(2\arctan\sqrt{h^2+k^2}\right)\left(h\sin L-k\cos L\right)}{\tau\sqrt{f^2+g^2}}\right] \end{bmatrix}
$$

Equation 76
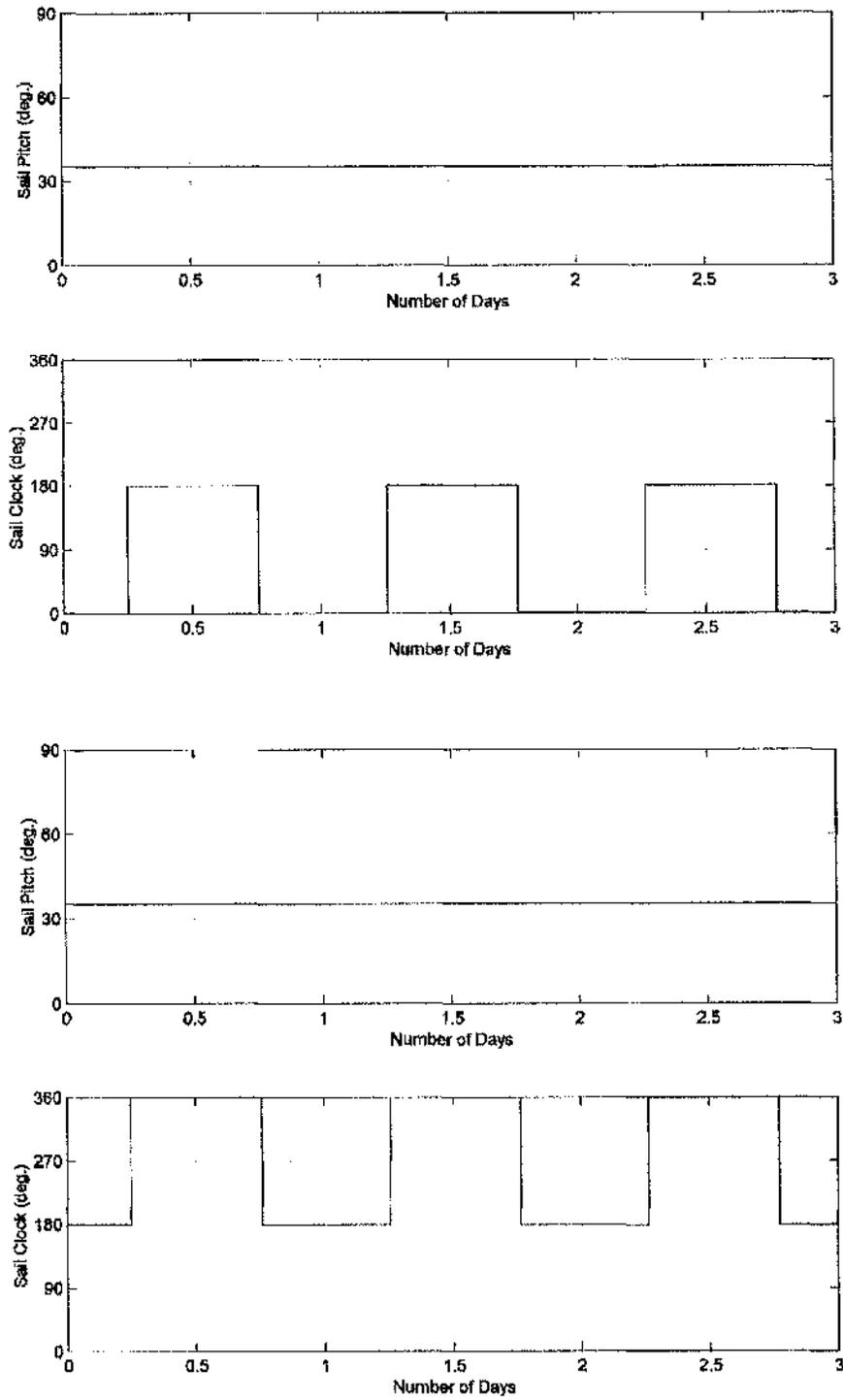
Figure 26 Locally optimal variation of inclination control angles; gain (top) and decrease (bottom).
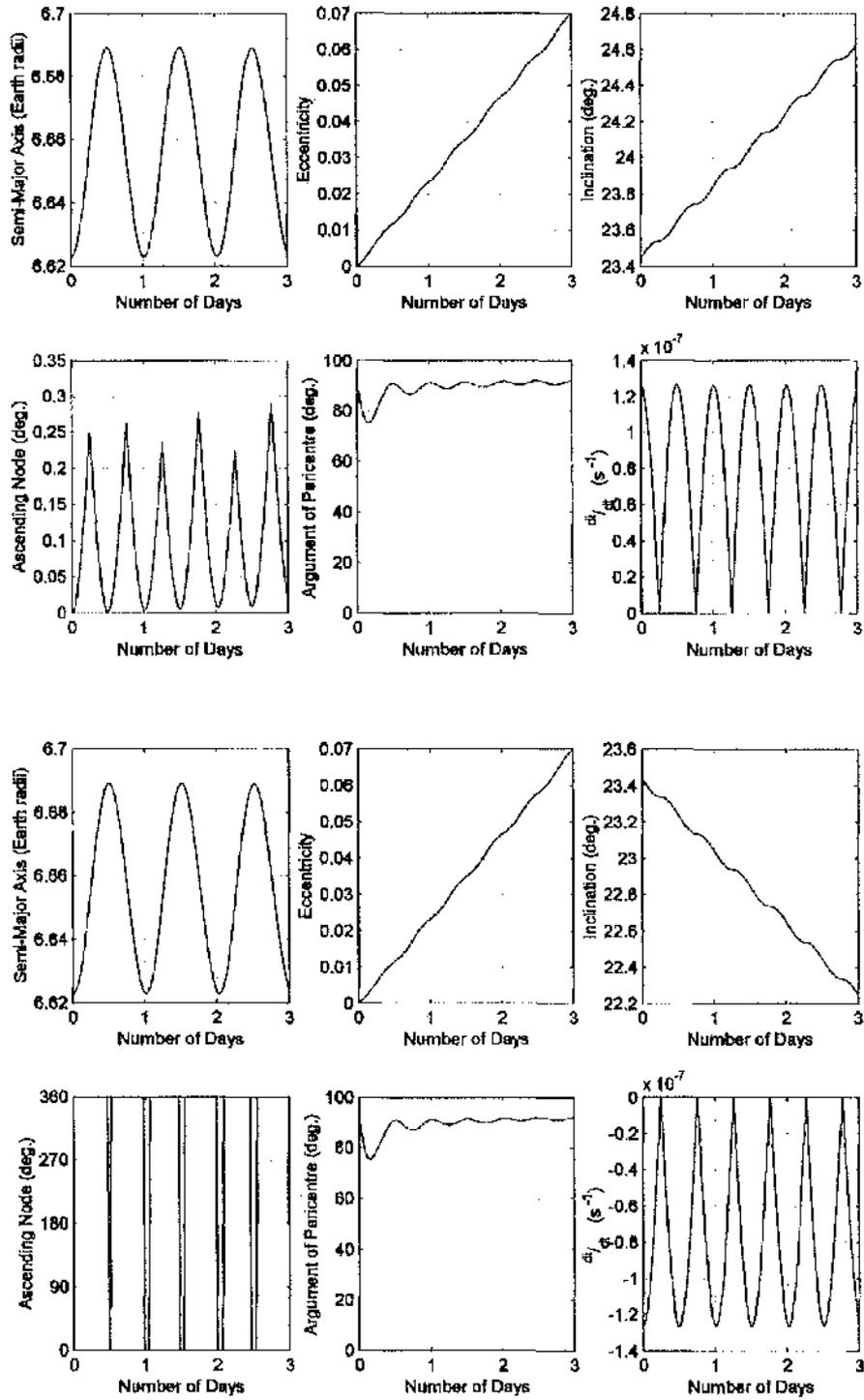
Figure 27 Orbit elements resulting from control angles in Figure 26; gain (top) and decrease (bottom).
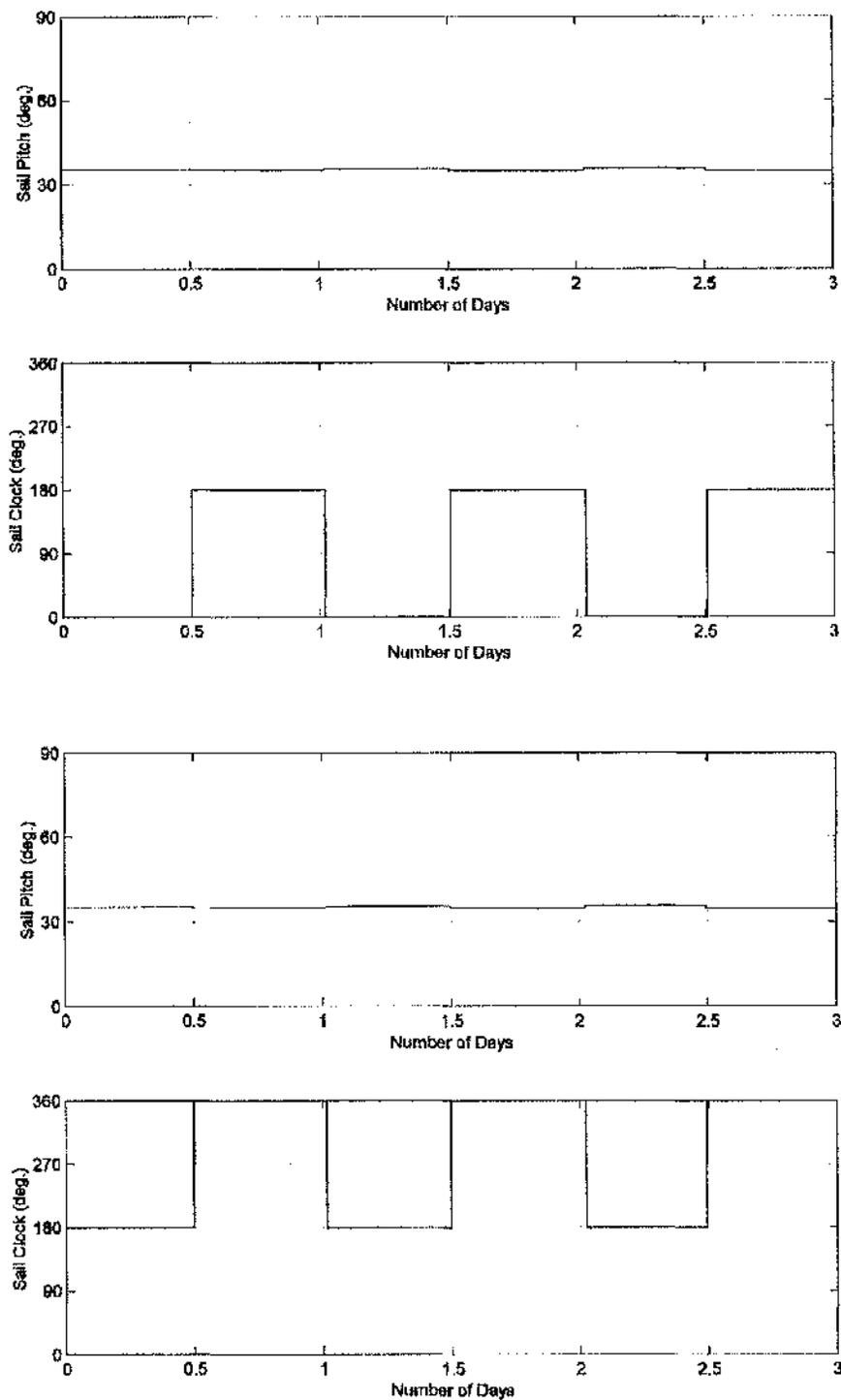
Figure 28 Locally optimal variation of ascending node control angles; gain (top) and decrease (bottom).
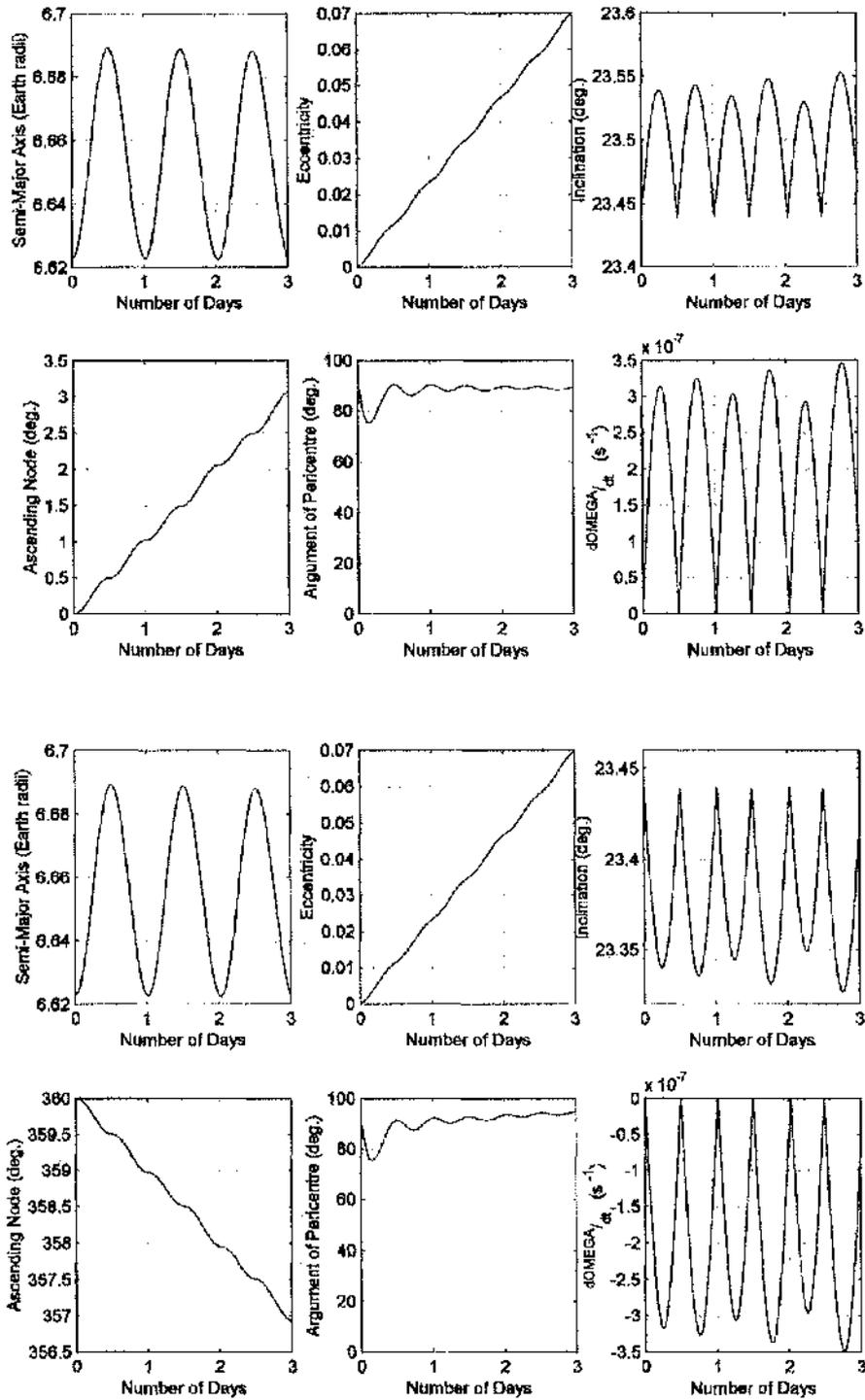
Figure 29 Orbit elements resulting from control angles in Figure 28; gain (top) and decrease (bottom).
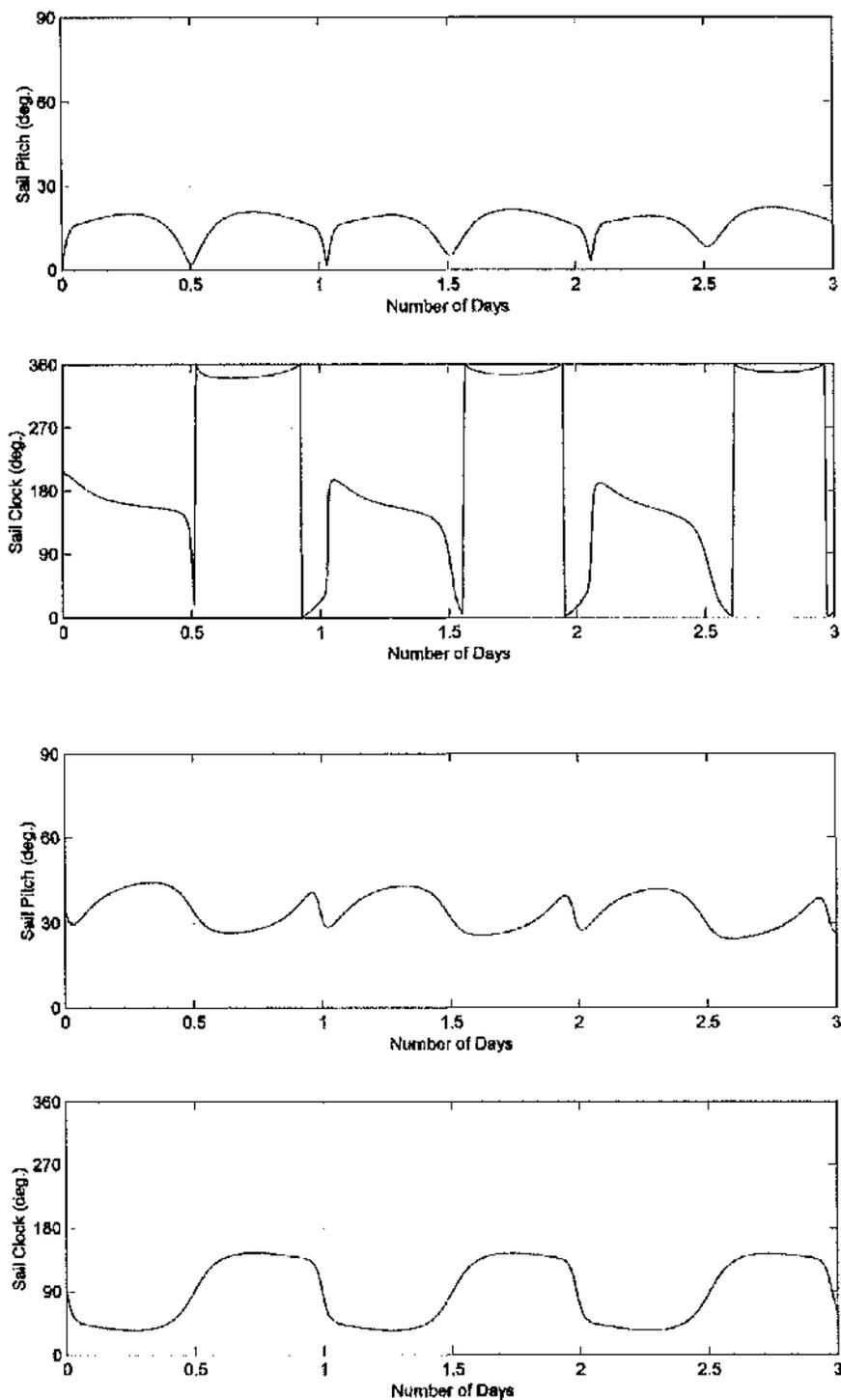
Figure 30 Locally optimal variation of argument of pericentre control angles; gain (top) and decrease (bottom).
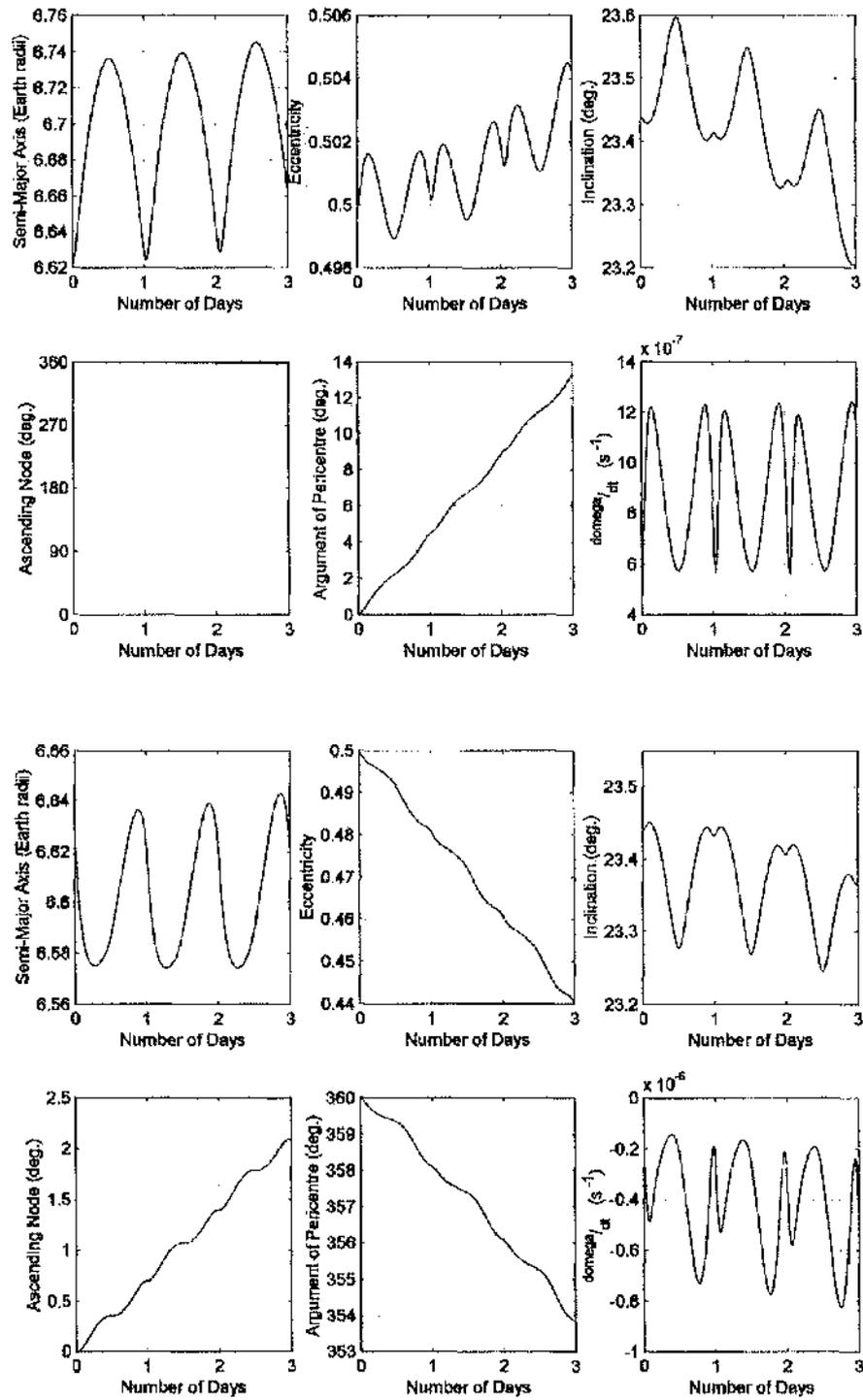
Figure 31 Orbit elements resulting from control angles in Figure 30; gain (top) and decrease (bottom).

2.8   Locally Optimal Radius of Pericentre Control Law

Identifying the vector λ_{r_p} from Equation 70 one gains Equation 77. Ensuring that the vector λ_{r_p} is defined within the Sun-sail line reference frame, and converting it to this frame if not as will be discussed in Section 3.4.1, Equation 77 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64. Figure 32 and Figure 33 illustrate the locally optimal radius of pericentre control law from the same initial orbit and conditions as the increase scenario in Section 2.7.

$$\lambda_{r_p} = \begin{bmatrix} \dfrac{2ae(1-e)\sin v}{p} - \sin v \\[2mm] \dfrac{2a(1-e)(1+e\cos v)}{p} - (\cos v + \cos E) \\[2mm] 0 \end{bmatrix}$$

$$= \begin{bmatrix} (f\sin L - g\cos L)\left[\dfrac{2\left(1-\sqrt{f^2+g^2}\right)}{1-f^2-g^2} - \dfrac{1}{\sqrt{f^2+g^2}}\right] \\[4mm] \dfrac{2\left(1-\sqrt{f^2+g^2}\right)(1+f\cos L + g\sin L)}{1-f^2-g^2} - \left(\dfrac{f\cos L + g\sin L}{\sqrt{f^2+g^2}} + \cos E\right) \\[4mm] 0 \end{bmatrix} \qquad \text{Equation 77}$$

2.9   Locally Optimal Radius of Apocentre Control Law

Identifying the vector λ_{r_a} from Equation 71 one gains Equation 78. Ensuring that the vector λ_{r_a} is defined within the Sun-sail line reference frame, and converting it to this frame if not as will be discussed in Section 3.4.1, Equation 78 allows Equation 62 to be used which allows the locally optimal sail pitch angle to be found using Equation 63. The locally optimal sail clock angle is found using Equation 64. Figure 34 and Figure 35 illustrate the locally optimal radius of apocentre control law from the same initial orbit as in Section 2.8.
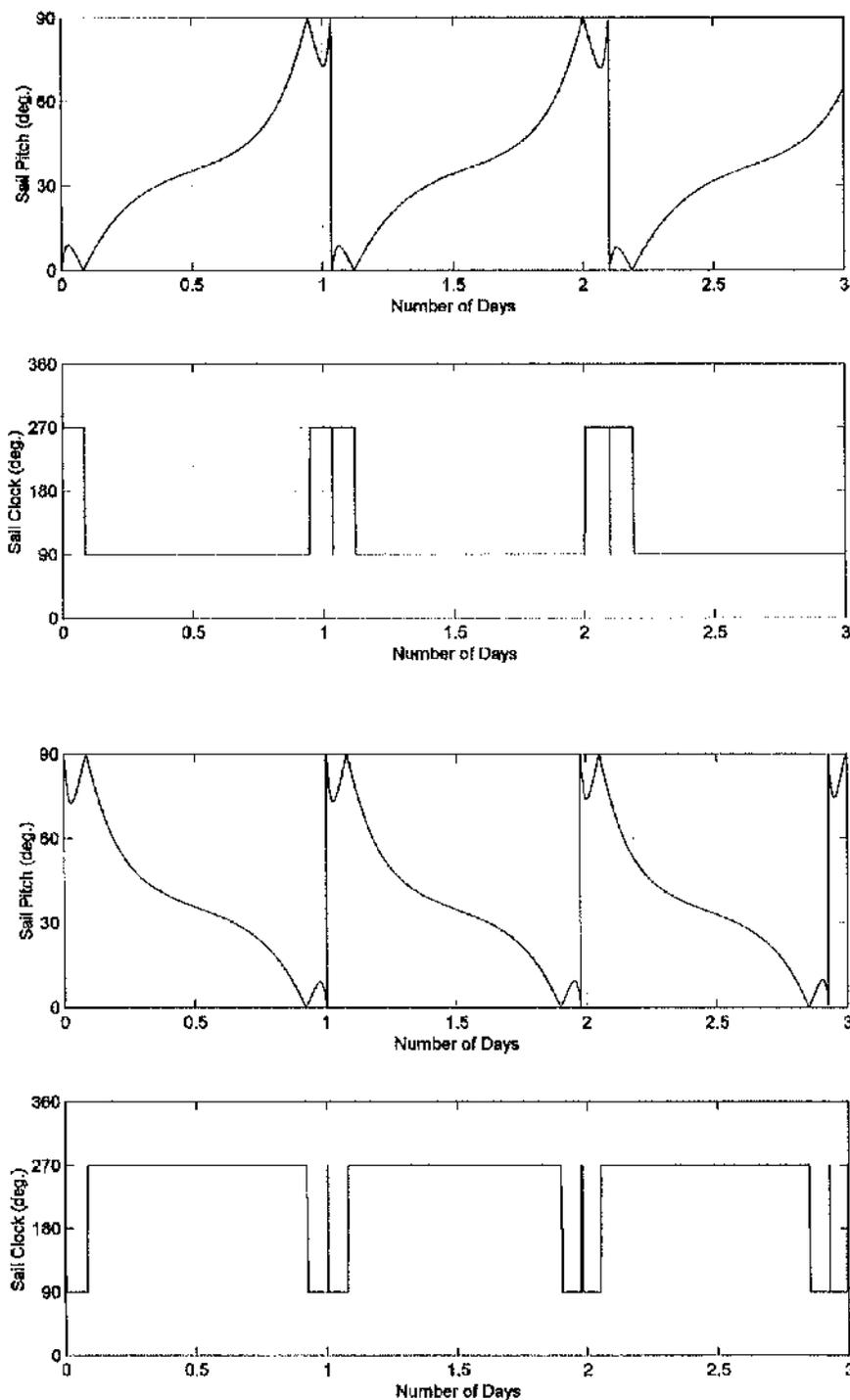
Figure 32 Locally optimal variation of radius of pericentre control angles; gain (top) and decrease (bottom).
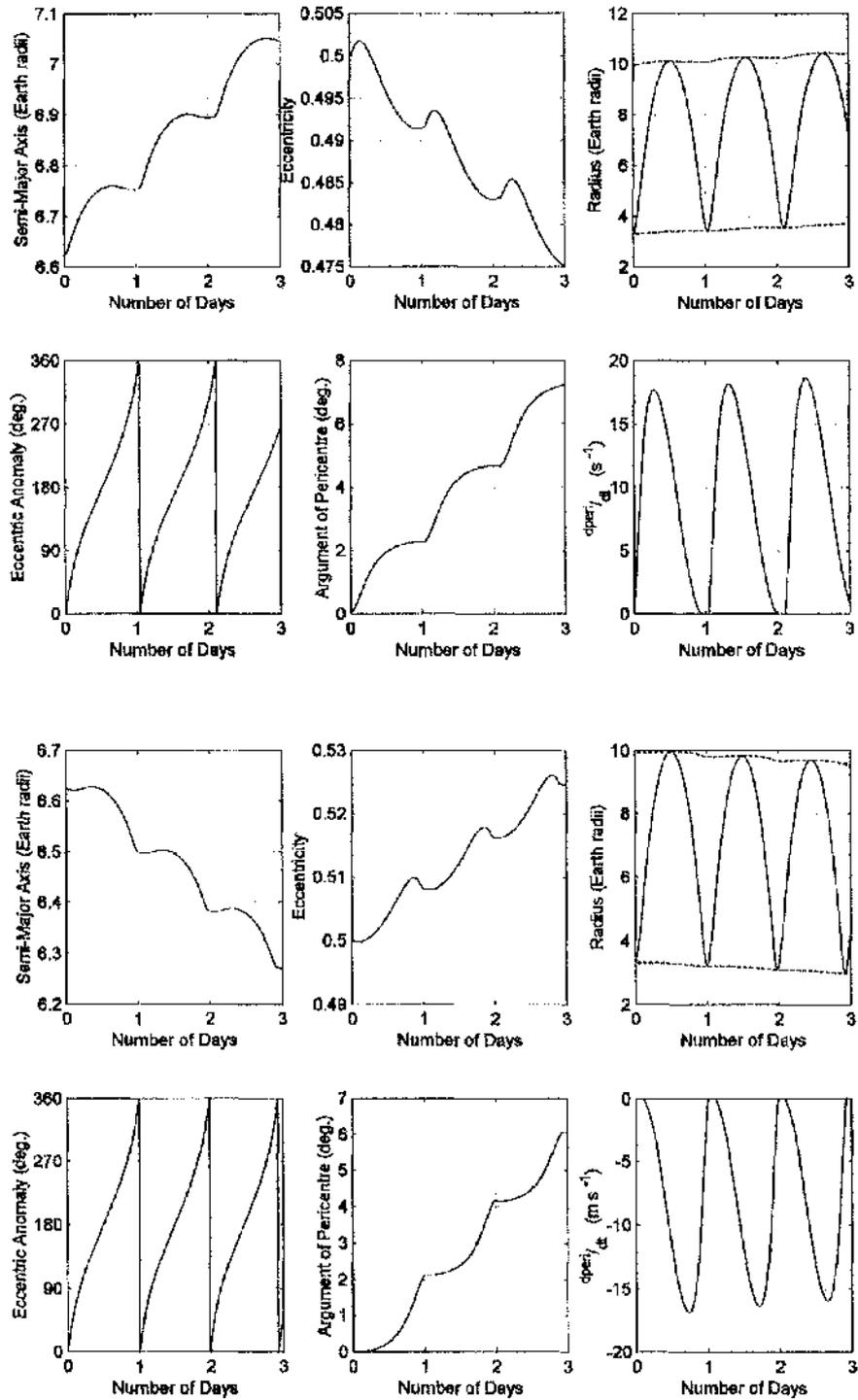
Figure 33 Orbit elements resulting from control angles in Figure 32; gain (top) and decrease (bottom).
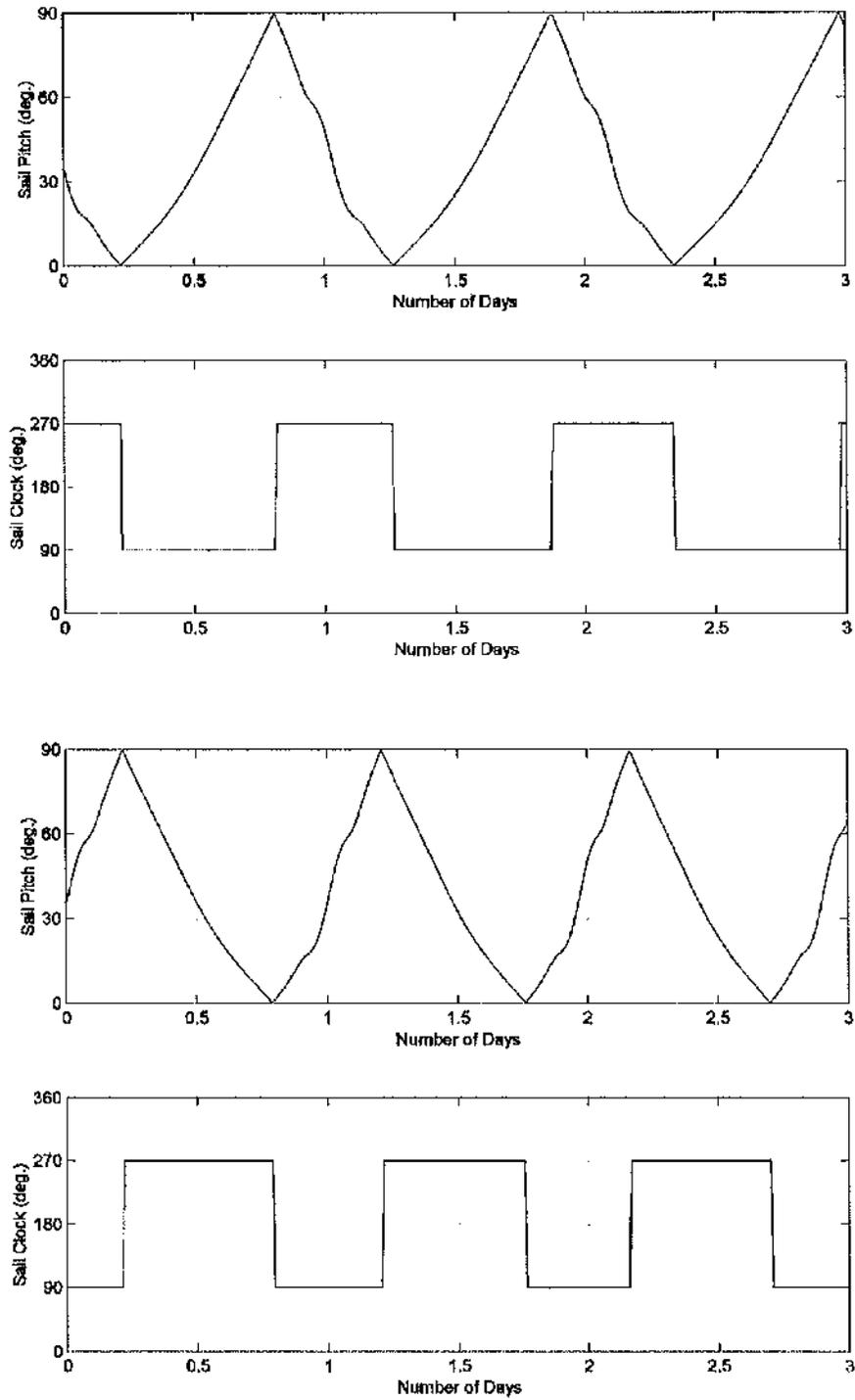
Figure 34 Locally optimal variation of radius of apocentre control angles; gain (top) and decrease (bottom).
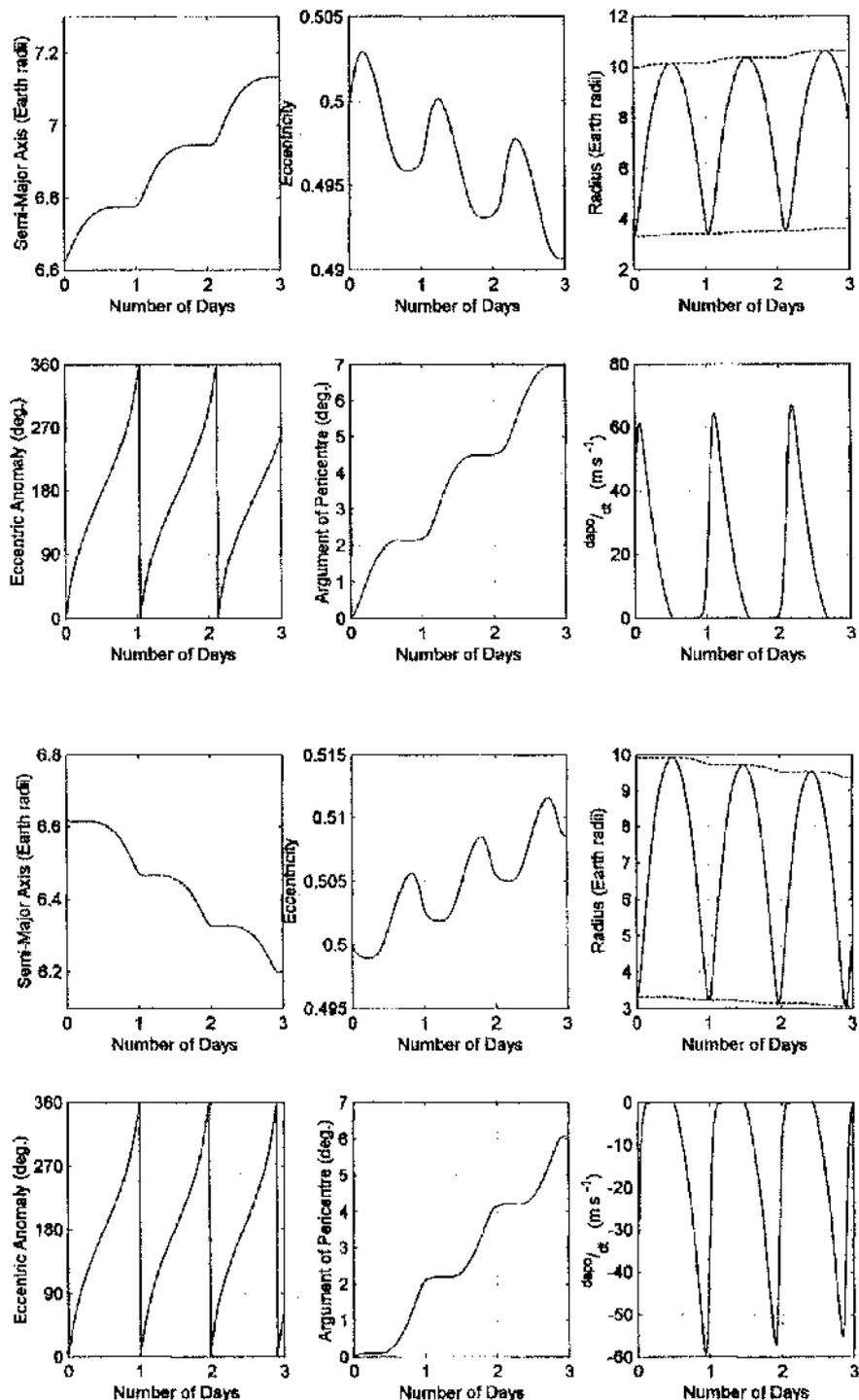
Figure 35 Orbit elements resulting from control angles in Figure 34; gain (top) and decrease (bottom).

$$
\lambda_{r_a} = \begin{bmatrix} \dfrac{2ae(1+e)\sin\nu}{p} + \sin\nu \\[2mm] \dfrac{2a(1+e)(1+e\cos\nu)}{P} + (\cos\nu + \cos E) \\[2mm] 0 \end{bmatrix}
$$

$$
= \begin{bmatrix} (f\sin L - g\cos L)\left[\dfrac{2\left(1+\sqrt{f^2+g^2}\right)}{1-f^2-g^2} + \dfrac{1}{\sqrt{f^2+g^2}} \right] \\[4mm] \dfrac{2\left(1+\sqrt{f^2+g^2}\right)(1+f\cos L + g\sin L)}{1-f^2-g^2} + \left(\dfrac{f\cos L + g\sin L}{\sqrt{f^2+g^2}} + \cos E \right) \\[4mm] 0 \end{bmatrix}
$$

Equation 78

2.10   Blending Control Laws

The blending of control laws is accomplished by initially calculating the vector λ_σ, in the Sun-sail line reference frame for each control law to be blended; obtaining a separate unit vector for each control law. The blended vector is thus computed as,

$$
\lambda_b = \frac{\displaystyle\sum_\sigma W_\sigma \hat{\lambda}_\sigma}{\displaystyle\sum_\sigma W_\sigma}
$$

Equation 79

where σ once again represents each orbit element. From $\hat{\lambda}_b$, rather than $\hat{\lambda}_\sigma$, one can directly define $\tilde{\alpha}$ using Equation 62 which allows the locally optimal blended sail pitch angle to be found using Equation 63. The locally optimal blended sail clock angle is found using Equation 64. The optimality of the blended system depends heavily on the weight functions applied in gaining the blended locally optimal thrust vector, that is to say λ_b. As already stated the blending procedure outlined for solar sail applications uses the orbital elements to define the weight of each control law, rather than defining each weight as a function of time from the start epoch as traditionally performed.[163 - 167] The methods of weight function definition developed within this dissertation are discussed in Sections 4 − 6.

3 Solar Sail Trajectory Model

In order to investigate the use of locally optimal control laws a special perturbations type problem was formulated. Special perturbations allows the position and velocity conditions of a body at a given epoch to be propagated over a small time interval, accounting for all the forces on the body during this interval, using the equations of motion. This calculation can be performed by one of a variety of methods, enabling the new positions and velocities at the end of the time interval to be found. A second computation using the new positions and velocities enables the process to be carried forward through another time interval. Each computation is called a step and in theory the numerical integration can be continued as long as desired. In reality rounding errors are introduced and the accuracy of any calculation decreases with every step. A potential, partial, solution to this error is to work with more significant figures than required such that the final rounding error does not influence the calculation when rounded to the required number of significant figures. Additionally, the error can be further alleviated by the use of as large a time step as possible during each calculation step, thus minimising the number of occasions on which the solution is rounded. Both of these alleviating methods are taken in all calculations in this dissertation to minimise the error of presented solutions and the error of all results presented is verified negligible with respect to the given number of significant figures. It can be shown that the probable error of a double integral is $0.1124n^{3/2}$, where n is the number of integration steps.[171] That is to say, after numerically integrating the second-order (x, y, z) equations of motion through 100 steps there is an even odds chance that the rounding error is smaller than 112.4 in units of the last decimal.[168] Reference 171 also shows that the mean error of the osculating elements of a body obtained by numerically integrating the Lagrange planetary equations, which are 1st order, will be proportional to $n^{1/2}$, except the mean orbital longitude (or whatever position fixing element is selected) where the mean error is again proportional to $n^{3/2}$ as this is a result of a double integral.

Perhaps the most straightforward method of determining the position and velocity of a body is to directly integrate the equations of motion in rectangular coordinates as first performed for a space body in 1908 by Cowell and Crommelin.[172, 173] The integration formulas used by Cowell and Crommelin were actually first given by Carl Friedrich Gauss. Cowell and Crommelin formulated their equations in rectangular coordinates and integrated them numerically by means of a multi-step algorithm. Since the publication of the paper by Cowell and Crommelin the use of the term Cowell's method has become ambiguous, within numerical analysis texts *"Cowell-type methods"* refer to multi-step algorithms similar to those used in the original paper.[168] However, in celestial mechanics the term *"Cowell's method"* refers to the formulation of the equations in a rectangular coordinate system and the subsequent integration using any technique whatsoever, for example by Runge-Kutta formulae.[168, 169, 174] Such a method is good for scenarios where the disturbing force or acceleration is of the same or higher order as that due to the central body, as the method does not distinguish between the two.[168] This however is also the primary disadvantage as a large number of significant figures have to be carried due to the large central force term, requiring many more time steps when the disturbing force or acceleration is small else a significant loss of accuracy occurs.[168] As a solar sail provides only a small perturbing acceleration the Cowell type model is not ideal for solar sail trajectory

propagation as the perturbing acceleration is often an order of magnitude (or more) lower than the central body acceleration. This is particularly true in the planetary escape and other complex multi-revolution trajectories which this dissertation focuses on. Note however that Cowell's method has been adopted recently for solar sail trajectory calculations, including Earth escape, however no justification was provided.[35]

If only the differential accelerations rather than the total acceleration are integrated considerable accuracy can be obtained with a larger time interval when the disturbing force or acceleration is small. This method is known as Encke's Method, however it was actually first proposed two years before Encke's work became known by Bond and Bond of Harvard University in 1849.[168, 169, 174] To a first approximation an orbit is a conic-section, this assumption is at the nucleus of Encke's method. Integrating the difference between the primary acceleration and the perturbing acceleration implies a reference orbit must be employed, along which the body would move in the absence of any perturbations. The integration gives the difference between the real coordinates and the conic-section coordinates. The conic-section orbit is an osculating orbit, thus at the epoch of osculation the differences vanish. As time from the initial epoch increases so the difference between the real coordinates and the conic-section coordinates increases, until it becomes necessary to derive a new osculating orbit. If a new osculating orbit is not derived the various accelerations will grow in magnitude and the process becomes cumbersome. The process of selecting a new conic orbit from which to calculate deviations is called *"rectification of the orbit"*.[168] Following rectification of the orbit the initial conditions for the deviation vector differential equation are again zero and the only non-zero acceleration is the disturbing acceleration. The error in determining the position and velocity of the osculating orbit is subject only to round-off errors and is independent of the integration technique used.[168] The accuracy of calculation of the deviation from the osculating orbit is limited by both round-off and truncation errors. The integrated quantities are small with respect to the osculating quantities and have little effect on the determination of the true orbit as before the errors become significant a new osculating orbit is selected through the process of rectification. The main advantage of Encke's method is the larger integration intervals which can be adopted compared to Cowell's method. However, the computational cost of a single Encke integration step is much greater than that of a Cowell step. The greater computational cost per step is typically more than compensated for by the larger step size. Encke's method has many applications, for example orbit determination of highly eccentric comets, such as the analysis performed by Encke on a comet later to be named after him. The method can also be used to analyse orbits in Earth – Moon space, where the Moon is taken as a perturbing body.

When propagating a near-Earth satellite it has been shown that the inclusion of the first-order effects of Earth oblateness in the reference orbit greatly improves the Encke method by increasing both the interval between rectifications of the reference orbit and the accuracy of the integration compared with the classic Encke method.[175] It has also been shown that the calculation time for the integration of the motion of four or more bodies can be reduced by an order of magnitude by comparison to the original Encke method if the reference orbit is taken to be a combination of several Keplerian orbits.[176] It is thus clear that the Encke method is optimised when the reference orbit is known and remains very close to the real evolving orbit for a significant period. It should also be noted that there is no necessity that the position and velocity in the reference orbit at any desired time be calculated from analytical expressions.[168] The Encke type model is

well suited to solar sail trajectory calculations and would form a valid base on which to build a solar sail trajectory model. However, the use of Variation of Parameters was selected for this dissertation as many of the traditional disadvantages of this system have been eliminated by either advances in theory or computational capabilities. Variation of Parameters also allows the analysis to be performed directly in terms of orbit elements and thus the use of locally optimal control laws is much simplified. Initially, variation of parameters may appear more problematical to implement than Encke's method, however it has some advantages when the perturbing acceleration is quite small. One of the primary differences is that the Encke reference orbit is constant until rectification occurs, however in variation of parameters the reference orbit is continuously changing and may thus be regarded as a form of Encke's method.[169]

3.1   Variation of Parameters

This method is also called the *"variation of orbital elements"* or the slightly paradoxical *"variation of constants"*, referring to the integration constants. In 1782 Lagrange completely developed for the first time the method of variation of parameters while studying the elliptical motion of comets. Lagrange's planetary equations (see Section 2.2) can be analytically integrated, as in the method of General Perturbations, or they can be integrated numerically step by step, with the new elements at the end of each step being used in the computation of the next step. Since Lagrange first introduced his planetary equations, where the rates of change of the osculating elements of a planet's orbit are given in terms of the elements of that planet and of the planets disturbing its heliocentric orbit, various attempts have been made to overcome some of the serious problems associated with the method. Some of the advantages of the variation of parameters method are that it is strictly a perturbation method and as such bypasses the central-body acceleration. For moderate perturbations the differentials of the elements are small and as such a larger step size can be used than in a rectangular coordinate method where the central-body acceleration must be calculated each step. Furthermore, the integration immediately exhibits the behaviour of the elements which is beneficial for the application of locally optimal control laws. Among the perceived disadvantages of the method is the more complicated nature of the right-hand side of the equations compared to those of the rectangular coordinates equations of motion, including the presence of sine and cosine terms. Additionally, the need to solve Kepler's equation, the break-down of the equations when orbit eccentricity is zero or one, or orbit inclination is zero, and the fact that the equations are usually given in elliptical elements and are thus inapplicable to parabolic, hyperbolic or rectilinear orbits are traditionally perceived disadvantages.[168] The disadvantages above regarding computational difficulties offset some of the benefits of a larger time-step than a Cowell type model, however such issues can be minimised with modern computing capabilities and prudent programming. As the orbit eccentricity drops towards zero the position of the apse becomes indeterminable, see Equation 69, similarly as the inclination drops to zero the ascending node becomes indeterminable, see Equation 68. The obvious solution is thus to define the orbit through a change in variables, for example noting symmetries one can apply standard transformations to make a change of variable from Keplerian to Delaunay variables.[177] Similarly, the solution of Kepler's equation can be avoided by changing the independent variable from time to a position fixing element, such as true or eccentric anomaly or the true longitude.[168]

3.2   Non-Singular Elements and the Equations of Motion

To derive variational equations which are non-singular one must find combinations of the classical elements which do not depend on either the line of nodes or the apsidal line. Adding the variational equations for Ω and ω (Equation 68 and Equation 69) one finds the resulting equation displays no singularity at zero inclination,

$$\frac{d\varpi}{dt} = \frac{1}{nabe}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} -p\cos\nu \\ (p+r)\sin\nu \\ er\sin(\omega+\nu)\tan\frac{i}{2} \end{bmatrix}$$ Equation 80

Noting,

$$\frac{dM}{dt} = n + \frac{1}{a^2 en}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} p\cos\nu - 2re \\ -(p+r)\sin\nu \\ 0 \end{bmatrix}$$ Equation 81

the variational equations for ϖ and M can thus be added to obtain an equation which also removes the singularity due to zero eccentricity,

$$\frac{dl}{dt} = n + \frac{1}{n}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} -\left(\frac{ep\cos\nu}{b(a+b)}+\frac{2r}{a^2}\right) \\ \left(\frac{e(p+r)\sin\nu}{b(a+b)}\right) \\ \left(\frac{r\sin(\omega+\nu)\tan\frac{i}{2}}{ab}\right) \end{bmatrix}$$ Equation 82

As Equation 82 is a function of the true anomaly, which is referenced to pericentre, further development is thus required. Kepler's equation can be written in the augmented form of Equation 83,

$$\begin{aligned} l &= \varpi + M = \varpi + E - e\sin E \\ &= (\varpi + E) + e\sin\varpi\cos(\varpi + E) - e\cos\varpi\sin(\varpi + E) \end{aligned}$$ Equation 83

Note that the orbit radius may be written as shown in Equation 84.

$$r = a(1 - e\sin\varpi\sin K - e\cos\varpi\cos K) = \frac{p}{1 + e\sin\varpi\sin L + e\cos\varpi\cos L}$$ Equation 84

From Kepler's equation, Equation 83, and the equation of an orbit, Equation 84, note that the eccentricity equivalent term and the longitude of pericentre equivalent term only appear in the combinations $e\sin\varpi$ and $e\cos\varpi$. These functions are thus selected to replace e and ϖ respectively. Following a similar process, and

writing the argument of latitude in terms of the true longitude one can select ($\tan{^1/_2} \sin \Omega$) and ($\tan{^1/_2} \cos \Omega$) to replace Ω and i. This element set is referred to as "*equinoctial elements*".[178] The equinoctial elements are non-singular except for rectilinear orbits and when $i = \pi$. This element set was in-fact first introduced by Lagrange in 1774 for his study of secular variations. Lagrange used i rather than $^1/_2$, however the inclusion of the half-angle simplifies the resulting Gaussian equations of motion and allows the use of Allan's expansion of the geopotential, if desired.[179]

3.2.1   Modified Equinoctial Elements

Employing a 'fast variable' (phase angle) as the sixth or position fixing element allows a regular perturbation technique to be used, with the fast variable as the independent variable. It thus becomes logical to modify the equinoctial elements by choosing true longitude in place of mean anomaly as the position fixing element. Furthermore, by replacing the semi-major axis with the semi-latus rectum one obtains a set of orbit elements which are non-singular for all orbits excluding $i = \pi$; however this singularity can be handled by an appropriate re-definition.[180] The "*modified equinoctial elements*" are thus defined in Equation 85 to Equation 90.[181]

$$p = a\left(1 - e^2\right)$$

Equation 85

$$f = e\cos(\omega + \Omega)$$

Equation 86

$$g = e\sin(\omega + \Omega)$$

Equation 87

$$h = \tan{^i/_2}\cos\Omega$$

Equation 88

$$k = \tan{^i/_2}\sin\Omega$$

Equation 89

$$L = \Omega + \omega + \nu$$

Equation 90

The auxiliary (positive) variables are defined in Equation 91 to Equation 95, noting that Equation 93 is simply the orbit radius.

$$s^2 = 1 + h^2 + k^2$$

Equation 91

$$w = 1 + f\cos L + g\sin L$$

Equation 92

$$r = \frac{p}{w}$$

Equation 93

$$\tau = \sqrt{h^2 + k^2} \qquad\qquad\qquad \text{Equation 94}$$

$$\alpha^2 = h^2 - k^2 \qquad\qquad\qquad \text{Equation 95}$$

The modified equinoctial elements equations of motion in the Gauss' form are found to reduce to

$$\frac{dp}{dt} = \frac{2p}{w}\sqrt{\frac{p}{\mu}}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \qquad\qquad \text{Equation 96}$$

$$\frac{df}{dt} = \sqrt{\frac{p}{\mu}}\frac{1}{w}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} w\sin L \\ (w+1)\cos L + f \\ -(h\sin L - k\cos L)g \end{bmatrix} \qquad\qquad \text{Equation 97}$$

$$\frac{dg}{dt} = \sqrt{\frac{p}{\mu}}\frac{1}{w}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} -w\cos L \\ (w+1)\sin L + g \\ (h\sin L - k\cos L)f \end{bmatrix} \qquad\qquad \text{Equation 98}$$

$$\frac{dh}{dt} = \sqrt{\frac{p}{\mu}}\frac{s^2}{2w}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \cos L \end{bmatrix} \qquad\qquad \text{Equation 99}$$

$$\frac{dh}{dt} = \sqrt{\frac{p}{\mu}}\frac{s^2}{2w}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ \sin L \end{bmatrix} \qquad\qquad \text{Equation 100}$$

$$\frac{dL}{dt} = \sqrt{\mu p}\left(\frac{w}{p}\right)^2 + \sqrt{\frac{p}{\mu}}\frac{1}{w}\begin{bmatrix} R & T & N \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ (h\sin L - k\cos L) \end{bmatrix} \qquad \text{Equation 101}$$

Notice, when the disturbing acceleration is zero Equation 96 – Equation 100 equal zero, while Equation 101 reduces to the angular momentum term. The validity of Equation 96 - Equation 101 was demonstrated in Reference 181; although typographical errors are present in the equations of motion presented in Reference 181 these errors have since been corrected.[182]

3.2.2   Transformation from Modified Equinoctial Elements to Classical Elements

The inverse transformation of Equation 85 – Equation 90 are obtained as,

$$a = \frac{p}{1 - f^2 - g^2} \qquad\qquad\qquad \text{Equation 102}$$

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

Equation 103

$$i = 2\arctan(\tau) = 2\arctan\left(\sqrt{h^2 + k^2}\right)$$

Equation 104

$$\Omega = \arctan\left(\frac{k}{h}\right)$$

Equation 105

$$\tan(\omega + \Omega) = \frac{g}{f}$$

Equation 106

$$\omega = \arctan\left(\frac{g}{f}\right) - \arctan\left(\frac{k}{h}\right)$$

Equation 107

$$\nu = L - \arctan\left(\frac{g}{f}\right)$$

Equation 108

Using Equation 102 – Equation 108 one can derive the following identities which are useful in the derivation of the locally optimal control laws in Section 2.

$$\cos\nu = \frac{f\cos L + g\sin L}{\sqrt{f^2 + g^2}}$$

Equation 109

$$\sin\nu = \frac{f\sin L - g\cos L}{\sqrt{f^2 + g^2}}$$

Equation 110

$$\cos\Omega = \frac{h}{\tau}$$

Equation 111

$$\sin\Omega = \frac{k}{\tau}$$

Equation 112

$$\cos i = \frac{1 - \tau^2}{1 + \tau^2}$$

Equation 113

$$\sin i = \frac{2\tau}{1 + \tau^2}$$

Equation 114

$$\cos\omega = \frac{fh + gk}{\tau\sqrt{f^2 + g^2}}$$

Equation 115

$$\sin \omega = \frac{gh - fk}{\tau \sqrt{f^2 + g^2}}$$

Equation 116

$$\cos(\omega + \upsilon) = \frac{h \cos L + k \sin L}{\tau}$$

Equation 117

$$\sin(\omega + \upsilon) = \frac{h \sin L - k \cos L}{\tau}$$

Equation 118

3.2.3   Transformation from Modified Equinoctial Elements to Cartesian Form

Using the relationships outlined in Section 3.2.2 one can define Cartesian state vectors in modified equinoctial elements as,

$$\mathbf{r} = \frac{r}{s^2} \begin{bmatrix} \cos L + \alpha^2 \cos L + 2hk \sin L \\ \sin L - \alpha^2 \sin L + 2hk \cos L \\ 2(h \sin L - k \cos L) \end{bmatrix}$$

Equation 119

$$\mathbf{v} = -\frac{1}{s^2} \sqrt{\frac{\mu}{p}} \begin{bmatrix} \alpha^2 \sin L + \sin L - 2hk \cos L + g - 2fhk + \alpha^2 g \\ \alpha^2 \cos L - \cos L + 2hk \sin L - f + 2ghk + \alpha^2 f \\ -2(h \cos L + k \sin L + fh + gk) \end{bmatrix}$$

Equation 120

Note that it is also possible to compute the inverse transformation; however the true longitude can only be defined to within a multiple of 2π and thus the reference epoch must be known to resolve its actual value.[182]

3.3   Numerical Integration

Numerical integration methods can be divided into either the single-step or multi-step class. The difference between these two methods is well illustrated in Reference 168. However, one can summarise the difference by noting that a single-step method is a self-starting method which only uses data from the beginning of the current step in the calculation of the variable values at the end of the step. Furthermore, changing the step-size to match a defined error criterion poses no difficulties, allowing the interval step-size to easily be halved or doubled. The primary difficulty with a single-step method is that if the equations are non-linear, such as Lagrange's planetary equation of motion, then it may become a time-consuming and unwieldy process to calculate the higher order terms of the expansion.[168] A multi-step method allows larger interval step-sizes to be adopted even when the higher order terms of the expansion are calculated. However, the law of diminishing returns sets in. Furthermore, stability considerations mean that it is wise to keep the order below double figures.[168] A multi-step procedure involves fewer computations than a single-step method, correct to the same order, subject to the constraint of not being self-starting and that special procedures are required to half or double the step-size. Therefore multi-step methods are best suited to scenarios where the step-size changes can be removed or minimised, such as almost circular orbits or when the equations have been regularised. Within this dissertation the equations of motion are propagated using an explicit, variable step

size Runge-Kutta (4, 5) formula, the Dormand-Prince pair; a single-step method.[183] Relative and absolute error tolerances of 10^{-8} and 10^{-12} are set for planet-centred and Sun-centred calculations respectively.

3.4   Orbit Perturbations

An unperturbed trajectory can be propagated using Equation 96 – Equation 101, thus by calculating the perturbing force on each of the *RTN* axis a solar sail trajectory can be propagated by considering the sail thrust a perturbation. Similarly, the inclusion of third body gravity effects and other relevant orbit perturbations can be included within the trajectory model. The sail thrust perturbation is denoted the "*primary*" perturbation, thus models with only the sail thrust perturbation included are referred to as "*unperturbed*". The inclusion of orbit perturbations is intended to demonstrate the ability of the trajectory design methods developed in later chapters to cope with non-ideal scenarios, rather than to actually provide definitive answers to the effects of such perturbations.

3.4.1   Solar Sail Thrust

3.4.1.1   Ideal Model

The orientation of the solar sail control angles, which define the sail normal, are given in Figure 12. The sail control angles are defined within the *RTN* reference frame for heliocentric trajectories, thus the sail normal vector given in Equation 121 can be directly applied to Equation 22 (page 15) to find the sail thrust vector for input into the equations of motion in Equation 96 – Equation 101.

$$\mathbf{n} = \begin{bmatrix} \cos\alpha \\ \sin\alpha\sin\delta \\ \sin\alpha\cos\delta \end{bmatrix} \qquad \text{Equation 121}$$

If the sail is in a planet-centred trajectory then the normal vector orientation given by Equation 121 must be transformed into planet-centred *RTN* axis, that is to say, the same coordinate system as the equations of motion. This transformation is performed in two steps and assuming the Sun-sail line is coincident with the Sun-planet line; the first step converts the thrust vector orientation into planet-centred inertial coordinates, the second then converts from planet-centred inertial coordinates into *RTN* coordinates. The transformation from planet-centred inertial (for instance, Earth-centred inertial) to Sun-line coordinates is performed as a rotation about the x-axis through the obliquity of the ecliptic and then a rotation about the new z-axis through an angle measured from the first point of Aries to the planet. The transformation matrix is thus found to be,

$$O_{SUN} = [T]O_{ECI} = \begin{bmatrix} \cos\vartheta & \sin\vartheta\cos\varepsilon & \sin\vartheta\sin\varepsilon \\ -\sin\vartheta & \cos\vartheta\cos\varepsilon & \cos\vartheta\sin\varepsilon \\ 0 & -\sin\varepsilon & \cos\varepsilon \end{bmatrix} O_{ECI} \qquad \text{Equation 122}$$

The transformation from Sun-line coordinates to planet-centred inertial coordinates is thus simply the inverse of the transformation matrix given in Equation 122.

The transformation from Earth-centred inertial (*ECI*) to *RTN* coordinates is a standard transformation matrix.[184] Application of Equation 122 and Equation 123 to Equation 121 gives the ideal sail thrust vector in the same coordinates as the equations of motion are defined, allowing the sail thrust effect to be incorporated into the trajectory model using Equation 22 (page 15).

$$O_{RTN} = [T] O_{ECI}$$

$$= \begin{bmatrix} \cos(\omega+\nu)\cos\Omega - \sin(\omega+\nu)\cos i \sin\Omega & \cos(\omega+\nu)\sin\Omega + \sin(\omega+\nu)\cos i \cos\Omega & \sin(\omega+\nu)\sin i \\ -\sin(\omega+\nu)\cos\Omega - \cos(\omega+\nu)\cos i \sin\Omega & -\sin(\omega+\nu)\sin\Omega + \cos(\omega+\nu)\cos i \cos\Omega & \cos(\omega+\nu)\sin i \\ \sin i \sin\Omega & -\sin i \cos\Omega & \cos i \end{bmatrix} O_{ECI}$$

Equation 123

The implementation of the solar sail force model was validated through comparison with analytical expressions for the effect of solar radiation pressure on spacecraft and by comparison with previously published work as illustrated in Figure 23.[1, 35, 69, 168, 169, 174, 185]

3.4.1.2   Non-Ideal Model

The orientation of the sail control angles define the sail thrust vector orientation and magnitude using Equation 39 – Equation 42 for a non-ideal sail; the sail optical parameters required by these Equations are detailed in Table 1. Note the degradation model discussed in Section 1.3.4 may also be applied to these coefficients. Equation 22 cannot be used as the direction of the force vector will not be normal to the sail surface, as discussed in Section 1.3.3 and illustrated in Figure 11. Equation 43 gives the centre-line angle for the sail thrust vector, which thus allows the sail cone angle to be defined. Recall, the cone angle is the angle from the Sun-line to the sail thrust vector. With the direction and magnitude of the sail thrust defined following Section 1.3.3.1 it becomes a simple matter of ensuring the vector is defined within the same reference frame as the equations of motion prior to applying the sail thrust vector to them. The conversion from Sun-line coordinates through planet-centred inertial coordinates to *RTN* coordinates follows that outline in Section 3.4.1.1. Note that the sail acceleration input to the coding for a non-ideal sail with the optical parameters defined in Table 1 is related to the characteristic acceleration by an efficiency factor, η, of 0.908156. The efficiency factor is due to non-perfect sail reflectivity at zero pitch. This distinction will be noted where required throughout the later sections such as to maintain a consistent definition of characteristic acceleration.

3.4.2   Occulation of Sunlight

Knowledge of any period of sunlight occulation is imperative during solar sail trajectory design. Assuming no secondary propulsion system is included within the spacecraft design no thrust is available for either orbit manoeuvring or attitude control during periods of shadow passage. The shadow model considers both umbra and penumbra shadow conditions, as illustrated in Figure 36. The theory is presented for Earth shadow, with the position vectors given in a geocentric frame; however the theory is easily extended to any other planet or to lunar shadow were position vectors are planet-centred or Selenocentric respectively. Note that while
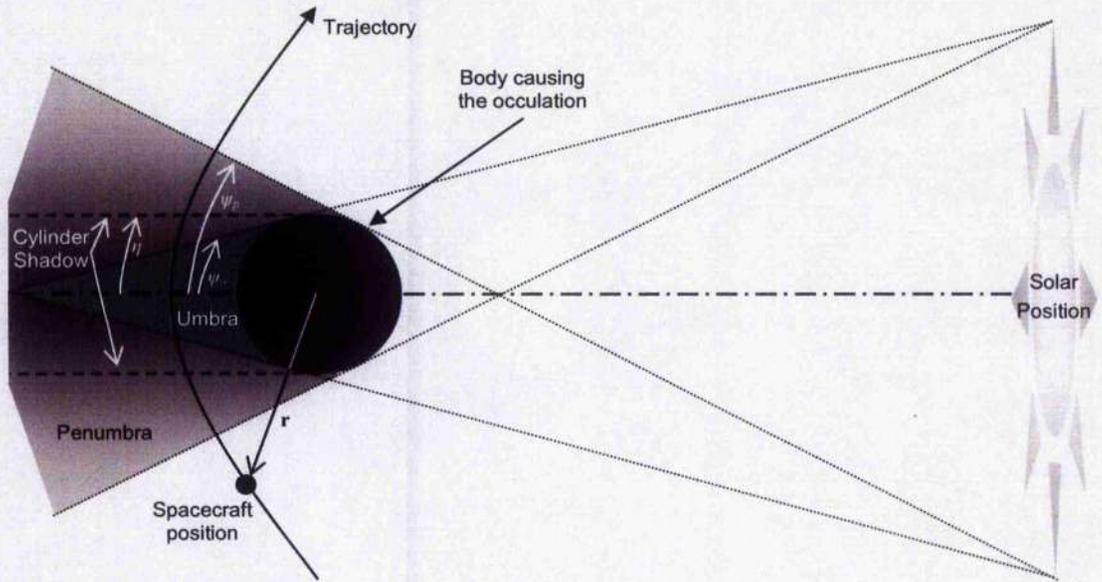
Figure 36 Shadow conditions. Not to scale.

sunlight occulation could potentially exist for heliocentric trajectories it is at no time included within the heliocentric model and as such shadow effects are only ever considered within planet-centred trajectories.

The shadow parameter is defined as,

$$\varsigma = -\frac{|\mathbf{r} \times \mathbf{r_h}|}{|\mathbf{r_h}|} \operatorname{sgn}(\mathbf{r} . \mathbf{r_h})$$

Equation 124

Within this analysis the coordinate system used is geocentric, thus $\mathbf{r}$ is the orbit radius from the shadow casting body, the Earth, while $\mathbf{r_h}$ is the radius of the spacecraft from the Sun. The critical shadow parameters for umbra and penumbra, respectively, are

$$\varsigma_u = |\mathbf{r}| \sin \psi_u$$

Equation 125

$$\varsigma_p = |\mathbf{r}| \sin \psi_p$$

Equation 126

The umbra and penumbra shadow angles, shown in Figure 36, are found to be,

$$\psi_u = \eta + \theta_u$$

Equation 127

$$\psi_p = \eta + \theta_p$$

Equation 128

where η is found by,

$$\eta = \arcsin\left(\frac{1.02 R_E}{|\mathbf{r}|}\right)$$

<div align="right">Equation 129</div>

Note that the cylinder shadow angle is found using the radius of the Earth plus 2 % to account for the Earth's atmosphere; this correction factor must be adjusted for each body considered. The umbra and penumbra cone angles, required in Equation 127 and Equation 128 are,

$$\theta_u = \arcsin\left(\frac{R_S + R_E}{\mathbf{R}_e}\right)$$

<div align="right">Equation 130</div>

$$\theta_p = \arcsin\left(\frac{R_S - R_E}{\mathbf{R}_e}\right)$$

<div align="right">Equation 131</div>

Thus, the shadow condition is defined through comparison of the critical shadow parameters with the shadow parameter, as detailed in Table 4. If the sail is in the umbra then the sail acceleration is set at zero. If the sail is in the penumbra the amount of visible sunlight is determined by linear interpolation of the shadow parameters and the sail acceleration is scaled accordingly, linear interpolation can be used due to the very narrow nature of the penumbra.

3.4.3   Third-Body Gravity Effects

The equations of motion in Equation 96 – Equation 101 model the spacecraft trajectory in a 2-body scenario; however Newton's law of universal gravitation states that all bodies exert a gravitational pull on each other. Most objects can be neglected from trajectory calculations due to the very small magnitude of the force which they exert, however in certain scenarios such simplifications cannot be made. For example, in Earth orbit the Moon and the Sun can exert a force of similar, or greater, magnitude to that of a solar sail. The force of n bodies on the spacecraft can be written as,[174]

$$\mathbf{a_j} = G\sum_{j}^{n} m_j \left(\frac{\mathbf{r_j} - \mathbf{r}}{|\mathbf{r_j} - \mathbf{r}|^3} - \frac{\mathbf{r_j}}{|\mathbf{r_j}|^3}\right)$$

<div align="right">Equation 132</div>

Parameter Condition	Shadow Condition
$0 \leq \varsigma \leq \varsigma_u$	In umbra
$\varsigma_u \leq \varsigma \leq \varsigma_p$	In penumbra
$|\varsigma| > |\varsigma_p|$	In full sunlight
$\varsigma < 0$	In full sunlight

Table 4 Shadow definition criteria by parameter comparison.

Thus, if the vectors **r** and **r**ⱼ are given in, say, *ECI* coordinates then the perturbation on the spacecraft is also given in *ECI* coordinates; requiring the perturbation vector be converted into *RTN* coordinates using Equation 123.

The implementation of third-body gravity effects into the trajectory model was validated through comparison with analytical solutions to the problem. Using Lagrange's planetary equations (Equation 65 – Equation 69) analytical expressions for the secular variation of orbit eccentricity, inclination, ascending node and argument of perigee can be determined in terms of directional cosines of the disturbing body.[186, 187]

3.4.4   Planetary Oblateness Effects

The selected non-singular elements allow for the use of Allan's expansion of the geopotential,[179] however several other models are available, including a model that gives the geopotential entirely in terms of non-singular orbital elements.[188] While accuracy of model and model integrity are of great importance, recall that the trajectory calculations within this dissertation are intended to demonstrate a trajectory design method and not create highly accurate sail trajectories, as would be required in a real-mission. As such the geopotential is modelled using a standard spherical harmonic representation due to the speed of calculation that this method allows.[168, 169, 174, 184, 187]

The use of perturbation equations of motion means the Earth's geopotential function is modelled with only the high order terms. The geopotential function is,

$$
\Phi(r,\phi,\lambda) = \frac{\mu}{r} \sum_{n=1}^{\infty} C_n^0 \left(\frac{R_E}{r}\right)^n P_n^0(u)
$$
$$
+ \frac{\mu}{r} \sum_{n=1}^{\infty} \sum_{m=1}^{n} \left(\frac{R_E}{r}\right)^n P_n^m(u) \left[S_n^m \sin(m\lambda) + C_n^m \cos(m\lambda) \right]
$$

Equation 133

where,

$$
\phi = \arcsin\left(\frac{2(h\sin L - k\cos L)}{s^2}\right)
$$

Equation 134

The gravitational harmonic coefficients are taken from the Earth Gravity Model 1996 (EGM 1996), the model uses an 18×18 matrix for the harmonic coefficients.[184] The partial derivatives of the geopotential function are,

$$
\frac{\partial \Phi}{\partial r} = -\frac{1}{r}\left(\frac{\mu}{r}\right) \sum_{n=2}^{N} \left(\frac{R_e}{r}\right)^n (n-1) \sum_{m=0}^{n} \left(C_n^m \cos m\lambda + S_n^m \sin m\lambda\right) P_n^m(u)
$$

Equation 135

$$\frac{\partial \Phi}{\partial \phi} = \left(\frac{\mu}{r}\right)\sum_{n=2}^{N}\left(\frac{R_e}{r}\right)^{n}\sum_{m=0}^{n}\left(C_n^m \cos m\lambda + S_n^m \sin m\lambda\right)\left(P_n^{m+1}(u) - m\tan\phi P_n^m(u)\right)$$ Equation 136

$$\frac{\partial \Phi}{\partial \lambda} = \left(\frac{\mu}{r}\right)\sum_{n=2}^{N}\left(\frac{R_e}{r}\right)^{n}\sum_{m=0}^{n} m\left(S_n^m \cos m\lambda - C_n^m \sin m\lambda\right)P_n^m(u)$$ Equation 137

Thus, the disturbing accelerations in ECI coordinates can be found to be,

$$\ddot{x} = \left(\frac{1}{r}\frac{\partial\Phi}{\partial r} - \frac{z}{r^2\sqrt{x^2+y^2}}\frac{\partial\Phi}{\partial\phi}\right)x - \left(\frac{1}{x^2+y^2}\frac{\partial\Phi}{\partial\lambda}\right)y$$ Equation 138

$$\ddot{y} = \left(\frac{1}{r}\frac{\partial\Phi}{\partial r} - \frac{z}{r^2\sqrt{x^2+y^2}}\frac{\partial\Phi}{\partial\phi}\right)y + \left(\frac{1}{x^2+y^2}\frac{\partial\Phi}{\partial\lambda}\right)x$$ Equation 139

$$\ddot{z} = \left(\frac{1}{r}\frac{\partial\Phi}{\partial r}\right)z + \left(\frac{\sqrt{x^2+y^2}}{r^2}\frac{\partial\Phi}{\partial\phi}\right)$$ Equation 140

Finally, the geopotential disturbing function accelerations are found by transformation of this vector into the geocentric satellite RTN reference frame, using Equation 123. The Earth's non-spherical perturbations were modelled using an adapted subroutine taken from Reference 184 and similar to the third-body gravity effects the implementation was validated using analytical approximations generated from Lagrange's planetary equations.[168, 169, 174, 187] Note that at Mercury the reciprocal of flattening is over eighteen times that of Earth, with a J_2 value of only 60×10^{-6}. Thus Mercury oblateness is modelled using only the J_2 coefficient.

Following validation of the coding of each individual perturbation discussed above (Sections 3.4.1 – 3.4.4) a further all inclusive test was performed. The trajectory simulation results presented in Reference 102 were generated using a commercial trajectory calculation software tool which allows the combined effects of the orbit perturbations examined in Sections 3.4.1 – 3.4.4 to be analysed. The results generated in Reference 102 were thus reproduced by the candidate in Reference 163 using the methods discussed within this dissertation as further validation of implementation.

3.4.5   Planetary Atmosphere Effects

The atmosphere of a body results in aerodynamic forces on any other body which passes close enough to it. The principle aerodynamic load is typically due to drag, acting in the opposite sense to the spacecraft motion and resulting in a reduction in orbit energy. Spacecraft also experience an aerodynamic lift effect; however this is normally small and hence neglected.[168, 169, 174, 187] Aerodynamic forces are difficult to analyse for conventional spacecraft and even more difficult for a class of spacecraft which has never flown, for instance a solar sail. The atmosphere is a dynamic phenomenon which is influenced by many parameters, such as solar activity, latitude, attitude, atmospheric rotation and tides, the diurnal bulge and seasonal

variations.[168, 169, 174, 187] Recalling once again that the trajectory calculations within this dissertation are intended to demonstrate a trajectory design method and not create highly accurate sail trajectories, atmospheric loads are not included in any trajectory model within this dissertation due to the difficulties outlined above. As such, a brief analysis is presented to quantify the limiting effect atmospheric drag would have on solar sail trajectories. Modelling the atmosphere as a free molecular flow and the solar sail as a rigid plate, the lift and drag coefficients, for a flat plate in a hypersonic flow, are,

$$C_L = 2\sin^2 \alpha_h \cos \alpha_h \qquad\qquad \text{Equation 141}$$

$$C_D = 2\sin^3 \alpha_h \qquad\qquad \text{Equation 142}$$

where α_h is the pitch of the sail with respect to the hypersonic flow. As would be expected the maximum atmospheric drag occurs when $\alpha_h = {}^\pi/_2$, when the sail is face-on to the flow; at maximum drag the sail experiences no lift force. Note the definition of pitch used within this short sub-section differs from that used for sail pitch within the rest of this dissertation.

Assuming a circular orbit and that C_D is approximately 2, that is to say the drag coefficient of a flat plate, the maximum drag pressure on the plate is,

$$D_{max} = \frac{1}{2}\rho v^2 C_D = \rho v^2 \qquad\qquad \text{Equation 143}$$

For a given sail loading the maximum drag acceleration for a given orbit radius is,

$$a_{drag} = \rho v^2 \left(\frac{A}{m}\right) = \frac{\rho v^2}{\sigma} \qquad\qquad \text{Equation 144}$$

Assuming a uniform, mean density profile based on the hydrostatic equilibrium of an isothermal atmosphere, the atmospheric density can be approximated as a function of altitude as,

$$\rho(h) = \rho_0 \exp\left(\frac{-(h - h_0)}{H_0}\right) \qquad\qquad \text{Equation 145}$$

where $\rho_0 = 7.25 \times 10^{-10}$ kg m^{-3}, the reference density, $h_0 = 140$ km, the reference altitude and $H_0 = 11.127$ km, the reference density scale height. Figure 37 gives the maximum atmospheric drag for a given altitude over a range of solar activity levels. Note that the perturbation is a strong function of solar activity, which is modulated in an 11-year cycle. At periods of low solar activity the atmospheric drag and sail acceleration balance at ~ 430 km, as shown in Figure 37. At mean solar activity the balance points rises to ~ 560 km and on up to ~ 940 km at high activity periods. A safe mean altitude is above 800 km, although at times of solar maximum this may double. Furthermore, it is noted that at no time can a solar sail be expected to survive at an altitude similar to that of the International Space Station (ISS) as the atmospheric drag would render the

sail uncontrollable. Similar analysis shows that at Mars the minimum solar sail altitude is ~ 300 km, while at Venus it is ~ 900 km.

3.4.6   Solar Wind and Other Forces

The solar wind would be expected to exert a small force due to the momentum transported by particles within it. At high solar wind speeds the mean proton number density at Earth is of order 10^6 m^{-3}, with a wind speed of order 700 km s^{-1}. The solar wind pressure exerted on the solar sail can thus be estimated from the transported momentum. It can be shown that a solar wind pressure of order 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.[1] The solar wind is thus of negligible effect when it comes to trajectory design, however it is an important perturbation when considering attitude control as it has a significant accumulative effect on sail attitude.

First order relativistic effects are proportional to the ratio of the solar sail velocity to the speed of light; typically once again this is of order 10^{-4} and can be neglected.[1] For solar sails in planetary orbit the secondary pressure due to radiation scattered from the planet is also small, typically at least three orders of magnitude less than that due to the direct solar radiation pressure. Although it should be noted that above the sub-solar point of Mercury the reflected radiation can be significantly larger than at other planets. No account is taken as to the effect of reflected radiation.

3.5   Trajectory Model Names

For simplicity of reference each of the trajectory models used within the dissertation have been given a name in Table 5. Note that the central body of each model is not included as the central body should be obvious from the context in which the model is used.
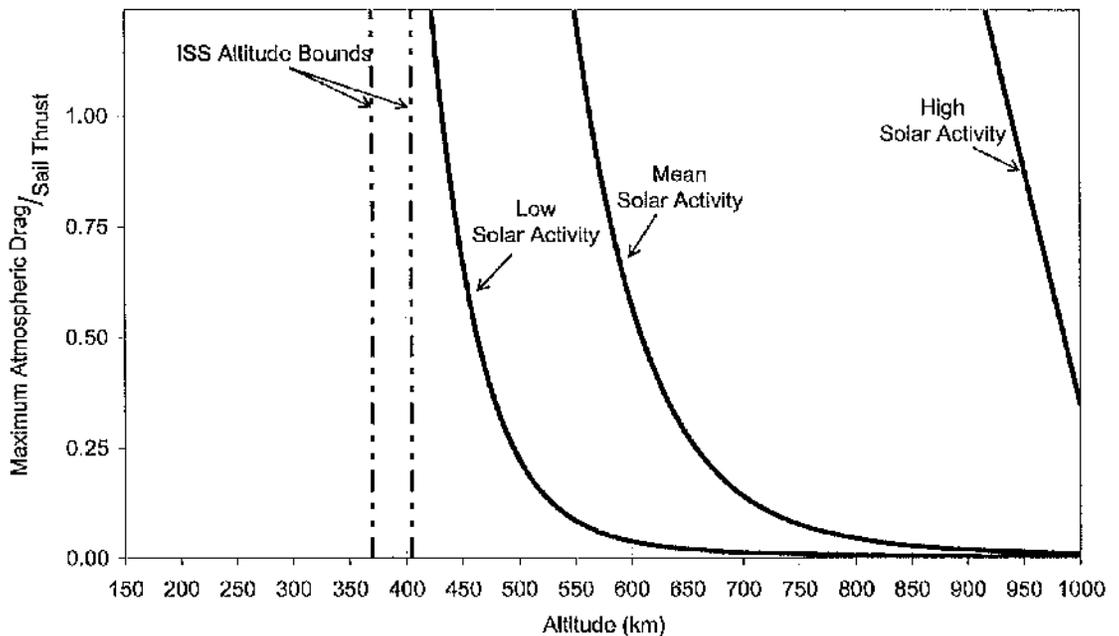
Figure 37 Effect of Earth's atmosphere on a solar sail at a range of altitudes and solar activities.

Name	Description
Model 1	Unperturbed solar sail trajectory model. Ideal sail force model and a point source Sun. All periods of occulation are neglected. The sail – Sun distance is corrected for the true distance.
Model 2	As for Model 1, but with consideration of occulation due to the central planet and its natural satellites as discussed in Section 3.4.2. This model is used only in planet-centred trajectories within this dissertation.
Model 3	Perturbed solar sail trajectory model; modelling the perturbations discussed in Sections 3.4.3 (natural satellites and Sun only) and Section 3.4.4. Occulation due to the central-body and its natural satellites are included. The Sun is modelled as a uniformly bright finite disc as discussed in Section 1.3.5. This model is used only in planet-centred trajectories within this dissertation.
Model 4	As for Model 1, but with a non-ideal sail.
Model 5	As for Model 1, but with the Sun modelled as a uniformly bright finite disc as discussed in Section 1.3.5. This model is used only in Sun-centred trajectories within this dissertation.
Model 6	As for Model 3, but with a non-ideal sail. Once again this is exclusively planet-centred.
Model 7	As for Model 4, but with optical surface degradation modelled following Section 1.3.4.

Table 5   Description of trajectory models.

4   Planetary Escape Using a Solar Sail and Locally Optimal Control Laws

With the increased interest in solar sailing from the science community (see Section 1.5) it becomes necessary to clarify previous anomalies observed in mathematical models and eliminate previous simplifications in order to aid future analysis of solar sailing missions. It has been noted that slight variations, up to 7 %, exist for Earth escape times depending on the launch date. However, no adequate explanation has been offered for the presence of this variation.[33 - 35] The number of eclipse events or the rate of energy gain by the sail have both been suggested as possible explanations. It is noted in Reference 33 that the variation appears to fall as sail acceleration increases. However, it is hypothesised in Reference 35 that an increase in sail acceleration may extenuate the variation, presumably as a greater proportion of the much shorter trajectory will now be in shadow. Furthermore, it has been noted that for geocentric spiral trajectories to both the lunar distance and a sub-escape point, the time of flight tends to be minimum for orbits within the ecliptic plane. It will be shown in this section that these two factors are related by the derivation of an optimal inclination for solar sail manoeuvring and that the effect is amplified at low sail accelerations, thus solving the anomaly.[163, 189]

A simple, autonomous solution to the problem of planetary escape will then be developed through the use of blending different locally optimal control laws. The use of blended control allows for a more realistic set of orbit goals to be defined than just simple energy gain. For example, a minimum pericentre altitude can be set, thus ensuring the sail remains above the upper-atmosphere or to expedite the sail orbit out of this region and then maintain so thereafter. Following the generation of algorithms to provide safe planetary escape trajectories; Earth escape trajectories without Earth occultation of the sail are examined. This is a potentially beneficial scenario for attitude control, thermal and other sub-system design.

4.1   Optimal Inclination for Planet-Centred Solar Sailing

Using a different derivation of the locally optimal energy-gain steering law to that in Section 2.3, the rate of energy variation is shown to be related to both the sail acceleration and the orbit inclination; thus confirming the presence of a theoretically optimal inclination. The time until Earth escape is then investigated over a range of sail accelerations and inclinations using Model 1, in order to corroborate the effect of orbit inclination and consequently time of year on escape duration. The effect of introducing Earth-eclipse is investigated, using Model 2, in order to understand and quantify the effect this may have on escape times throughout the year.

In order to derive the locally optimal energy gain control law, following Reference 35, the definition of a Sun – Vector coordinate system is required; the system is illustrated in Figure 38. The origin of the Sun – Vector coordinate system is defined as the spacecraft centre-of-mass and aligns the positive X_{sun} axis with the instantaneous direction of the Sun. The Z_{sun} axis is defined as the cross product of the velocity vector and the Sun unit vector; with the Y_{sun} axis completing the right-hand Cartesian coordinate system. Thus the Z_{sun} axis velocity component is always zero in the Sun – Vector coordinate system. The sail clock angle is taken

from the $\mathbf{Y_{sun}}$ axis towards the projection of the sail normal into the plane defined by the $\mathbf{Y_{sun}}$ and $\mathbf{Z_{sun}}$ axis, as seen in Figure 38. The sail normal vector, within the Sun – Vector coordinate system, is described as,

$$\mathbf{n} = \begin{bmatrix} \cos\alpha \\ \sin\alpha\cos\delta \\ \sin\alpha\sin\delta \end{bmatrix}$$

Equation 146

Note Equation 146 differs from Equation 121 due to the different definition of the coordinate system. Furthermore, when the velocity vector and the Sun unit-vector are parallel this coordinate system is not defined, thus the system is used only for the development of theory and not for orbit propagation.

Using the Sun-Vector coordinate system and following Reference 35 the function to be maximised is defined as,

$$F(\alpha,\delta) = \mathbf{a_s}.\mathbf{v}$$

Equation 147

where, $\mathbf{v} = [v_x \ v_y \ v_z]$. Assuming a circular Earth orbit the sail acceleration vector is defined as $\mathbf{a_s} = a_{s_c}(\mathbf{n}.\mathbf{X_{sun}})^2\mathbf{n}$, combining this with Equation 146 the sail acceleration vector is obtained as,

$$\mathbf{a_s} = a_{s_c}\cos^2\alpha\begin{bmatrix} \cos\alpha \\ \sin\alpha\cos\delta \\ \sin\alpha\sin\delta \end{bmatrix}$$

Equation 148

From the definition of the coordinate system recall that the $\mathbf{Z_{sun}}$ axis velocity component, v_z, is zero, thus,

$$F(\alpha,\delta) = a_{s_c}\left(v_x\cos^3\alpha + v_y\cos^2\alpha\sin\alpha\cos\delta\right)$$

Equation 149
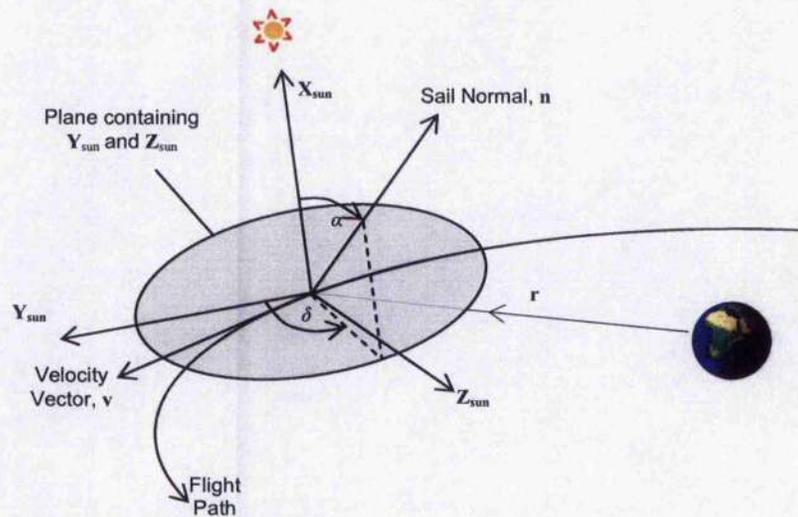
Figure 38 The Sun-Vector Coordinate System, with the plane normal to $\mathbf{X_{sun}}$ illustrated.

Forming the first derivatives with respect to α and δ gives the conditions for a turning point of the function.

$$\frac{\partial F(\alpha,\delta)}{\partial \delta} = -a_{s_c} v_y \cos^2 \alpha \sin \alpha \sin \delta = 0 \qquad \text{Equation 150}$$

$$\frac{\partial F(\alpha,\delta)}{\partial \alpha} = -a_{s_c} \left[3v_x \cos^2 \alpha \sin \alpha + v_y \cos \delta \left(2\cos \alpha \sin^2 \alpha - \cos^3 \alpha \right) \right] = 0 \qquad \text{Equation 151}$$

Rearranging Equation 151, with $\cos^3\alpha \neq 0$ or $\alpha \neq 0$ and $\cos \delta \neq 0$ or $\delta \neq 90$ deg, 270 deg, and $v_y \neq 0$ gives,

$$a_{s_c}\left(\tan^2 \alpha + \frac{3v_x \tan \alpha}{2v_y \cos \delta} - \frac{1}{2} \right) = 0 \qquad \text{Equation 152}$$

Solving for α and δ from Equation 152 and Equation 150 respectively,

$$\alpha_{1,2} = \arctan\left[-\frac{3v_x}{4v_y \cos \delta} \pm \sqrt{\frac{1}{2} + \left(\frac{3v_x}{4v_y \cos \delta} \right)^2} \right] \qquad \text{Equation 153}$$

$$\delta_{1,2} = \arcsin(0) \Rightarrow \delta_1 = 0 \text{deg.}, \delta_2 = 180 \text{deg}. \qquad \text{Equation 154}$$

As obtained in Reference 35, Equation 153 and Equation 154 have two solutions. Equation 153 allows the optimal sail pitch angle to be found for the special case where $v_x = 0$. This angle is often quoted in literature as the optimal fixed sail pitch angle.

$$\alpha_{opt} = \left| \arctan\left(\pm \frac{1}{\sqrt{2}} \right) \right| \cong 35.264 \text{deg}. \qquad \text{Equation 155}$$

The solution for the sail clock angle, Equation 154, states that an optimal steering law is achieved if the sail normal vector, the velocity vector and the Sun vector are all within the same plane. The Sun vector and the velocity vector orientations cannot be altered or optimised. Therefore, the optimal condition defined by Equation 154 can be achieved only by aligning the sail normal vector within the plane defined by the other vectors, thus requiring a fixed sail clock angle of 0 or 180 deg.

Recall from Section 2.3 that the rate of change of semi-major axis depends only on the radial and transverse perturbing accelerations and not on the out-of-plane acceleration. It therefore follows that in order to maximise the rate of change of semi-major axis, and hence orbit energy, the sail force should ideally be oriented entirely within the orbit plane. However, the orbit-plane and the plane defined by the velocity and Sun vectors are coincident only if the sail orbit lies within the ecliptic plane. When the sail orbit is outside the ecliptic plane an angle exists between the orbit plane and the velocity/Sun vector plane. Thus it is not always possible to maximise the sail force within both required planes at all times. In order to ensure the

local maximum rate of energy change, the sail force vector is optimised such that the maximum sail force is directed along the orbit velocity vector. Recall, the sail pitch angle is found using a standard optimisation process given in Equation 63 (page 42). If the orbit-plane is not coincident with the ecliptic plane Equation 63 rotates the sail normal vector out of the orbit-plane towards the plane defined by the velocity and Sun vectors, hence generating an out-of-orbit-plane force. As noted, the rate of change of semi-major axis depends only on the orbit perturbations within the plane of motion and hence the generation of an out-of-plane sail force reduces optimality. Thus, the optimal orbit inclination is defined such that the plane of motion is coincident with the ecliptic plane.

The definition of the ecliptic plane as the optimal plane for solar sail in-plane orbit manoeuvring allows us to explain the apparent seasonal variation of sail escape times from Earth orbit. A 7 deg orbit inclination at the northern hemisphere winter solstice results in an orbit inclination of 16.4 deg from the ecliptic plane. However, at the northern hemisphere summer solstice the inclination to the ecliptic is now 30.4 deg. Thus, the increased inclination with respect to the ecliptic plane should result in a greater out-of-orbit-plane force and hence an increased escape time for a June/July launch, as found in References 33 and 34. As the sail acceleration is increased the difference between escape time for June and December launch should decrease, as the number of orbits until escape is reduced, hence minimising the effect of the out-of-plane sail force. The defined optimal inclination holds true for the locally optimal variation of any orbit element or parameter where the rate of change is dependent on only the in-plane perturbing forces, these include eccentricity, radius of pericentre and radius of apocentre.

4.2   Earth Escape Time, Sail Characteristic Acceleration and Orbit Inclination

Initially Model 1 is used to allow the nature of the relationship between orbit inclination and sail efficiency to be seen without the background effects generated by orbit perturbations. Earth shadow is then introduced through application of Model 2 to provide a comparison, as sail propulsion efficiency is reduced for orbits within the ecliptic plane, due to the large fraction of time spent in Earth's shadow.[35] Such a reduction in sail efficiency could be expected to influence the escape times, hence altering the optimal inclination. However this is shown not to be true. Figure 39 shows the time until escape from GEO radius for a range of sail characteristic accelerations, from 0.15 mm s^{-2} to 2.0 mm s^{-2}, using the semi-major axis controller exclusively. It is seen that for sail characteristic accelerations of 0.75 mm s^{-2} and greater that the minimum escape time corresponds to an orbit inclination within the ecliptic plane, as predicted. The minimum is visible on the surface plot as a groove on the otherwise reasonably smooth surface. However, for sail characteristic accelerations below 0.75 mm s^{-2} the minimum is not evident, though the orbits near the optimal inclination do tend to be the quickest to escape. This breakdown is due to the relatively low level of sail acceleration compared to local gravity and as a result the optimal inclination effect is lost during the high number of orbit revolutions required to gain escape energy. The breakdown in the predicted relationship between sail performance and orbit inclination is reflected by the much more irregular nature of the surface plot at low sail accelerations. Note from Figure 39 that the irregular surface continues into higher sail characteristic accelerations for inclinations between 45 deg and 90 deg. The reason for this anomaly remains unclear and no satisfactory explanation could be derived, however calculation error was eliminated as a possible cause. As the orbit inclination increases from zero to ε the angle between the orbit plane and the plane defined by

the Sun/velocity vectors decreases to zero, hence the optimal inclination. As inclination then continues to increase to 2ε, this angle increases symmetrically with $i < \varepsilon$. This symmetry is reflected in the time until escape, seen by taking a section through Figure 39 at characteristic acceleration 0.75 mm s^{-2}, as shown in Figure 40. As the inclination continues to increase it is seen in Figure 40 that the time until escape continues to rises until the orbit inclination reaches 180 deg. Furthermore, note a change in orbit inclination can be as influential on escape time as a modest increase in sail characteristic acceleration of up to 0.25 mm s^{-2}; this is an important consideration that should be taken into account in the early stages of any solar sail planet-centred mission analysis. Figure 39 and Figure 40 show the exact relationship between sail performance and orbit inclination and clearly shows an optimal sail inclination of $i = \varepsilon$. However, when passing through the Earth's shadow cone no propulsion is provided and sail propulsion efficiency has been shown to be lower for orbits within the ecliptic plane.[35] Therefore, while the basic orbital mechanics suggests an inclination within the ecliptic plane to be optimal, the introduction of Earth shadow could be expected to alter this.

Figure 41 shows the time until escape from GEO radius for a range of sail characteristic accelerations from 0.15 mm s^{-2} to 2.0 mm s^{-2}, using Model 2 and the semi-major axis controller exclusively. Figure 41 shows that shadow does not alter the optimal inclination, with the surface channel still visible at $i = \varepsilon$. However, it is also noticed that the time until escape is increased for orbits near to the ecliptic plane due to the presence of shadow in this region. Figure 41 shows the surface to be much more uneven than before, with the surface remaining irregular up to much higher sail characteristic accelerations. The irregular surface structure has been noted in the shadow free case to be an indicator that the relationship between orbit inclination and sail performance is starting to breakdown. Note once again however that a change in orbit inclination can still be as influential on escape time as an increase in sail characteristic acceleration. Furthermore, the symmetrical nature of escape time about the ecliptic plane is now much more visible than in Figure 39 and the increase in escape times for increasing orbit inclination can once again be seen.

As sail characteristic acceleration is increased it is seen in Figure 42 the difference between escape time at optimal inclination and worst-case inclination is confirmed to fall. This is analogous to the seasonal variation in Earth escape times found previously and confirms that this variation reduces as sail performance is increased.

At Mercury the obliquity of the ecliptic is 0.01 deg, hence the optimal inclination with respect to the planets equator is $i = 0.01$ deg. However, due to the proximity of the optimal inclination to zero, a reduced inclination scan is presented in Figure 43 using Model 1 and the semi-major axis controller exclusively to escape from a 1000 km circular orbit. Additionally, Mercury's highly eccentric orbit, $e = 0.2056$, results in a large variation in solar flux and hence sail acceleration through the Hermian year, thus Figure 43 shows data for two start epochs, corresponding to Mercury perihelion and aphelion passage. Figure 43 shows once again the effect of inclination on escape time, with $i = \varepsilon$ clearly optimal. Once more it is also seen that the relationship breaks down at low sail characteristic accelerations, where the number of orbit revolutions prior to escape is much larger than for the high sail characteristic accelerations.
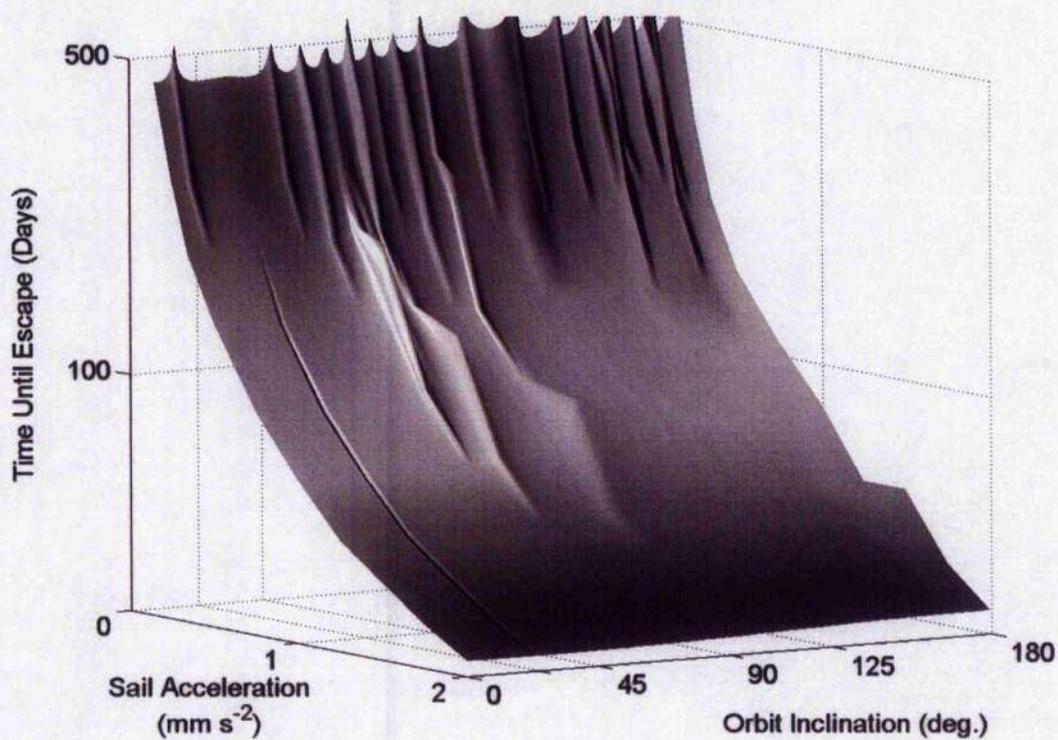
Figure 39   Solar sail escape time from GEO radius, without shadow effects.
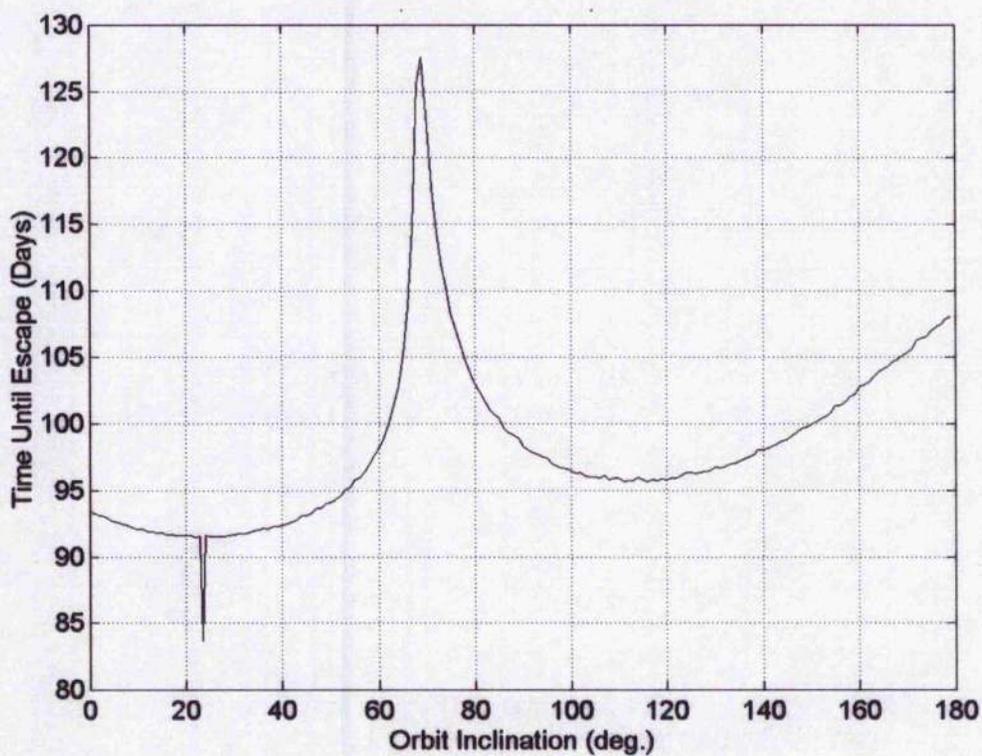
Figure 40   Section through Figure 39 corresponding to sail characteristic acceleration 0.75 mm s^{-2} versus

inclination, without shadow effects.
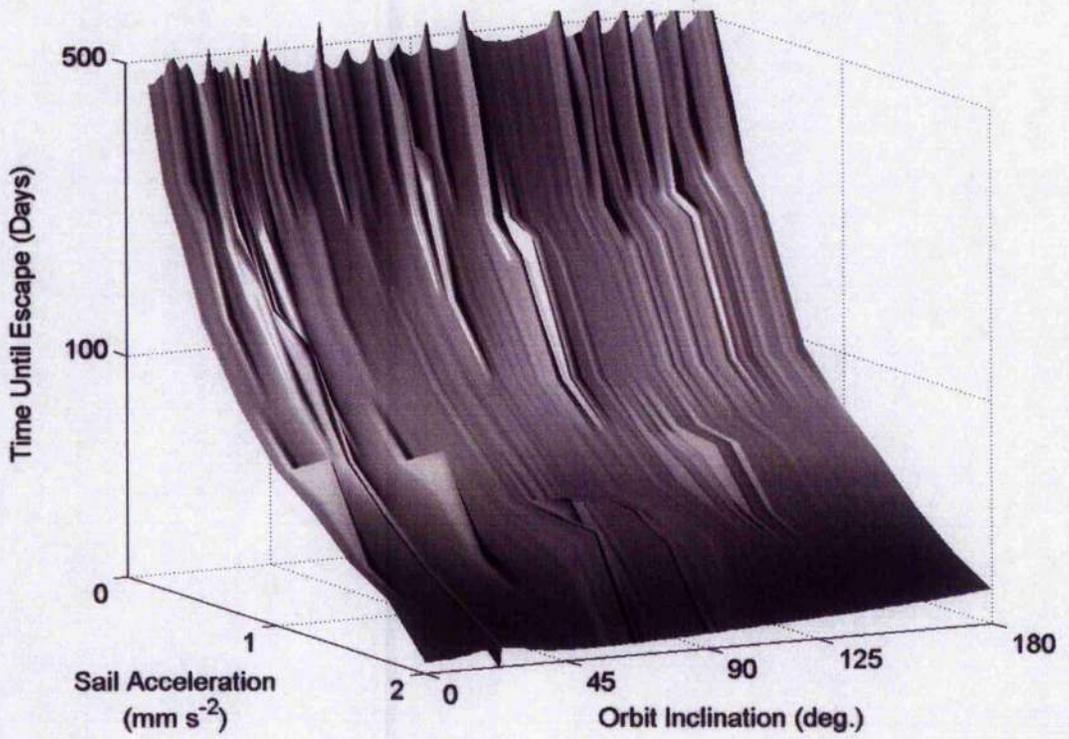
Figure 41 Solar sail escape time from GEO radius, with shadow.
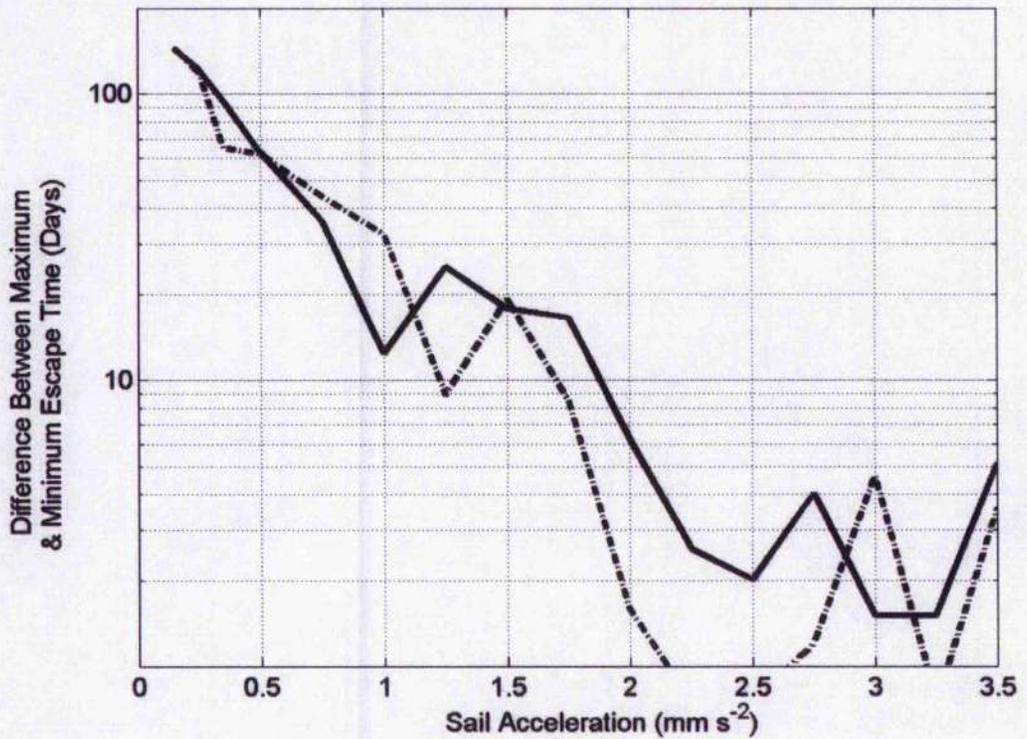
Figure 42 Difference between maximum and minimum escape time. Shadow case (——) and without shadow (– · –).
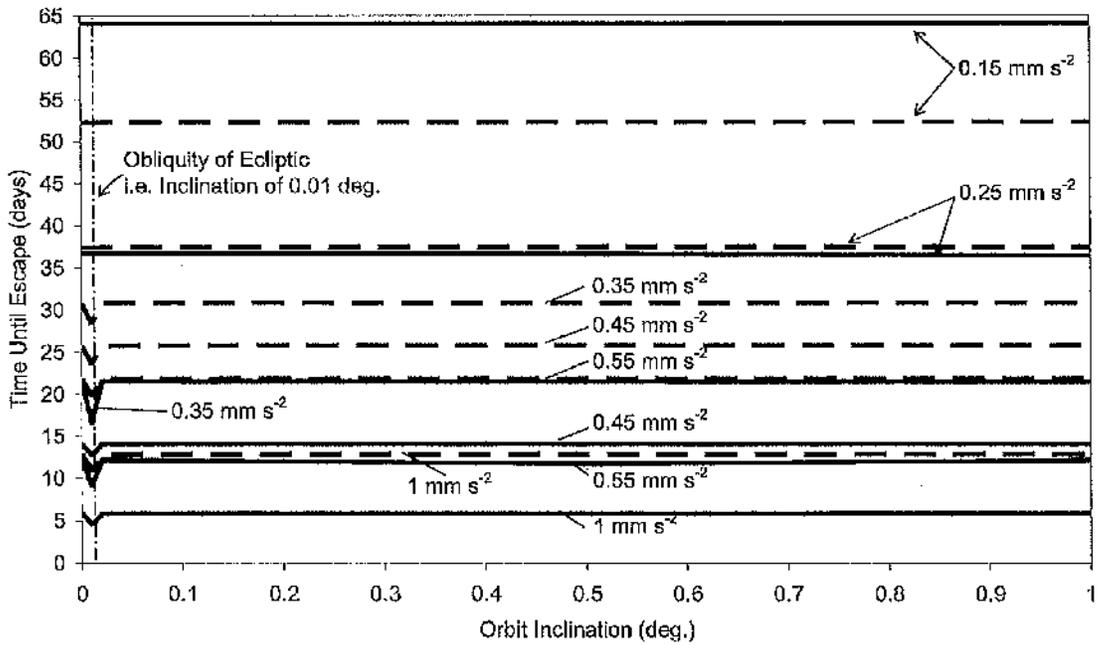
Figure 43   Mercury escape time from 1000 km altitude circular orbit using Model 1. Start epoch coincident with Mercury perihelion passage (—) and aphelion passage (– –) both shown.

The two start epochs used in Figure 43 show an interesting anomaly of trajectory manoeuvres at Mercury when using solar sail propulsion. Note for $a_{s_c} = 0.15$ mm s^{-2} the aphelion start epoch reaches escape twelve days quicker than the perihelion start epoch, however for $a_{s_c} = 1.0$ mm s^{-2} it is seen that the aphelion start epoch is now seven days slower than the perihelion start epoch. Thus, one can say for low sail characteristic accelerations an aphelion start epoch provides escape conditions quickest, yet for high sail characteristic acceleration a perihelion start epoch provides escape conditions quickest. This relationship will now be examined in further detail.

4.3   Mercury Escape Times through the Hermian Year

Figure 44 shows the distance of Mercury from the Sun through one full orbit revolution of 88 days along with the associated solar radiation pressure. From Figure 44 it is clearly seen that solar radiation pressure and hence sail acceleration are maximum at Mercury perihelion, suggesting sail escape trajectories should be of minimum duration during Mercury perihelion passage.

Using Model 1 and the semi-major axis controller exclusively, Mercury escape times were calculated for a fixed initial orbit inclination of 0.01 deg. The calculation start epoch was incremented in one-day intervals from 01 January 2015 (JD 2457023.5) for 115 days, thus the time interval corresponds exactly to that used in Figure 44. Perihelion passage occurs at day 20.86 with aphelion passage 44 days later at day 64.84, while the Hermian year is completed on day 108.83. Figure 45 shows the escape time from a 500 km altitude circular orbit. The time until escape varies with start date in a sinusoidal fashion, with period equal to one Hermian

year. The variation corresponds to the shape of the SRP distribution through the orbit period of Mercury as would be expected. However, it is also noted from Figure 45 that both the contour plot and surface plot are very jagged, with spikes in escape time up to 10 – 50% greater than the mean surface value. These spikes are a unique characteristic of solar sail propulsion and are caused by the inability of a solar sail to gain orbit energy while travelling towards the Sun. Hence, if a sail falls just short of escape energy as it reaches the maximum distance from the Sun it must then complete a half revolution about the planet before gaining the required orbit escape energy. However, a small increase in sail performance will result in the sail acquiring escape energy just before this maximum turning point and time until escape thus appearing much reduced.

Note in Figure 45 the escape time for a low characteristic acceleration sail has maximum just after Mercury perihelion passage, yet for high sail characteristic accelerations the maximum has migrated through the orbit to aphelion passage. Similarly, the minimum escape time for a low characteristic acceleration sail is just after aphelion passage, yet for high sail characteristic accelerations it has migrated to perihelion. The migration of maximum and minimum escape times is shown in Figure 46 for a selection of initial altitudes. Figure 46 shows the day number from Mercury perihelion passage plotted against sail characteristic acceleration. The start date of the maximum duration escape trajectory tends towards Mercury aphelion passage as sail characteristic acceleration is increased. Similarly, the start date of the shortest escape trajectory tends towards Mercury perihelion passage as sail characteristic acceleration is increased. Thus explaining the apparent anomalous situation found in Figure 43.
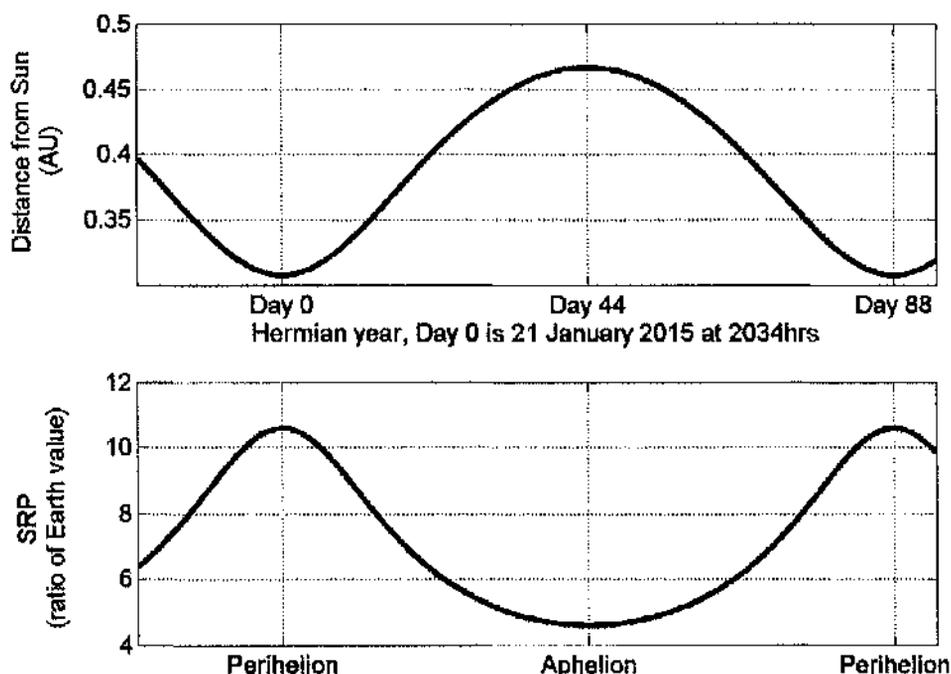
Figure 44 Mercury's heliocentric orbit radius though Hermian year (top) and the associated solar radiation pressure over the same period (bottom).
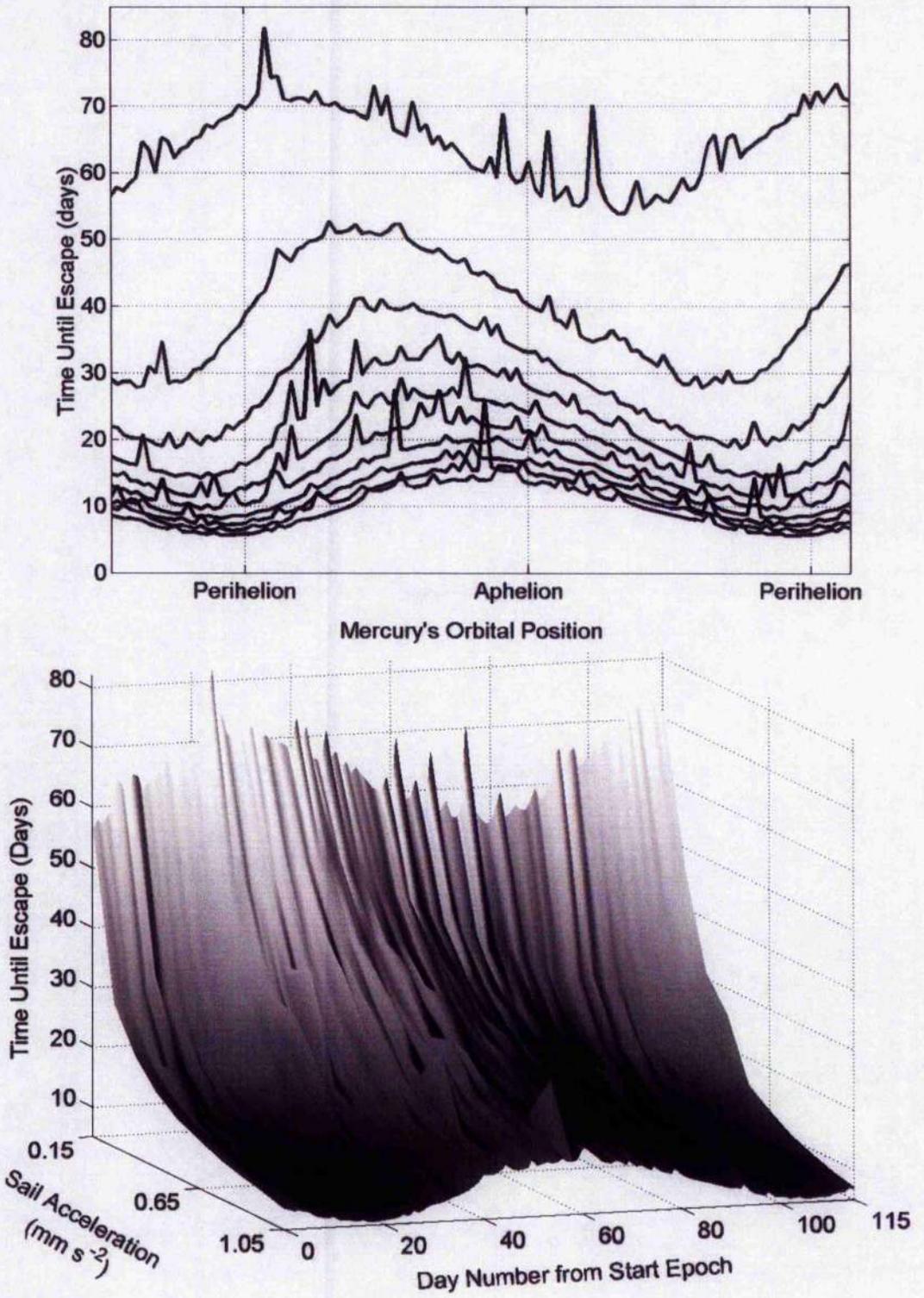
Figure 45   Escape time from ecliptic 500 km circular orbit at Mercury, $a_{s_c} = 0.15$ mm s^{-2} to 1.05 mm s^{-2}, in

0.1 mm s^{-2} increments, from the top.   Surface plot corresponds to same data as contour plot.
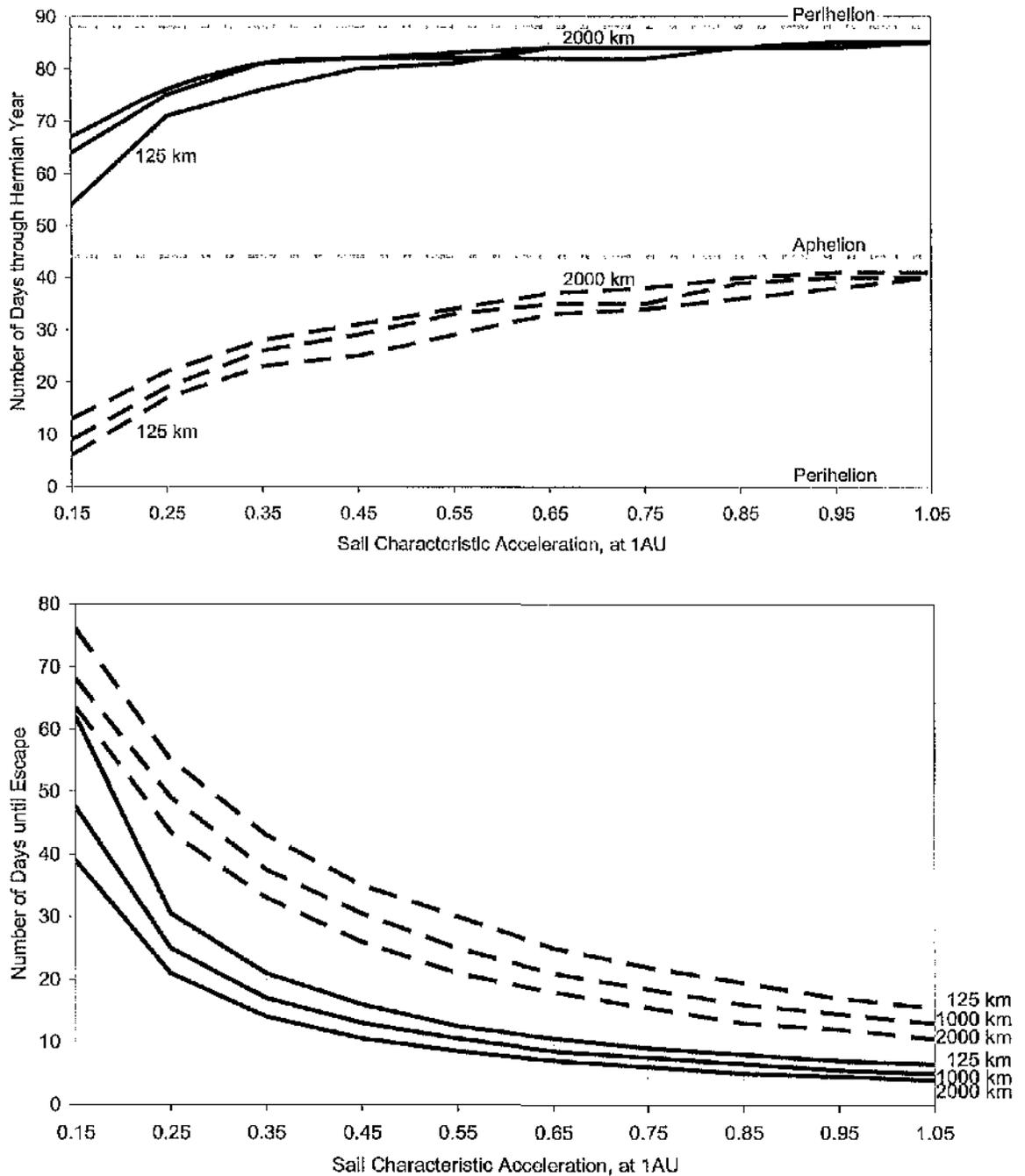
Figure 46 Launch date of maximum (– –) & minimum (—) escape time trajectories, through Hermian year, top. Time until escape of maximum & minimum trajectories, bottom. Initial altitudes of 125 km; 1000 km; 2000 km.

Finally, the duration between minimum and maximum escape times was examined and is shown in Figure 47. The difference between maximum and minimum decreases as the sail characteristic acceleration is increased. However, it is noted that even at low sail characteristic accelerations the difference is less than 44 days. Moreover, if the shortest escape trajectory is longer than 44 days then the difference between maximum and minimum tends to be much smaller. That is, a maximum difference can be identified at the

point when the minimum escape duration is just below one half an orbit revolution of Mercury about the Sun and tends to be equal to approximately 24 days. Thus, while the variation in escape time is substantial throughout the Mercury orbit, the optimal start epoch for an escape trajectory to minimise Julian Day at point of escape is independent of launch date.

4.4   Near-Optimal Earth Escape with Minimum Altitude Constraint

Using only the semi-major axis control law it is found that the orbit eccentricity tends to rapidly increase during the final few orbits prior to escape, resulting in a corresponding rapid decrease in perigee altitude. Due to the nature of solar sail propulsion it is possible to gain energy for only half an orbit, as the sail travels away from the Sun. Thus, if the sail is slightly below the energy required to escape at the end of this half orbit, the sail requires another pass of the planet prior to escape. If the radius of perigee is less than the radius of the planet this will result in a collision with the planet, as has been seen in previous work by the candidate (and others), where a negative altitude was noted prior to escape.[163] It follows that the locally optimal strategy used should be altered such that negative altitudes no longer become possible; in effect one wishes to set a minimum radius of perigee. Thus, Earth escape using the semi-major axis control law (Section 2.3) blended with the radius of pericentre control law (Section 2.8) is considered. It is possible to obtain planetary escape through use of only the pericentre controller, thus assuring a positive altitude is maintained throughout. This however would result in a greatly increased escape time due to the inefficiency of the pericentre controller in gaining orbit energy. The most advantageous strategy is to use the semi-major axis control law to gain orbit energy whenever possible and to use the pericentre control law only when it is absolutely required.
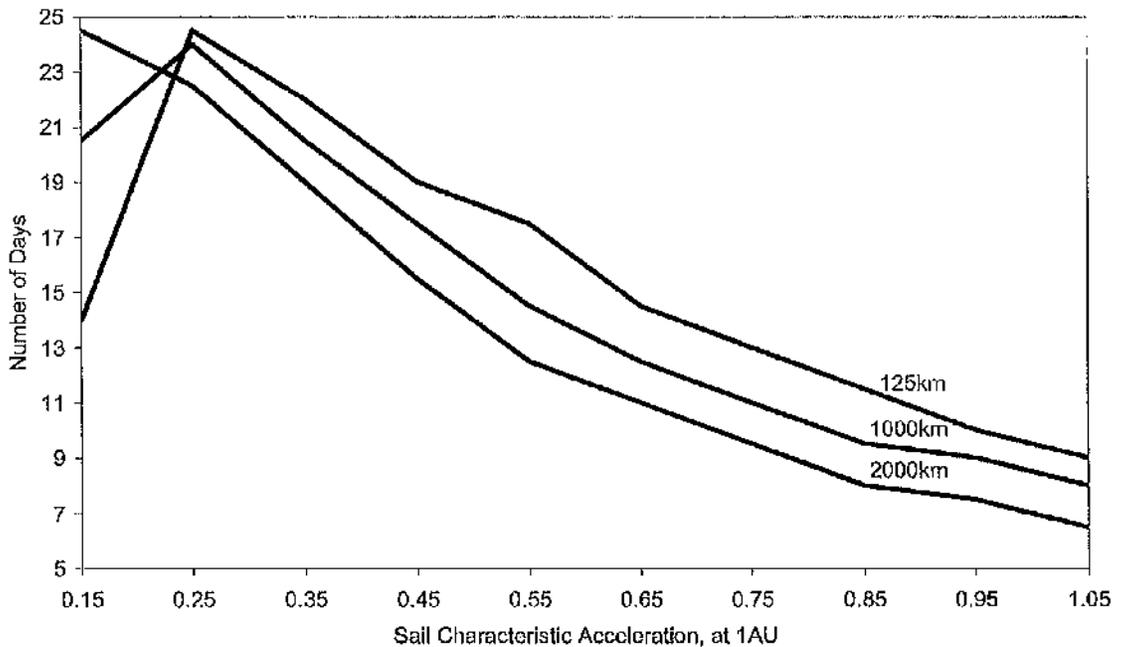
Figure 47 Difference between minimum & maximum duration escape against sail characteristic acceleration. Altitudes are 125 km; 1000 km; 2000 km.

Recall from Section 2 that the approach adopted within this dissertation is that the weight functions should be independent of time, thus using the osculating orbit elements to set the weight functions of each control law prior to blending. Accordingly, Equation 156 to Equation 158 show the weight of each control law as a function of the radius of pericentre only. As the pericentre drops towards undesirable values the pericentre controller becomes more prominent than the energy-gain controller and as pericentre increases the energy gain controller becomes more prominent. Furthermore, in order to ensure a rapid changeover between the control laws a set of exponential weight functions are employed. A rapid changeover between controllers is desirable in this scenario due to the rapid nature of pericentre decrease towards the end of the escape trajectory. In a more general sense a difficulty encountered in generating transfer trajectories with more than one control law is that the controller can become stuck in a dead-band region, where it is caught between the selection of each control law and the orbit elements alter very little. This characteristic however can be turned into an advantage, where blended control laws have been used to generate station-keeping algorithms for potential future solar sail missions, such as GeoSail and a Mercury Sun-Synchronous Orbiter as will be discussed in Section 5.

$$W_a = \exp\left(\frac{W_f}{10}\right)$$

Equation 156

$$W_{r_p} = \frac{2500}{\exp(W_f)}$$

Equation 157

$$W_f = \left(\frac{r_p}{10^6} - 2.5\right)$$

Equation 158

The weight functions defined in Equation 156 to Equation 158 were found using engineering judgement, experience of the system and some trial and error. Note that the units used Equation 156 to Equation 158 are metres. The ratio of W_{r_p} / W_a against instantaneous altitude of perigee is illustrated in Figure 48 where it is seen that large values of r_p result in the semi-major axis control law dominating the blending process. Similarly, small values of r_p result in the pericentre control law dominating the blending process.

4.4.1   Initial Orbit Selection

Low cost launch options are somewhat limited and tend to place the spacecraft into a prohibitive orbit for solar sail performance, due to Earth's steep gravity-well, short orbit periods that require rapid slew manoeuvres and the residual upper atmosphere. However, for completeness it is necessary to consider non-optimal initial orbits due to parallel applications in orbit about other planetary bodies, such as a Mercury sample return mission.[132 - 135] Earth escape from high energy orbits is however of practical interest, for example a piggyback launch opportunity to a 72-hr Earth orbit with a future science mission, similar to the INTEGRAL[190] spacecraft, with a perigee altitude of 10 000 km well above the upper atmosphere would provide an attractive initial orbit for solar sail operations. It was found however that even such high-energy
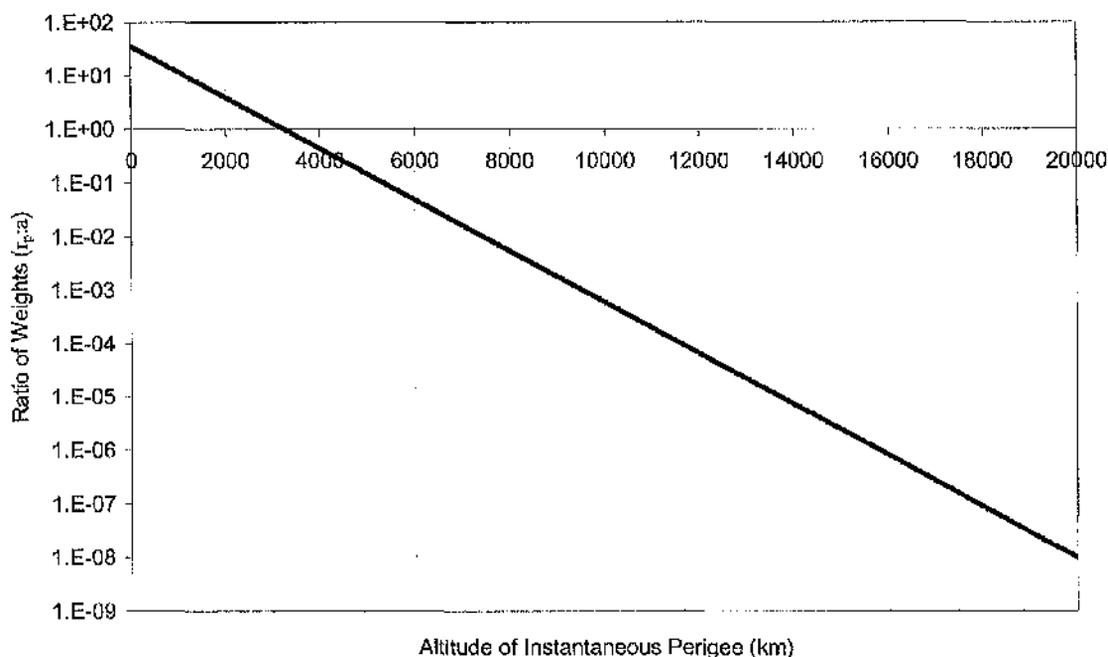
Figure 48 Ratio of W_{r_p} : W_a against instantaneous altitude of perigee.

orbits can have a rapid reduction in perigee and could still traverse the upper atmosphere. Four potential options are presented as initial orbits for solar sail Earth escape; the first is from GTO, the second is from GEO, the third a 1000 km altitude polar orbit and the fourth is a 72-hr Earth orbit with a perigee altitude of 10 000 km. The four potential orbits cover a wide range of initial conditions and thus provide a comprehensive test case for the control algorithms and weight functions presented.

4.4.2   Escape from GTO

GTO has been identified by many studies as a potential starting orbit for solar sail missions, particularly by the DLR ODISSEE concept.[99] GTO is taken to be similar to the original ASAP-5 delivery orbit, giving a perigee altitude of 560 km.[88] Note that Ariane 5 launches to GTO have in actual fact used a range of perigee altitudes from 241 km up to 1167 km, with the perigee altitude typically set at over 600 km in recent GTO launches, for example the SMART-1 perigee was 667 km altitude. At 560 km altitude the solar sail will experience air drag and aerodynamic torque; as such the blended sail control law is altered so that when the sail altitude is below 1000 km the sail is continually slewed to maintain a minimum drag, edge on, profile to the atmosphere. This minimum profile approach will have the additional benefit of significantly reducing gravity gradient effects across the sail surface, which will aid attitude control system design. The sail moves through the atmosphere with negligible aerodynamic loading on the sail structure and allows GTO to be considered as a realistic initial orbit, although it should be noted that sail slew rates are required to be high due to the short orbit period. This addition to the sail control strategy is adopted only for GTO escape trajectories; however it would be valid for any high eccentricity orbit with pericentre inside the upper regions of the planetary atmosphere, such as a Molniya orbit. Furthermore, with this modification to the control strategy the exclusion of atmospheric effects from Model 3 and Model 6 becomes a reasonable

simplification. A midnight launch places the payload on an orbit with a Sun-pointing apogee. The midnight option is considered along with the non-standard midday launch, which places the payload on an orbit with a Sun-pointing perigee and a prolonged shadow event at each orbit apogee. An ASAP launch would be an auxiliary payload and the GTO orbit alignment, along with perigee altitude, would be defined by the primary payload requirements. It was found that standard midnight launches result in the sail striking the Earth before escape for all sail accelerations above 0.3 mm s^{-2} when the semi-major axis controller is used exclusively. This is shown at the left hand side of Figure 49 where the escape duration drops to zero days, indicating an Earth collision; this convention is held for all trajectories within the remainder of Section 4. The midday launches however do not repeat this when using the single controller. Instead it is found that Earth collision occurs only for high sail characteristic accelerations, once the total number of orbit revolutions prior to escape is small. From Figure 49 it is seen that when the blended controller, described above and by Equation 156 to Equation 158 is used the trajectory no longer strikes the planet for either the midday or midnight launch options. Comparison of the blended control law with the energy gain control law shows only a small increase in escape time for midday launches as a result of raising and then maintaining perigee altitude above the upper atmosphere. Thus, one concludes that the inclusion of the additional steps taken to reduce aerodynamic loads on the sail have had negligible impact on sail escape performance, yet has a potentially significant impact on reducing sail loads. Model 6 was used to produce Figure 49. Note that due to the method of sail performance input to the model, the sail acceleration denoted in Figure 49 is related to the sail characteristic acceleration by an efficiency factor, η, of 0.908156, as defined by Table 1.
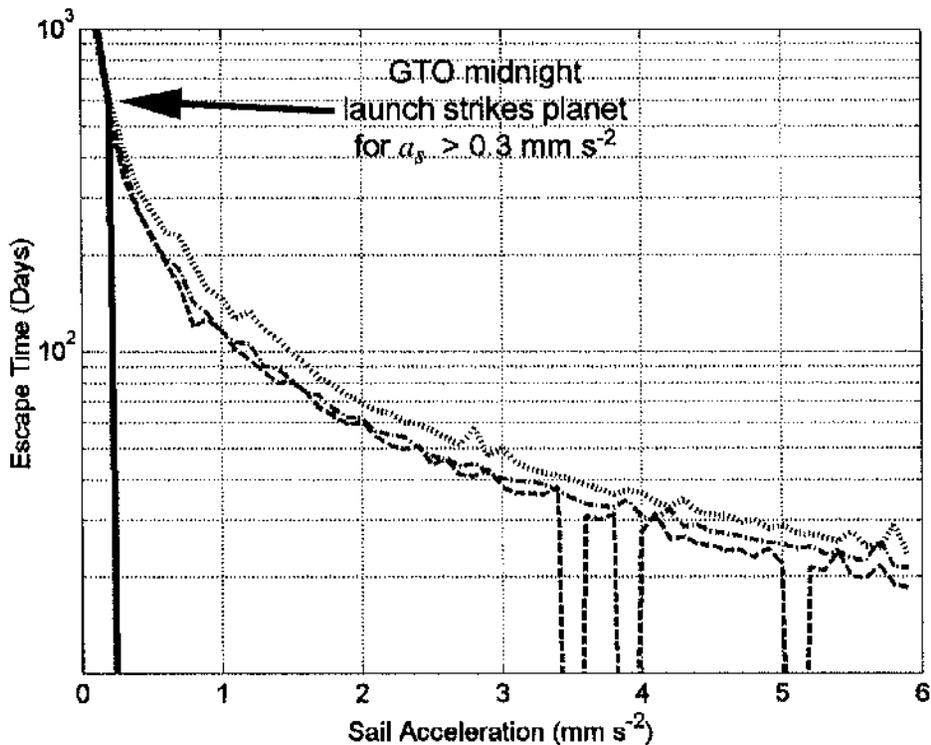
Figure 49 GTO escape times for midnight (—) and midday (--) launch using $^{da}/_{dt}$ control only and blending controllers for midnight (···) and midday (— · —) launch.

Selecting the specific case of a sail acceleration of 2.0 mm s^{-2} (characteristic acceleration 1.816 mm s^{-2}) one can examine in detail the behaviour of the control laws and the effect of the weight functions, using Model 6. The result of the controller is visualised in Figure 50 and Figure 51, where it is seen that the single controller allows the orbit to enter the forbidden region on the semi-latus rectum – eccentricity plot, corresponding to a minimum altitude of 200 km. The blended controller realises it is approaching this region and acts to avoid it, hence safely navigating the solar sail towards the target region and Earth escape. Note from Figure 51 the weight of pericentre steadily decreases through the trajectory to day number 20 when it begins to rise again to counter the reduction in perigee radius.

4.4.3   Escape from GEO

GEO represents perhaps the most attractive initial orbit, with a large orbit radius well outside the steep gravity well and air drag associated with LEO. Furthermore it is attainable at relatively low cost as a Delta IV auxiliary payload on the Secondary Attach Mounting (SAM).[191 - 193]

It has been assumed previously that the issues of air drag and aerodynamic torque on a solar sail need not be considered for an escape spiral beginning at GEO.[191] However, it has been found that this assumption breaks down for the locally optimal energy gain control law at high sail characteristic accelerations when the number of orbits until escape is low, causing a rapid variation in eccentricity and hence pericentre altitude during the short escape spiral. It is thus required that one use the blended control law. It is also found that the low perigee passages occurs just prior to a reduction in the number of orbits required for escape, as seen in Figure 52. Model 6 was used to produce Figure 52 which shows the required time until escape from GEO against sail acceleration with the typical exponential drop-off rate clearly visible, corresponding well with previous work.[35] Similar to Figure 49, the sail acceleration denoted in Figure 52 is related to the sail characteristic acceleration by an efficiency factor, η, of 0.908156.

It is noted in Figure 52 that the exponential drop-off in the required time until escape from GEO is mixed with a short period oscillation, seen as maximum and minimum within the exponential curve. Each maximum corresponds to a reduction by one in the number of orbits required to reach escape energy. This jagged curve is a unique characteristic of solar sail propulsion caused by the inability of a solar sail to gain orbit energy while travelling towards the Sun, as discussed in previous sections. In reality the exact locations of these spikes in escape time would be difficult to predict and hence take advantage of, or conversely insure against encountering, due to trajectory model uncertainties, calculation errors and launch date uncertainties. Thus, such maximum and minimum would make advanced mission planning awkward, as the exact escape epoch would be difficult to predict. Therefore, ensuring escape for an optimal planetary transfer trajectory would be problematic and hence require a margin in the planetary escape phase of the mission.

It is seen in Figure 52 that when the sail acceleration is low both controllers provide almost identical results, and up to an acceleration of 3 mm s^{-2} the escape times are similar. At no time does the blended controller allow the sail to pass below the 1000 km altitude limit.
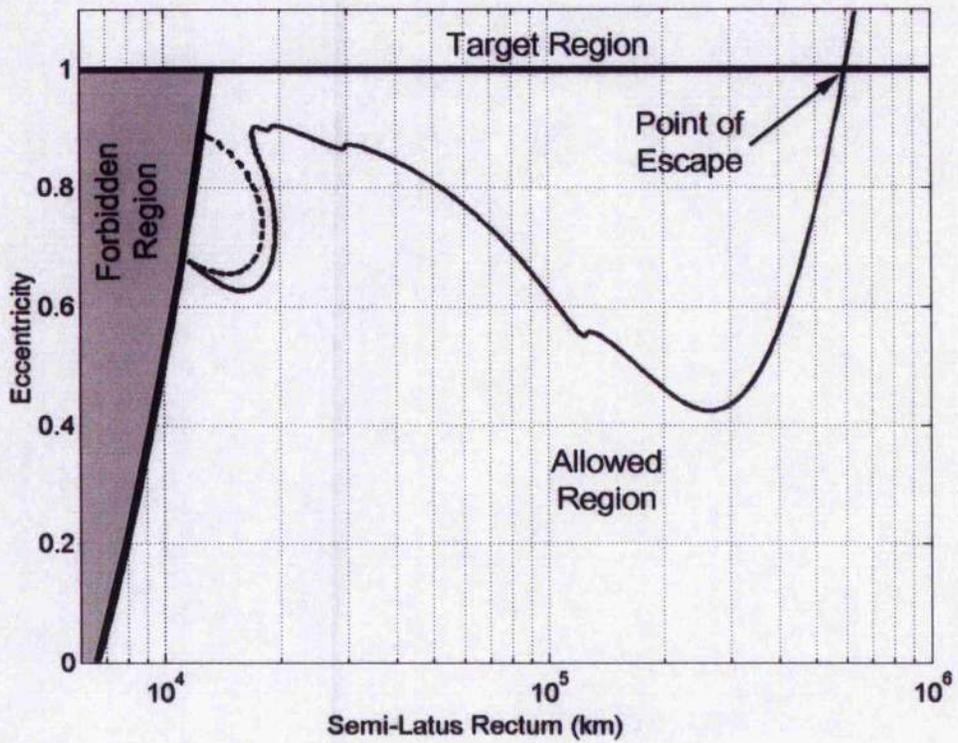
Figure 50 Visualisation of single (– –) and blended (——) control laws for escape spirals from midnight GTO.
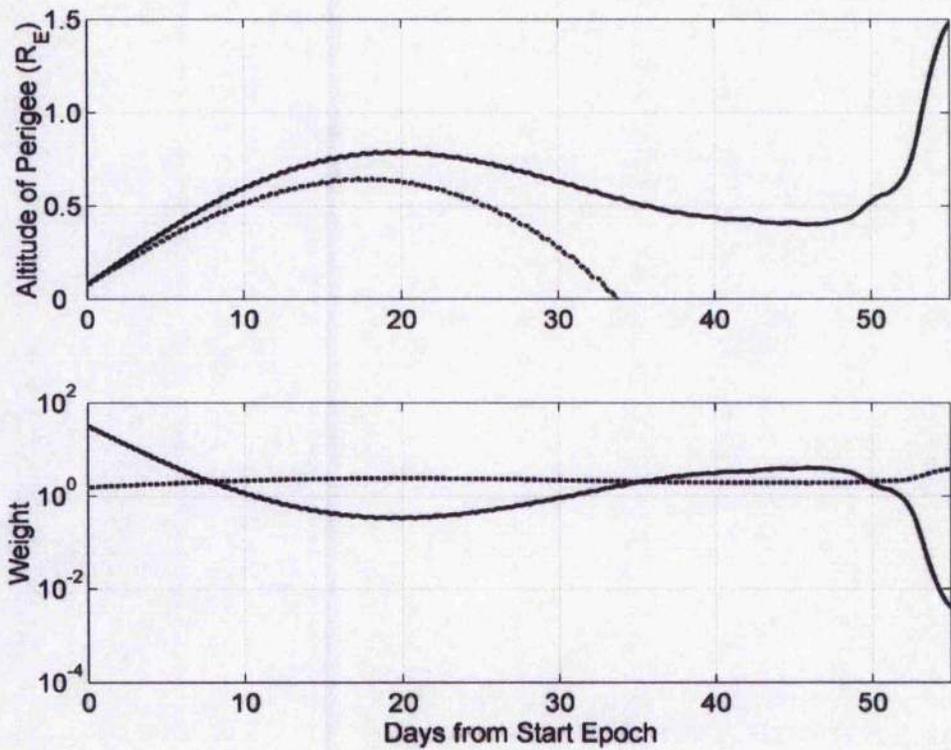
Figure 51 Visualisation of single (--) and blended (——) control laws for escape spirals from midnight GTO.
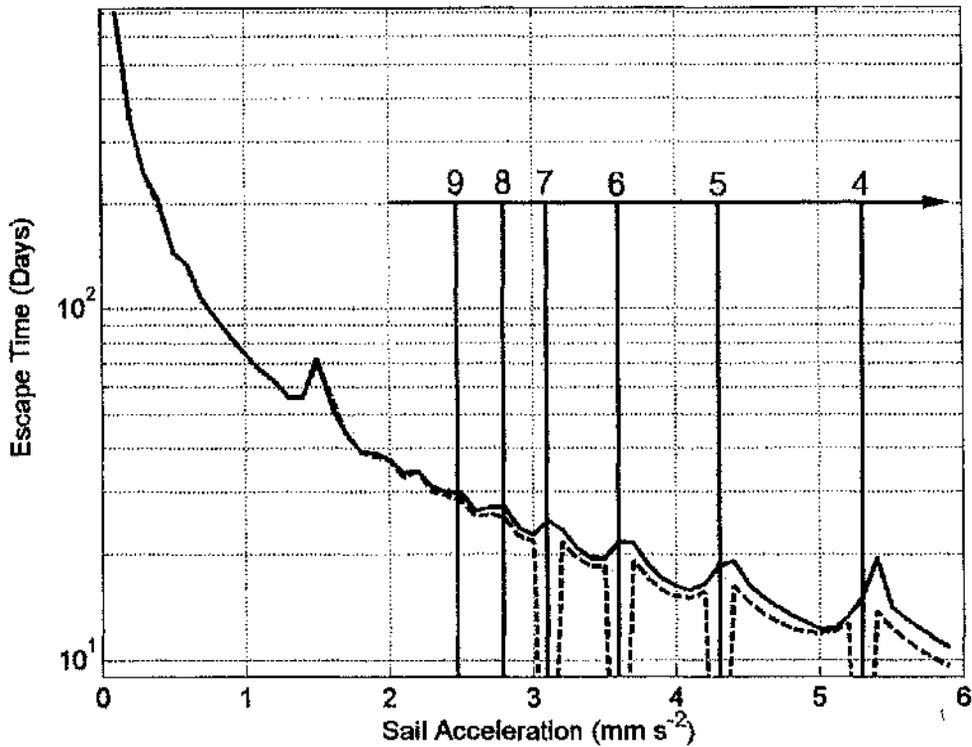
Figure 52 Escape time from GEO using $^{da}/_{dt}$ control only (--) and blended control (—). Orbit revolutions prior to escape indicated.

Figure 53 shows the escape time using the blended sail control in Model 1 and Model 6 where it is seen that the blended sail control law is able to adjust the sail orientation with respect to time, thus correcting for a different set of perturbations from those within the original design scenario. Note that within Figure 53 the sail acceleration equals the sail characteristic acceleration for Model 1, while the efficiency factor of 0.908156 must be used to find the characteristic acceleration for Model 6 due to the non-ideal sail model definition used within this section. The ability of the control system to adapt in the presence, or absence, of perturbations originally considered is due to the nature of the individual control laws, where the desired sail orientation is defined by the current orbital elements and not by a stored data file as would be required if attempting to follow a true-optimal trajectory. Thus, if the sail is not where it was originally predicted to be, then the on-board system automatically adjusts, correcting for the unforeseen perturbation while maintaining the near-optimal nature of the original trajectory. This self-correcting feature of the control system offers the potential to reduce the required uplink telemetry, as only the current sail state vectors are required, rather than an entire new set of control angles.

4.4.4   Escape from 1000 km Polar Orbit

A high polar orbit within the LEO environment can be achieved as a dedicated low cost launch, for example through use of a Dnepr launcher,[194] or the new Arianespace Vega launcher.[195] Several advantages have been identified which could make this an attractive option for future sail missions. However, as many problems as benefits would exist, as such escape from a 1000 km polar orbit would be a significant engineering

challenge.[191] The potential parallel applications at other planets however require that the control system be able to safely guide the sail to escape from such a low energy orbit.[132 - 135] Note that the control strategy documented within this section was successfully adapted, by re-tuning Equation 156 to Equation 158, in References 133 to 135 for Mercury capture and escape trajectories as part of a sample return mission study. Escape times from a 1000 km polar orbit are shown in Figure 54, where the initial orbit normal is aligned with the Earth-Sun line and calculation start epoch set at Vernal Equinox. Model 6 was used to produce Figure 54, with the sail characteristic acceleration related to the sail acceleration by $\eta = 0.908156$. It is seen that the semi-major axis controller causes the sail trajectory to intersect the Earth for most sail accelerations in the range 1 mm s^{-2} to 2 mm s^{-2}. The blended control system is able to steer the sail to escape without Earth collision. Additionally, it is found that the escape times are within 5 % of the semi-major axis control times, except close to the region when this controller breaks down and safe escape times tend towards 10 % longer in duration. The weights given in Equation 156 to Equation 158 were once again used for this escape time scan; however an additional condition was added such that the semi-major axis controller was used exclusively if eccentricity was less than 0.07 and perigee altitude was greater than 500 km. This additional condition was found to improve optimality which was compromised due to the low initial orbit energy both at Earth and at Mercury. Note in Figure 54 the escape time for an acceleration of 0.1 mm s^{-2} was found to exceed five years, thus maintaining calculation accuracy made calculation of escape time prohibitive.
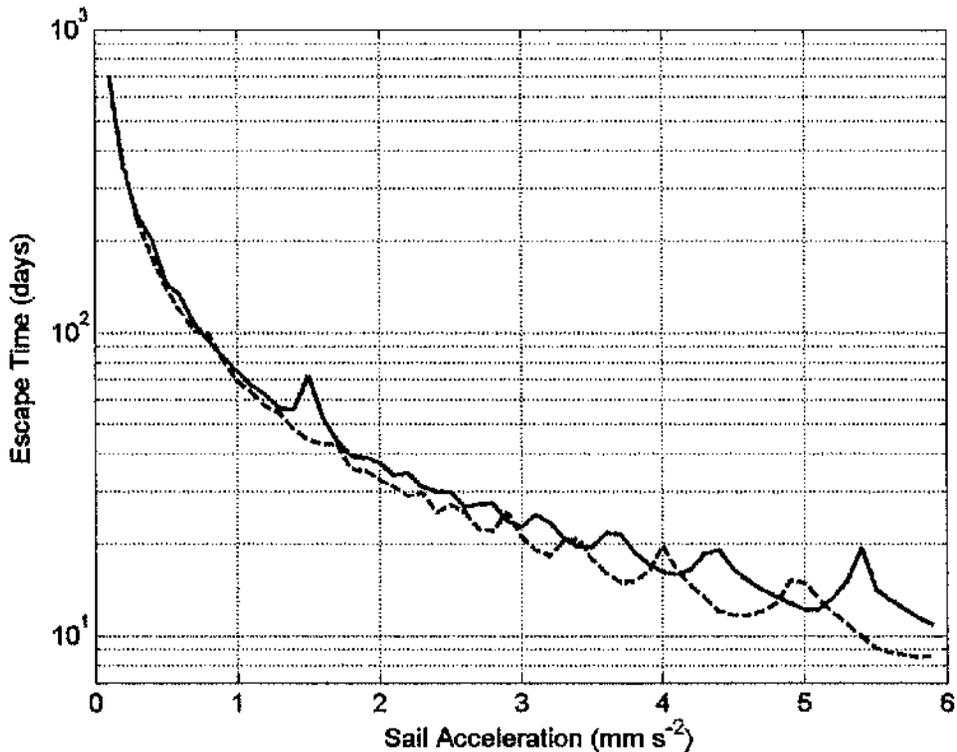
Figure 53   GEO escape using blended sail control in Model 1 (--) and Model 6 (—).
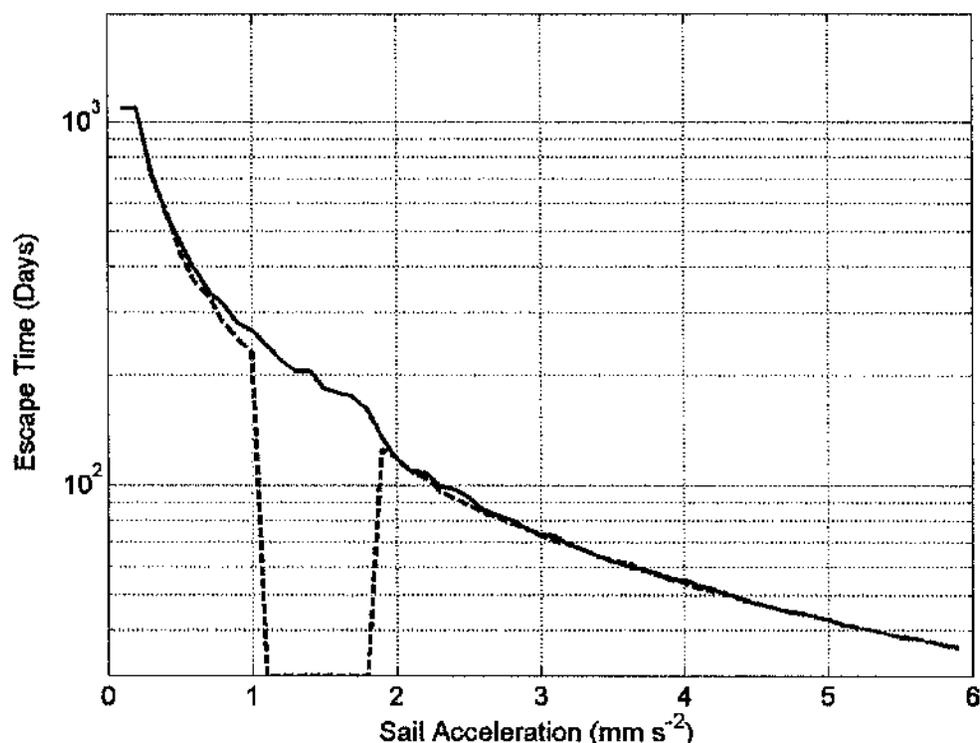
Figure 54   Escape times from a 1000 km polar orbit, for $^{da}/_{dt}$ (--) and blended (—) controllers.

4.4.5   Escape from 72-hr Earth Orbit

Earth escape using solar sail propulsion is a significant technological challenge; however escape from a large Earth orbit is much less challenging due to the long orbit period and the proximity to escape energy provided by the launch vehicle. This scenario is probably the only realistic near to mid-term option for an Earth escape trajectory using solar sail propulsion. To investigate this type of escape trajectory the INTEGRAL launch orbit is selected, defined as 10 000km × 153000 km × 51.6 deg.[190]

On investigation of solar sail escape times from the INTEGRAL orbit using only the semi-major axis control law it is found that at no time did the trajectory intersect the planet. However, as indicated by Figure 55 it is seen that there are regions where the trajectory comes close to the upper atmosphere. This is shown in Figure 55 where one sees the gap in escape times between the semi-major axis control and blended control. Model 6 was used to produce Figure 55, with the sail characteristic acceleration related to the sail acceleration by η = 0.908156. The blended sail control is normally within 5 % of the semi-major axis control, except in the range of sail accelerations between 2.9 mm s^{-2} and 3.8 mm s^{-2} where relative escape times rises to as much as 35 %. This range of apparently poor escape trajectories corresponds to the number of complete orbits until escape dropping from two to one and the associated rapid increase in orbit eccentricity which causes the sail to pass close to, or through, the upper atmosphere.
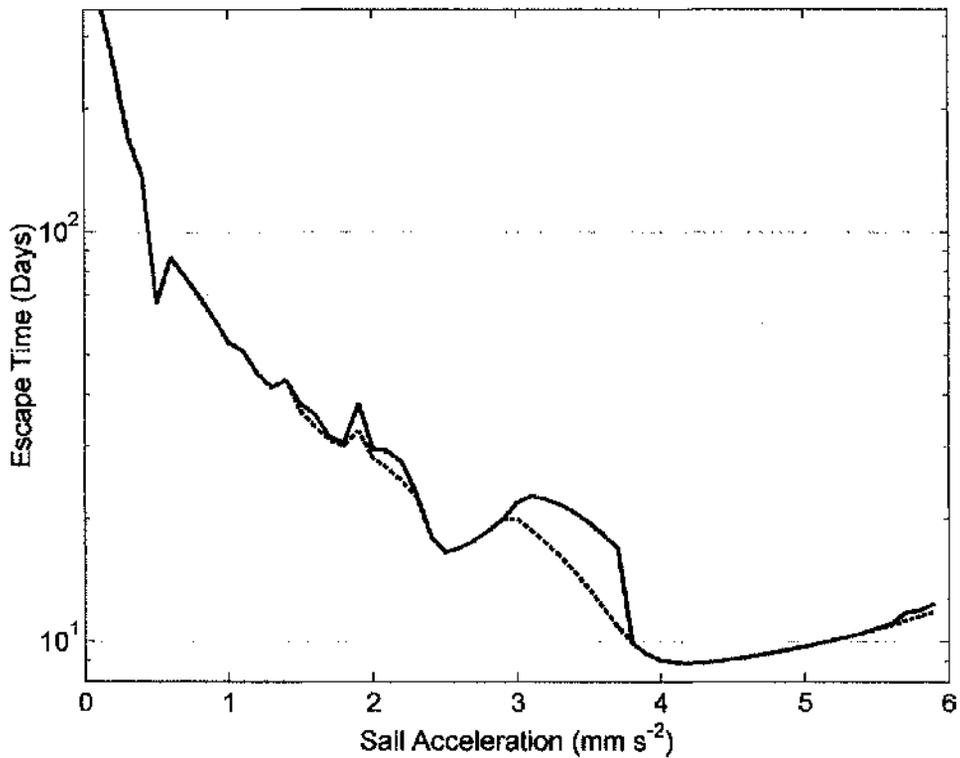
Figure 55 Escape times from a 72-hr Earth orbit with a perigee altitude of 10 000 km, for $^{da}/_{dt}$ (--) and blended (—) controllers.

4.5   Earth Escape without Shadow

As a solar sail passes through a planet's shadow cone the solar flux over the sail surface drops to zero; as does the thrust. Hence a secondary attitude control system may be required during shadow passage. This secondary system could take any of the standard forms,[53, 187] 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, thus stressing the sail and 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. Such a scenario would potentially enable a reduction in sail assembly loading and a corresponding increase in sail acceleration or payload capability.

Using the blended control algorithms outlined in Equation 156 to Equation 158 and Model 1, the required sail characteristic acceleration for escape from an Earth polar orbit at a range of altitudes was found, as shown in Figure 56. The initial orbit is defined similar to that used in Section 4.4.4 such that the initial orbit normal is aligned with the Earth-Sun line and calculation start epoch is set at the Vernal Equinox. The orbit model utilised only considers perturbations due to the sail thrust; 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. Furthermore, it has been shown in Section 4.4.3 (Figure 53) that the control system can correct for perturbations not included in the

original design strategy. The sail characteristic acceleration at each altitude was incremented in steps of 0.01 mm s^{-2} until an escape trajectory was achieved without any shadow events, the initial altitude was then increased by 50 km and the process repeated. In Section 2.6 (page 52) a locally optimal control law for the variation of the ascending node angle was presented. It was found however that introducing this into the blending equations produced an unnecessary complication within the control system, producing much slower escape times due to the tendency of this controller to cancel out any energy gain over the orbit period. Hence, only the locally optimal control laws used in Section 4.4 and blended through Equation 156 to Equation 158 are utilised in this section.

The required sail characteristic acceleration for a range of initial altitudes from 800 km to 25 000 km is shown in Figure 56. It is seen that an exponential increase in sail characteristic acceleration is required 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 performance or the exponential reduction in escape time for a given altitude as sail characteristic acceleration is increased, as seen in several figures within Section 4.4. The corresponding escape time for a shadow free trajectory from each altitude is shown in Figure 57. As would be expected from the exponential curve of required sail characteristic acceleration in Figure 56, the minimum sail performance escape time for shadow free trajectories is essentially independent of initial altitude; as the required sail characteristic 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. In Reference 35 a single shadow-free Earth escape trajectory is produced using a locally optimal radius of apocentre control law. The initial altitude of this trajectory was 20 000 km, for an ideal sail with no orbit perturbations and a characteristic acceleration of 0.85 mm s^{-2}. From Figure 56 it is seen that this point is above the presented curve and hence the two results correspond well. Furthermore, the trajectory presented in Reference 35 has an escape time of 146 days, which corresponds with Figure 57.

In order to quantify the true effect of neglecting all orbit perturbations other than sail thrust, the single case of escape from 20 000 km using Model 6 was investigated. Figure 56 indicates the ideal sail characteristic acceleration required is approximately 0.8 mm s^{-2}; however accounting for orbit perturbations and introducing a non-ideal sail model the input sail acceleration is increased to 0.85 mm s^{-2}, giving an actual sail characteristic acceleration of 0.772 mm s^{-2}. The escape trajectory is thus calculated for these initial conditions as seen in Figure 58, where the escape trajectory is viewed in a fixed Sun-axis reference frame looking from the Sun towards the Earth. It is seen from Figure 58 that at no time does the trajectory pass behind the Earth and hence no terrestrial shadow events are recorded. This result is verified by analytical analysis of Earth, Sun and spacecraft position vectors and including a 2 % addition to the Earth's radius; accounting for the increase in shadow size due to the atmosphere as discussed in Section 3.4.2. Figure 59 shows the orbit inclination and ascending node angles. It is seen that the ascending node angle initially increases slowly for the first 100-days, before then rapidly increasing for the final 40-days prior to escape on day 141. Note that the minimum altitude of this trajectory is 2397.2 km, on the 116th day of the trajectory.
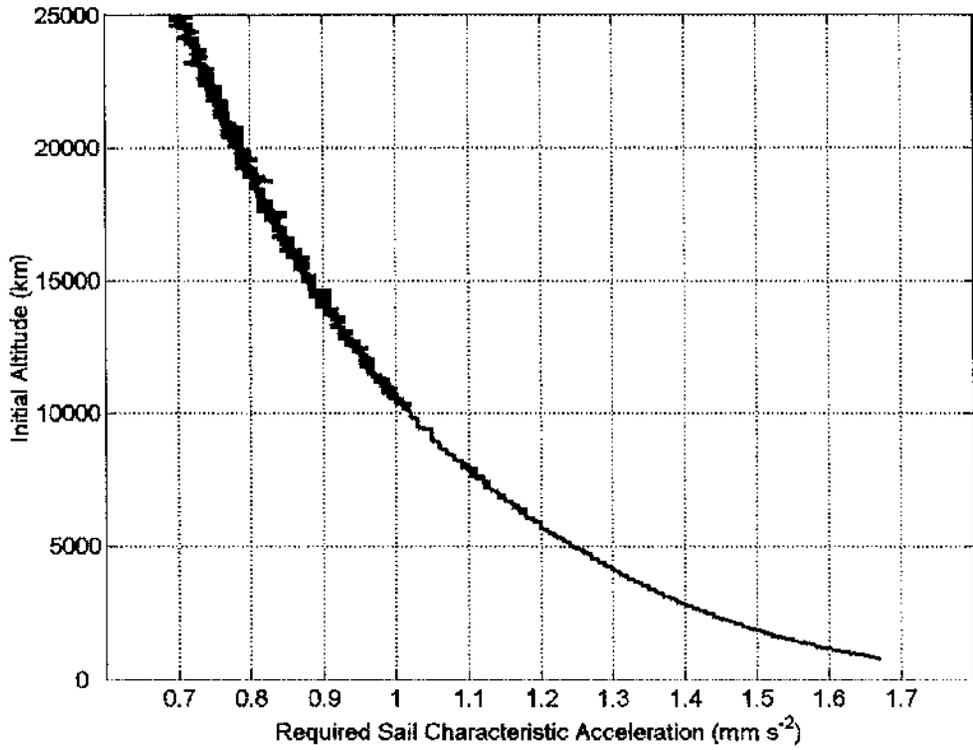
Figure 56 Minimum required sail characteristic acceleration for shadow free Earth escape from a polar orbit.
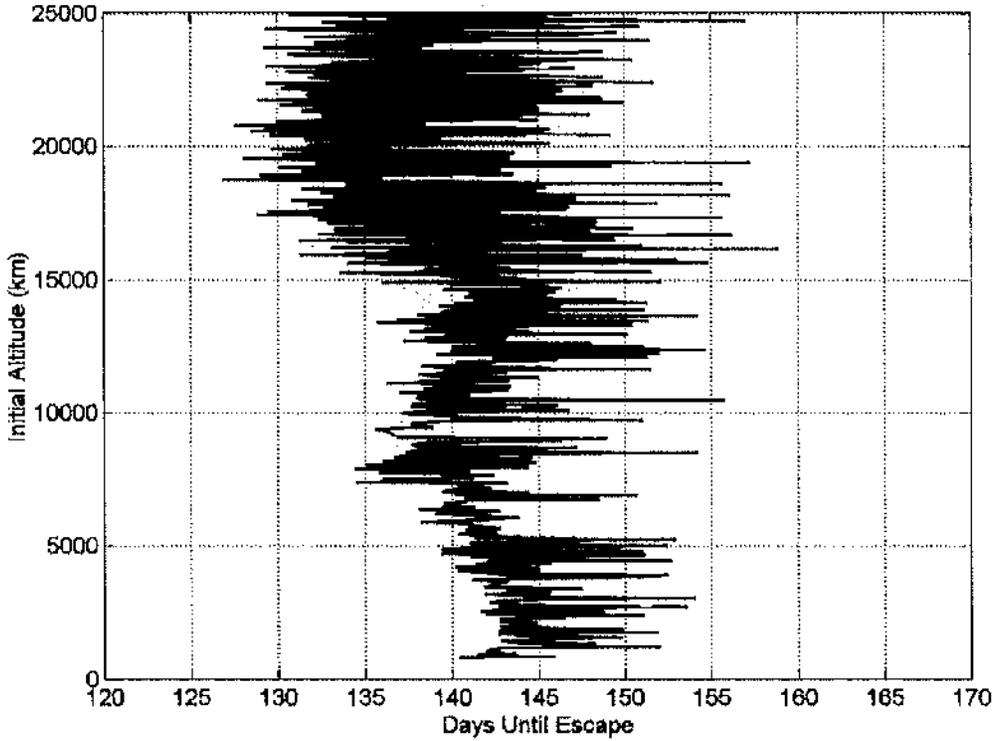
Figure 57 Escape time for each trajectory data point in Figure 56.
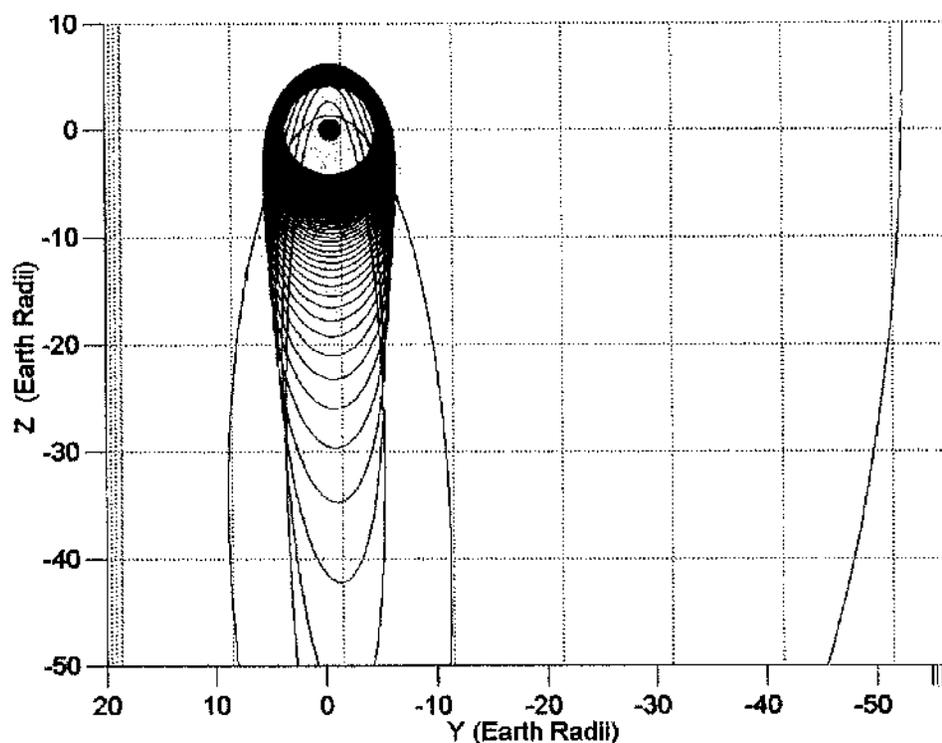
Figure 58 Shadow free escape from 20 000 km altitude seen from a fixed Sun-line coordinate system, with sail characteristic acceleration of 0.772 mm s^{-2}.
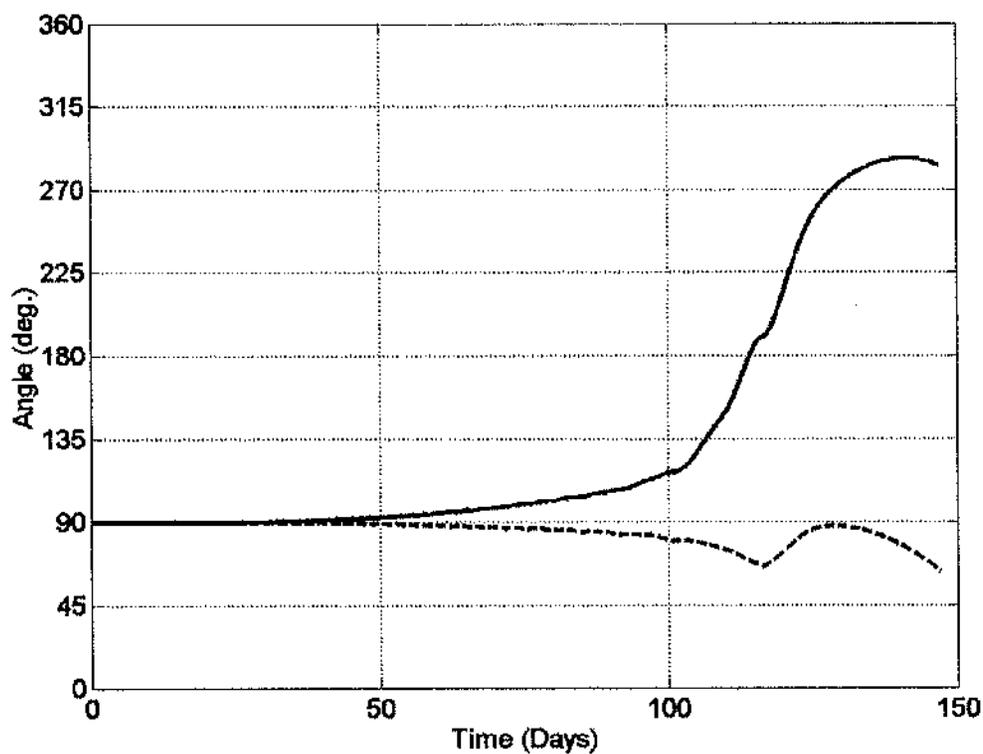
Figure 59 Inclination (--) and ascending node (—) angles for the trajectory in Figure 58.

5 Planet-Centred Orbit Transfers Using a Solar Sail and Locally Optimal Control Laws

The use of a solar sail in planetary escape trajectories was thoroughly investigated in Section 4, where a potential solution to the problem of maintaining a minimum altitude during Earth escape was demonstrated. This section follows on by considering the use of a solar sail for planet-centred transfers and station keeping. In Section 4 the weight functions were defined by a series of simple exponential relationships. However, in attempting to create a more complex and robust control system for orbit transfers and station keeping applications it is required to evolve the derivation and calculation of the weight functions. The use of locally optimal control laws for solar sailing is thus further developed through the formation of a new control method which calculates the weights to be used in the blending of the individual control laws to provide near-optimal results. The new control method will then be demonstrated for use in a complex orbit transfer at Mercury where the main constraints are thermal and not transfer duration, which can be considered secondary in many solar sail applications at Mercury as will be discussed next.[133 - 135] The control method will also be used for station keeping and investigation of how such a method could enhance the science return of previously proposed planetary solar sail missions.

5.1 Planet-Centred Orbit Transfers Using a Solar Sail

The generation of solar sail planet-centred orbit transfers are perhaps of limited purpose. For instance, lengthy Earth escape spirals at the beginning of a heliocentric mission should be eliminated by the launch vehicle whenever possible. If however the sail is the primary form of propulsion then during a reconnaissance or sample and return mission the sail must be utilised to correctly deliver the science payload to the target orbit about the destination body. Much prior work that has claimed to generate capture trajectories has simply been escape trajectories integrated over a negative time-span.[130, 133 - 135] Such an approach does however provide a suitable approximation to obtain the required timescale of any capture trajectory and is thus suitable for early mission analysis studies.

Only a few solar sail planet-centred orbit transfers have been previously published due primarily to the significant difficulty of generating trajectories which are optimised over numerous revolutions. A rudimentary transfer from GTO to the original GeoSail mission orbit of 10 × 30 Earth radii was generated by the candidate using locally optimal control laws.[163] The transfer trajectory in Reference 163 is not included as part of this dissertation. A recent doctoral thesis used extremal steering strategies for simulation and optimisation of Earth – Moon transfer trajectories using solar sailing.[36] This work resulted in the solution of a weak stability boundary problem and generated realistic transfer trajectories from GTO to a bound lunar orbit. One of the few other published planet-centred solar sail transfer trajectories; other than lunar flybys which are not actual orbit-to-orbit transfers, was in a 1977 study.[31, 32] This study developed a numerical algorithm to calculate optimal planet-centred trajectories. However, only one orbit-to-orbit transfer was generated as it was found that the algorithm required a very good initial guess before a solution was found and that eccentricity convergence was difficult when the target eccentricity was low. Perhaps with hindsight

one can now speculate that this is due to the anomaly discussed in Section 2.4 and that a different Earth – Sun – Sail orientation may have provided a solution. Transfer to sub-escape points presented no convergence difficulties; however orbit transfers were much more difficult and the authors were unable to generate more than one complete trajectory within the timeframe of the study.

It is the experience of the candidate that while planet-centred orbit transfers are of only limited purpose, when they are required the primary cost function is seldom time. For example, the primary cost function for sail operations at Mercury is typically thermal and avoidance of passage near the sub-solar point. During a recent Mercury Sample and Return mission study it was found that the surface thermal conditions were such that a lander could only survive at key specific times of the Hermian year.[133 - 135] It was also found that the optimal Earth – Mercury transfer and subsequent capture spiral resulted in arrival of the lander at an inappropriate time for landing. Thus, it is required that the lander wait in Mercury orbit until the surface thermal conditions are suitable. The optimal orbit for the sail to enter while waiting for the surface conditions to become suitable was found to be the Mercury-forced sun-synchronous orbit, hence minimising thermal loads on the sail and its systems.[35, 129 - 131] Thermal requirements thus necessitate an orbit-to-orbit transfer from the Sun-synchronous orbit to the low-circular near-polar orbit for deployment of the lander. The primary cost function of this transfer is not time, as the arrival time is fixed and the transfer is necessitated by a need to wait for the correct surface conditions.

5.2   Accessibility and Deficit Blending

The Accessibility and Deficit (A^nD) blending method seeks to give each individual control law a relative importance prior to defining the final weight functions and thus the blended control vector. The deficit of each element from the final target value is considered. Additionally, the efficiency or accessibility of any attempt to alter an orbital element is considered, thus avoiding inefficient use of the sail, such as in prolonged periods of high pitch.

The Deficit score is found not by consideration of each element's value, but instead by estimation of the time required to attain the target value using the locally optimal control law. By computing λ_σ the locally optimal pitch and clock angle for control law σ can be found. With the locally optimal pitch and clock angles calculated the sail perturbation vector in the Sun-line reference frame can be determined using Equation 121. Thus, using Equation 122 and Equation 123, the sail thrust vector is converted into sail RTN axis in order to calculate $d\sigma/dt$, the rate of change of element σ. Note that to find $d\sigma/dt$ the magnitude of the sail perturbation vector is corrected due to the locally optimal sail pitch angle following Equation 22. With knowledge of the locally optimal rate of change of element σ, the current value of σ and the target value of σ, it becomes a simple matter to estimate the time required to attain the target value assuming a constant rate of change. Repeating this process for each control law being blended allows one to normalise the time required with respect to the largest time. Thus, each control law gains a score between zero and one for the corresponding deficit, with zero meaning the element has attained its target value and one that it is the furthest, or has the greatest deficit, from its target value. Note that this assessment assumes an ideal sail force model. However, in Section 6 this assumption will be removed for heliocentric trajectories by using the same force model to set the deficit score as is used for orbit propagation.

The deficit of an element can appear excessively high if the corresponding locally optimal pitch angle is high, which results in a low rate of change and thus gives a high deficit score to an element which will poorly utilise the sail. Thus, one should also consider the accessibility of control law σ by consideration of the corresponding optimal pitch angle. Recall from Equation 22 that as the pitch angle is increased the corresponding sail acceleration drops off as the cosine squared of the pitch angle. The accessibility score is found by calculating the cosine squared of the optimal pitch angle for each control law and normalising with respect to the largest. Hence, the control law which best utilises the sail thrust gains a score of one, while the most inefficient gains a low, but not necessarily zero, score.

Using the A^nD blending method the weight functions are thus based not only on need but also on an opportunistic level. The deficit score gives the highest score to the element furthest from its target value, while the accessibility score gives priority to the element which most efficiently uses the sail. The combination of the two scores provides the weighted relevance of each control law. However, the method of combining the two scores must itself be carefully considered and rationalised. Multiplying the Accessibility and Deficit scores results in a low total score if either score is low. It has however been found that if an element has low accessibility for a given direction of change, say negative, this is because the orbit alignment about the planet and with respect to the Sun is typically close to optimal for the opposite direction of change, as discussed in Section 2.4. As a result of this, if an element has low accessibility and high deficit then multiplication of the two scores results in an increase in the deficit. It is thus found that even though the accessibility is low one cannot totally ignore the element. Addition of the two scores results in a low accessibility and high deficit scenario receiving a moderate score and was thus found to offer a better solution. The final A^nD score is thus found by addition of the two individual scores.

The final weight functions are found by multiplying an individual element A^nD score by a constant. For example, as will be seen later, the GeoSail mission primary requirement is to rotate the orbit argument of pericentre. Thus an additional importance is placed on this element and it is multiplied by a larger constant. Elements which are not being blended are multiplied by zero to remove them from consideration, while elements of lesser importance are multiplied by smaller constants. The use of constants allows the control system to be fine tuned to increase optimality and essentially reduces the trajectory optimisation problem from finding the cone and clock angle control history to finding a small set of constants. Thus, the optimisation process has only a few data points to determine which then in-turn determine the sail control angle history for the best-case trajectory, rather than the optimiser trying to find several hundreds of data points as traditionally performed in heliocentric trajectories, or thousands of data points as would be required in planet-centred trajectories. The selection of appropriate constants is intuitive and typically follows the mission goals, such as seen for the GeoSail mission. However, engineering judgement (or automation) allows a more rapid convergence towards the most favourable solution. If it is unclear which orbital elements should be focused on, an initial guess can be obtained by utilising the A^nD scores only, prior to then introducing the constants in order to improve optimality.

5.3   Mercury-Forced Sun-Synchronous Orbits

The close proximity of Mercury to the Sun means that even a relatively modest level of sail performance can provide a significant thrust. Mercury has a reciprocal of flattening over eighteen times that of Earth, with a J_2 value of only 60×10^{-6}, although the reciprocal of flattening is often mistakenly quoted as being infinite. Thus, once in orbit about Mercury it is not possible to create a natural sun-synchronous orbit; a supplemental thrust is required to replace the oblateness utilised at Earth. Leipold, et al showed that this supplemental thrust could be provided by a modest solar sail in a highly elliptical polar orbit, consequently allowing the spacecraft to maintain station at or near to the solar terminator of Mercury.[35, 129 - 131] As such the thermal loading due to reflection and re-radiation from Mercury's surface is significantly reduced, while the severe thermal cycling encountered by numerous passes through the shadow cone is also eliminated. It is considered that the optimal remote sensing orbit places the spacecraft at a small offset from the solar terminator, rather than directly overhead as the low Sun angles near the terminator on the dayside of the planet allow greater topographic discrimination in near-constant illumination conditions. Finally, due to Mercury's rotational and orbital period 3:2 resonance, the Sun-synchronous polar orbit allows complete surface coverage in only 88 days. However, due to the optimal orbit offset from the solar terminator it is necessary to remain in orbit for 176 days to acquire full surface visual coverage.

Recreating the trajectories published by Leipold, et al it is found that the orbit is an unstable equilibrium. It has been shown previously that the ascending node angle cannot be varied without also altering the argument of pericentre angle.[196, 197] It was similarly noted by Leipold, et al that the argument of pericentre experienced a long period oscillation due to the sail thrust vector. Over short timescales the small variation in argument of pericentre results in only small variations in semi-major axis and eccentricity. However as the argument of pericentre reaches the peaks and troughs of its long period oscillation, the nominal pericentre altitude of 200 km varies as low as 70 km and rises as high as 400 km, prior to collision with the Hermian surface. Collision typically occurs around 100 to 140 days from the initial start epoch, depending on initial conditions. It thus becomes clear that while the optimal science orbit has a very low pericentre, one requires either an active sail to achieve this, as will be discussed later, or an initially greater altitude.

The allowed Sun-synchronous orbits for a given sail characteristic acceleration can be determined through analysis of the variational equation of motion of the ascending node angle, as given in Equation 68. Following Leipold et al one can integrate Equation 68 over an orbit period, assuming inclination, semi-latus rectum and argument of pericentre are constant over the orbit period.[131] Further, the sail orientation is assumed constant over one orbit revolution and is directed normal to the orbit plane. That is to say, $\alpha = 0$ deg for orbits along the terminator and $\alpha = 10$ deg for orbits offset from the terminator by 10 deg. Changing the integration variable from time to true anomaly allows the derivation of Equation 159.

$$\Delta\Omega\Big|_{\omega+\nu=0}^{\omega+\nu=2\pi} = N\frac{p^2}{\mu\sin i}\int_{\omega+\nu=0}^{2\pi}\frac{\sin(\omega+\nu)}{(1+e\cos\nu)^3}d\nu \qquad\qquad \text{Equation 159}$$

The precession of the node can thus be determined using the orbit period as,

$$\dot{\Omega} = N \frac{\sqrt{a}\left(1-e^2\right)^2}{2\pi\sqrt{\mu}\sin i} \int_{\omega+v=0}^{2\pi} \frac{\sin(\omega+v)}{\left(1+e\cos v\right)^3} dv$$

Equation 160

Extending the analysis performed by Leipold et al in Reference 131 one can generate contours of the potential Sun-synchronous orbits for a given sail acceleration or offset angle from the solar terminator, thus enabling rapid assessment of different potential scenarios. Analysis of Equation 160 with $\omega = {}^{\pi}/_2$ or ${}^{3\pi}/_2$ and setting $i = {}^{\pi}/_2$, such that $\sin i = 1$, allows the derivation of Equation 161 as

$$a = \left[N \frac{\left(1-e^2\right)^2}{2\pi\dot{\Omega}\sqrt{\mu}} \int_0^{2\pi} \frac{\cos v}{\left(1+e\cos v\right)^3} dv \right]^{-2}$$

Equation 161

The required rate of change of ascending node angle for a Sun-synchronous orbit is found to be the specific angular momentum of Mercury with respect to the Sun divided by the square of the Mercury – Sun distance. Thus, the required rate of change of ascending node angle varies from 1.28×10^{-6} rad s^{-1} to 5.57×10^{-7} rad s^{-1} depending on Mercury's position. Note however that as Mercury orbits the Sun the sail acceleration varies as the inverse square of the solar distance, thus the induced rate of change of ascending node will vary similarly, as required. Thus, no alteration in the orbit is required during the Hermian year as was also found by Leipold, et al.[131]

Solving Equation 161 for a given eccentricity allows the corresponding semi-major axis for any given sail acceleration level to be determined. For example, the orbit defined previously by Leipold, et al for utilisation within a future science mission was for a sail with characteristic acceleration 0.25 mm s^{-2} at zero offset from the solar terminator.[129 - 131] Pericentre altitude was defined as 200 km and apocentre altitude is quoted as "*approximately 6350 km*". It is found that the actual idealised value is 6293.63 km for such a scenario. Extending the analysis for a sail characteristic acceleration of 0.25 mm s^{-2} Figure 60 shows a plot of pericentre altitudes versus apocentre altitudes for a range of solar terminator offset angles. Above a 50 deg offset angle from the solar terminator the orbit begins to intersect the shadow cone; recall that for an orbit offset from the solar terminator one requires to maintain the sail force vector normal to the orbit plane and as such the fixed sail pitch angle equals the nominal offset angle. Figure 61 shows a plot of pericentre altitudes versus apocentre altitudes for a range of sail accelerations and orbits with solar terminator offset angles of 0 deg and 10 deg; note that the 0.25mm s^{-2} contours correspond to the 0 deg and 10 deg contours shown in Figure 60. Figure 61 shows that the increased sail characteristic acceleration allows for the apocentre to be lowered for a given pericentre value, while the increase in solar terminator offset angle requires an increase in apocentre altitude for a given pericentre value. Note further that the orbits defined in Figure 60 and Figure 61 correspond to the required actual thrust vector and have not been corrected for an imperfect sail surface.
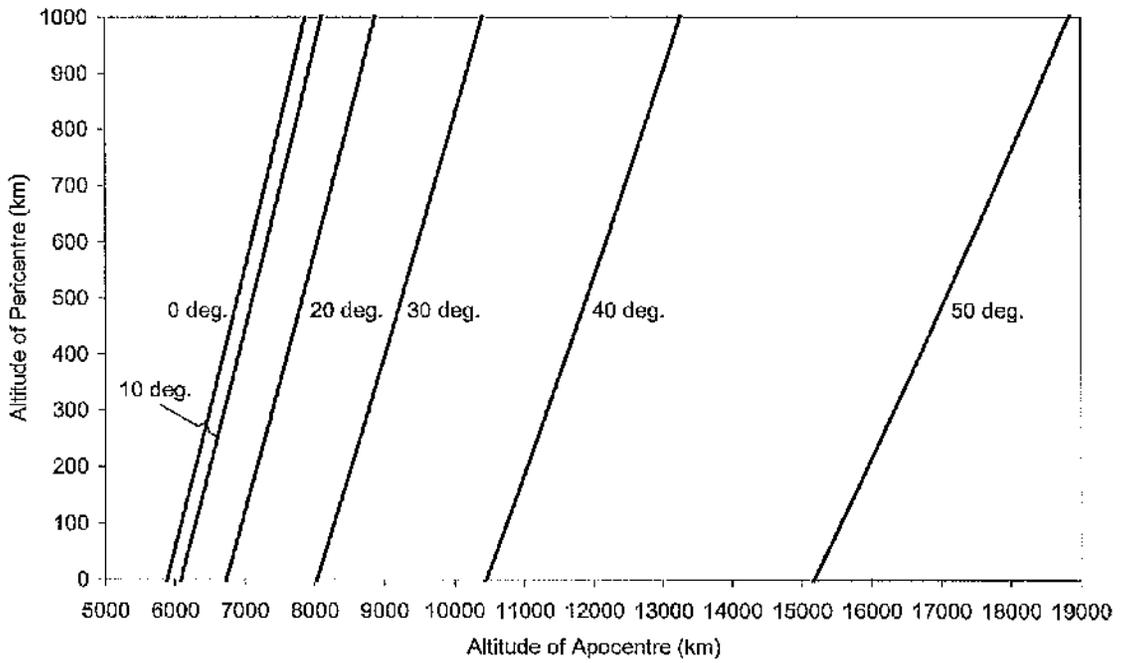
Figure 60 Pericentre versus apocentre contours for $a_{s_c} = 0.25$ mm s^{-2} forced Sun-synchronous orbits, at a range of offset angles from the solar terminator.
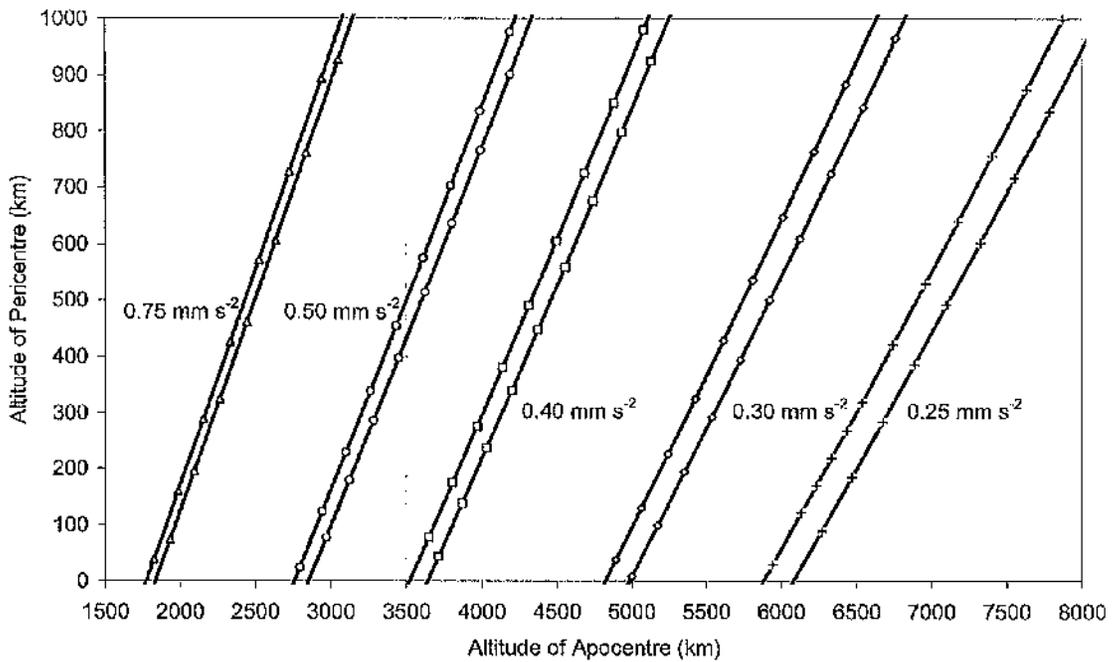
Figure 61 Pericentre versus apocentre contours for orbit planes at 0 deg (left line within each pair) and 10 deg (right line within each pair) offset from solar terminator at a range of sail characteristic accelerations.

For a fixed sail orientation an increased pericentre altitude of 500 km is selected; noting that Leipold, et al first suggested a pericentre altitude of 600 km before later reducing this to 200 km.[129] The sail orbit is displaced from the terminator by 10 deg. From Equation 161, Figure 60 and Figure 61 the required apocentre altitude is found to be 7109.9 km, for a sail characteristic acceleration of 0.25 mm s^{-2}. Note that the sail characteristic acceleration is fixed at 0.25 mm s^{-2} for the remainder of this section when discussing sail operations at Mercury. Utilising Model 3 the trajectory is propagated over a 180 day period for a forced Sun-synchronous orbiter with pericentre over the northern geographical pole. A start epoch at 05 July 2010 gives an initial ascending node angle of 54.5 deg. The ascending node angle varies from this initial value to 360 deg and back up to approximately 54.5 deg after one Hermian year. The rate of change of ascending node varies throughout the Hermian year, as expected, thus maintaining a separation angle from the solar terminator of approximately 10 deg. The orbit inclination remains within +0.45 deg and −0.27 deg of 90 deg, similarly the argument of pericentre varies minimally. It is found that both the pericentre and apocentre altitudes vary considerably from the initial values quoted. Large variations could potentially impact the science goals of the mission and as such active sail control could potentially be used to minimise pericentre altitude variation, while also lowering the pericentre altitude. Such an active sail control system will be discussed in Section 5.3.2. The 180 day propagation described above terminates on 01 January 2011 with the corresponding orbiter position elements defined in Table 6. Note that the final pericentre altitude is 361 km above its nominal value while apocentre is below its nominal value, thus giving a much more circular orbit. Inclination and argument of pericentre are both close to their nominal values, while the ascending node angle places the orbiter ground track approximately 10.3 deg ahead of the solar terminator.

The elements defined in Table 6 will now be used as starting conditions for an orbit transfer to a south-pole pericentre orbit. Such a transfer would potentially enable high-resolution mapping of the entire surface of Mercury with a single spacecraft. Note that Leipold, et al proposed using two spacecraft for such coverage.[130]

Orbit Element	Value	Unit
Semi-Major Axis	6258.2	km
Eccentricity	0.47	–
Altitude of Pericentre	861.2	km
Altitude of Apocentre	6775.8	km
Inclination	90.26	deg
Argument of Pericentre	89.84	deg
Ascending Node	75.11	deg
True Anomaly	228.44	deg

Table 6   Orbit elements after 180 day forced Sun-synchronous orbit.

5.3.1   Transfer between Sun-Synchronous Orbits Using A^nD Blending

A direct transfer from a north-pole pericentre to a south-pole pericentre forced sun-synchronous orbit is relatively simple. The argument of pericentre and ascending node angles must be rotated through 180 deg, while the nominal value of all the other elements remains unaltered. However, a direct and simple transfer would require the orbiter to pass directly over the sub-solar point and through the planetary shadow cone. The transfer orbit is thus optimised such that the orbiter does not pass through the planetary shadow cone at any point. The primary optimisation cost function hence becomes spacecraft thermal constraints rather than a minimum time transfer. Using the orbit elements in Table 6 as the initial conditions for the orbit transfer, the target elements are defined in Table 7. Several potential strategies can be adopted to eliminate planetary shadow from this transfer. The ascending node angle could be rotated very rapidly or the orbit velocity could be reduced to approximately zero and then the direction reversed on a parabolic orbit at a large orbit radius. This second option however would require the orbit energy to approach zero and would require a very high level of navigation accuracy. Alternatively, the adopted strategy was to raise the orbit energy and circularise the orbit, allowing the planet to rotate rapidly beneath the orbiter. The use of a circular orbit also simplifies the transfer, as the argument of pericentre can be selected as eccentricity rises again, rather than actually rotating the orbit through 180 deg. The targeting of a continually varying ascending node adds an additional complication to the optimisation process. It is found however that the A^nD blending method handles such a condition well.

The transfer trajectory is split into eight phases and propagated using Model 3. The first phase of the trajectory raises the orbit energy using the semi-major axis control law exclusively for approximately 27 days. Subsequent phases have similar intermediate aims, which all contribute towards the final complete trajectory. For example, the purpose of the second phase is the reduction of eccentricity to zero, while also targeting an ascending node value which aligns the orbit correctly for passage of the ascending node / Solar terminator offset angle through 90 deg, when avoidance of the shadow cone is critical. The constant by which the A^nD score is multiplied is determined by the relative importance of each element during that

Orbit Element	Value	Unit
Semi-Major Axis	6244.65	km
Eccentricity	0.53	–
Altitude of Pericentre	500.0	km
Altitude of Apocentre	7109.9	km
Inclination	90.00	deg
Argument of Pericentre	270.0	deg
Ascending Node	$(\Omega_{sun} + 10°) + 180°$ from N-pole value	deg
True Anomaly	228.44	deg

Table 7 Target orbit elements for forced Sun-synchronous orbit with south-pole pericentre.

particular phase of the trajectory, hence in phase 2 the eccentricity and right ascension angle have high constants and the semi-major axis constant is low. Additionally, at certain times the target eccentricity is set to zero, rather than that given in Table 7. This was found to reduce transfer time which while not the primary cost function of the transfer remains of critical importance. Furthermore, at certain times the control laws are not allowed to automatically determine which direction they should be driving the orbit; instead they are forced to always increase a certain element irrespective of current and target values. For example, the ascending node angle is always increased and never decreased. This has a slightly different effect from altering the target value, as the $\Lambda^a D$ score is affected in a different way but is done for similar reasons. Selection of such strategies is through engineering judgement.

Analytical analysis of Mercury, Sun and spacecraft position vectors can be used to confirm the entire transfer trajectory is shadow free. Recall from Section 3.4.2 the spacecraft shadow parameter is defined in Equation 124, thus using Equation 126, Equation 128, Equation 129 and Equation 131 the critical shadow parameter for penumbra conditions is defined as,

$$\varsigma_p = |\mathbf{r}| \sin\left[\arcsin\left(\frac{R_M}{|\mathbf{r}|} \right) + \arcsin\left(\frac{R_S - R_M}{|\mathbf{R_m}|} \right) \right]$$ Equation 162

Note in Equation 162 the radius of Mercury is not altered by an atmospheric constant as in Equation 129, since Mercury has only a very tenuous atmosphere. The analysis concentrates on penumbra shadow as this ensures the sail remains within full sunlight at all times. If the magnitude of the spacecraft shadow parameter is greater than the magnitude of the penumbra critical shadow parameter then the spacecraft is in complete Sun light on the night-side of the planet and if the shadow parameter is less than zero then the spacecraft is on the dayside of Mercury and thus in complete Sun-light, as discussed in Section 3.4.2. Post-processing the spacecraft shadow parameter output vector from the trajectory analysis one can remove the terms which correspond to the spacecraft being on the day-side of the planet. Figure 62 shows the penumbra critical shadow parameter through the transfer trajectory when on the planetary night-side. It is seen that at all times when the spacecraft is on the night-side of the planet the spacecraft shadow parameter is greater than the penumbra critical shadow parameter, thus confirming the trajectory is entirely shadow free.

Figure 63 shows the orbit argument of pericentre and inclination angles, which are shown to converge with their target values, while the right ascension angle is seen to terminate 10 deg ahead of the solar terminator. Finally, Figure 64 shows the altitude of pericentre and apocentre throughout the 142.3 day trajectory. Note the orbit eccentricity peaks at day 30, prior to an extended period where eccentricity is very low, which corresponds with the rapid variations in argument of pericentre angle in Figure 63. Figure 64 shows that both pericentre and apocentre converge well with the target values; with all the orbit elements reaching convergence with Table 7 values on day 142 of the simulation.
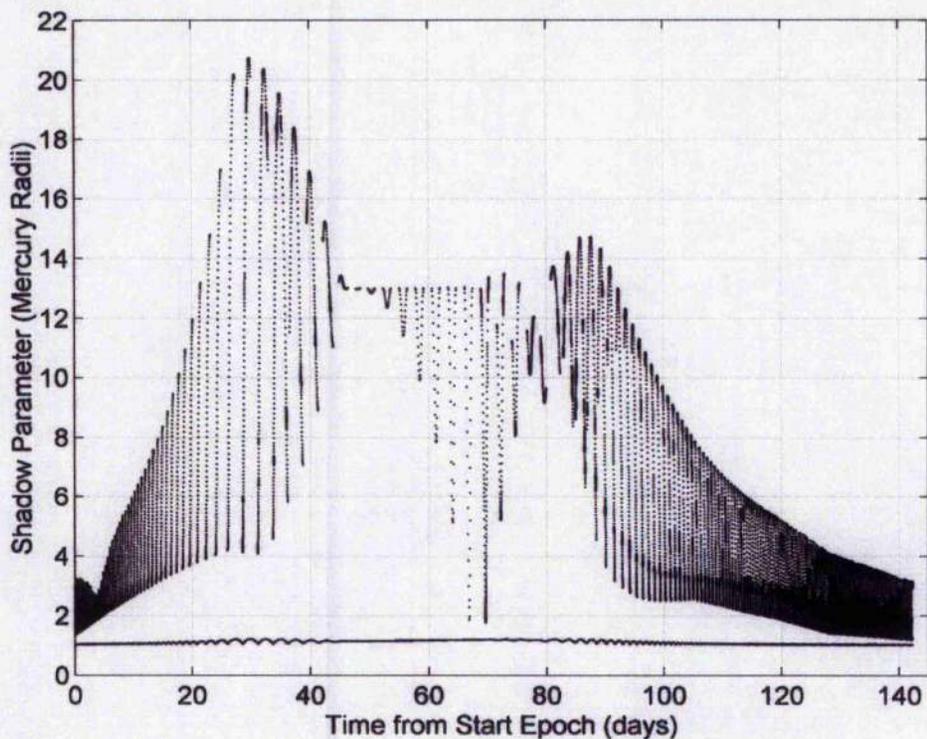
Figure 62 Shadow parameter (···), night-side of planet only, and penumbra critical shadow parameter (—).
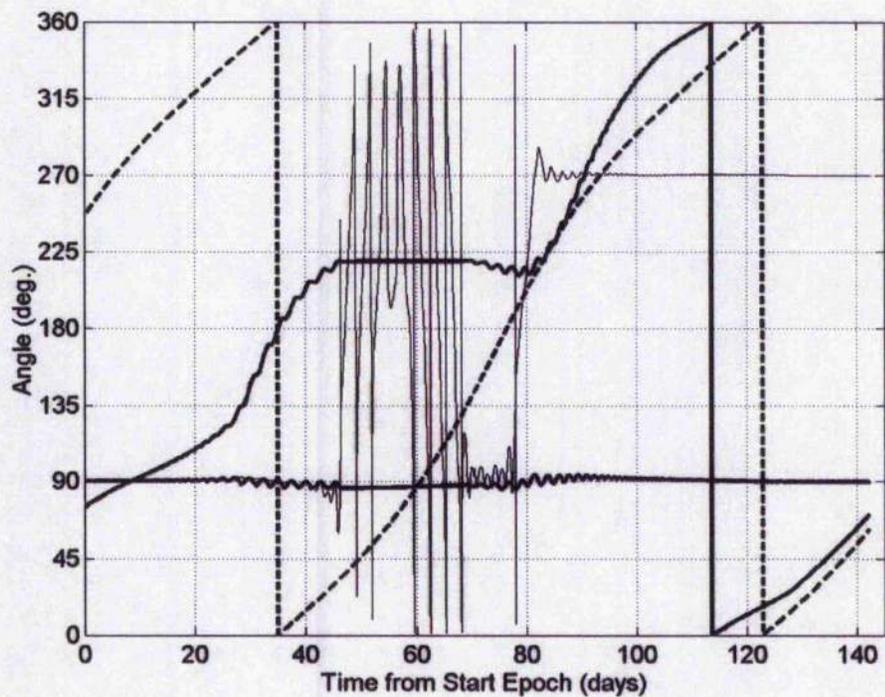
Figure 63 Ω (—, starting at ~54 deg); Ω of target solar terminator ground track (--); ω (light —) and i (—, starting at 90 deg and ending at 90 deg).
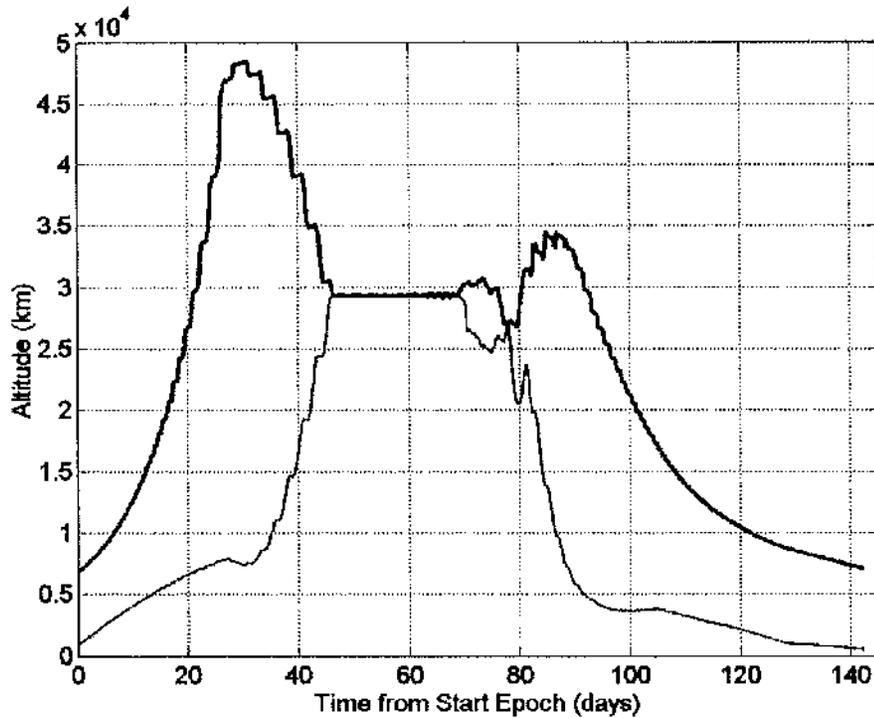
Figure 64 Instantaneous altitude of periherm and apoherm during transfer trajectory.

The transfer trajectory created using A^nD blending requires numerous rapid sail slew manoeuvres. Furthermore, the time optimality of the transfer at approximately 142 days is difficult to evaluate. However, recall that this was not the primary cost function of the transfer. The transfer trajectory is constrained by spacecraft thermal considerations and is verified as being shadow free. It is thus possible to generate a shadow free transfer trajectory between a north-pole pericentre forced Sun-synchronous orbit and an equivalent south-pole pericentre orbit. Moreover, one can state that the use of A^nD blending has enabled such a transfer to be generated much easier than would be possible with conventional optimisation tools.

5.3.2   Forced Sun-Synchronous Orbit with Active Sail Control Using A^nD Blending

One of the significant attractions of using a solar sail to generate a forced Sun-synchronous orbit is the lack of sail control requirements. The sail pitch angle is constant at all times through the orbit and as such sail attitude control could be maintained by mostly passive methods. If the sail were used to deliver the payload into Mercury orbit then an active sail control system would be required for the Earth – Mercury transit and the capture spiral. However, if the sail is delivered to Mercury by a chemical, electric or combination of systems then it could be deployed in Mercury orbit, allowing a relatively small, simple and low cost solar sail to be utilised. Such an approach would appear to be the best approach for a near-term, low to medium risk mission, as sail capture manoeuvres typically require a high sail slew rate capability and increase mission risk. It is worth noting however that sail deployment after being stowed for so long in space may present unique and complex design issues. The polar nature of the target orbit in this Sun-synchronous scenario allows a fixed sail pitch of $\arctan(1/\sqrt{2})$, the optimal fixed sail pitch angle as derived by Equation 155 in

Section 4.1, to be utilised while spinning the sail clock angle through 360 deg each orbit. Such a fixed pitch capture trajectory can actually be accomplished in a marginally shorter time than the locally optimal steering strategy requires. Both fixed pitch and locally optimal trajectories are shadow free.

The primary requirement of an active sail control system for a forced sun-synchronous orbit is clearly that the ascending node angle is rapidly rotated. It is therefore logical to investigate the use of the ascending node control law exclusively. The control law autonomously selects whether to increase or decrease the ascending node angle based on the current and target values. It was found however that the variation of pericentre and apocentre altitude through a single Hermian year is considerably more than the fixed sail pitch scenario over two Hermian years. This is a direct result of an increase in the amplitude of the long period oscillation of the argument of pericentre, as discussed earlier. It was also found that the right of ascension angle follows the solar terminator much more poorly as a direct result in the substantial variations in orbit size and shape. It is also noted that the inclination angle varies by as much as +5 deg and -6 deg from the nominal value of 90 deg. It is thus evident that an active sail control system must consider more than just the right ascension control law due to the unstable nature of the orbit.

It is found that in the fixed sail pitch scenario the orbit offset from the solar terminator is not a constant, instead varying from plus ten degrees, the nominal value, to as low as +3 deg and as high as +14 deg from the terminator. With such wide variations in the offset angle the surface illumination conditions would not be constant. Thus it is required that the active sail control system provide a much more constant solar terminator offset angle and hence a much more constant surface illumination angle. It was also mentioned earlier that an active sail control system could potentially allow a reduction in pericentre altitude. Such a reduction increases surface resolution imaging for science data. Accordingly, an active sail control system for a forced Sun-synchronous orbit is beneficial if it provides an improvement in surface illumination consistency and an increase in surface image resolutions, without any adverse effects on other orbit elements and parameters. During the fixed sail pitch scenario the offset angle from the solar terminator, the orbit eccentricity and semi-major axis each vary considerably. It is thus logical that when first attempting to generate an active sail control law one would attempt to control only these orbit elements. It was found however that when one controls only these elements that the orbit inclination and argument of pericentre vary significantly more than in the fixed sail pitch scenario, primarily due to the coupling in the out-of-plane perturbation terms. As a result of these initial findings it is apparent that an active sail control system must control orbit size and shape, while also considering the exact orbit plane location rather than just its ascending node angle.

When propagating station-keeping trajectories using the A^{n}D blending method the sail pitch angle is set at given discrete time-steps; typically these time steps are equal to or less than a quarter of the nominal orbit period. The pitch and clock angle are set by the current orbital elements without any forward-looking considerations through the duration of the time-step. The use of discrete time-steps has two objectives; primarily in the forced Sun-synchronous scenario it is implemented in order to reduce computation costs, as the variable step-size integrator has a tendency to take very small step sizes when the current and target elements are close together. By fixing the sail pitch for a given period of time the integrator can propagate

the orbit, while maintaining calculation accuracy without readjusting the sail at very small time-steps. Note that although the sail pitch is set in discrete time-steps the integrator is still variable step-size and typically takes many steps during each discrete pitch angle setting. The use of discrete sail pitch angle settings also has the effect of removing the potential requirement for multiple large angle slews in a very short period of time. Furthermore, if the discrete settings are suitably far apart the time required to slew the sail between settings can become significantly shorter than each discrete pitch time step; hence sail slew rate effects are minimised within the trajectory simulation. Such a scenario is investigated in Section 5.4 where A^nD blending is applied to station-keeping of the GeoSail trajectory.

The use of active sail control to generate a forced sun-synchronous orbit at Mercury allows the pericentre altitude to be significantly lowered. It was found that pericentre could be taken below 100 km and easily maintained within a narrow range of altitudes. However it was felt that altitudes below 100 km were undesirable, partly for science data acquisition and that in order to maintain very low pericentre values within safe bounds it was required to relax control of other orbit elements. As such the nominal forced Sun-synchronous orbit has as a 100 km pericentre altitude and a 7500 km apocentre altitude, with ground track displaced 10 deg ahead of the solar terminator. Using the constants detailed in Table 8 the final weight functions utilised to gain a forced Sun-synchronous orbit at Mercury with A^nD blending setting sail pitch and clock angle once per hour are gained. It was found that the control method had a tendency to bring the ascending node angle 10 deg ahead of the terminator and then allow it to drift backwards, before then acting to return it towards the nominal value. As a result the average offset angle from the solar terminator tended to be less than 10 deg. Therefore the target offset angle was set at 10.25 deg, forcing the ascending node ahead of its nominal value and then allowing it to drift backwards through the nominal value, hence gaining an average offset angle much closer to the nominal, as seen in Figure 65. It is also seen in Figure 65 that the offset angle is maintained within very tight bounds at certain times of the Hermian year, for example in days 0 to 20 and 80 to 100 the offset angle is maintained with $\pm$ 0.2 deg. Yet at other times the offset angle spikes at values as much as 0.5 deg away from the nominal. Such events are found to correspond to Mercury's perihelion passage and the significantly increased levels of solar radiation pressure acting on the sail. The ability of the A^nD blending method to maintain orbit control during such large fluctuations in orbit perturbations illustrates the adaptive nature of the control system. Recall from Figure 44 that the solar radiation flux varies from four times the value at Earth, at Mercury aphelion, to just over ten times at Mercury perihelion, a variation of 250 % in orbit perturbation magnitude during each Hermian year. Note that the maximum recorded sail pitch angle during the active sail control trajectory was 70 deg, thus while the active sail provides excellent orbit control it would require an agile sail.

Figure 65 shows the offset angle from the solar terminator for a fixed sail pitch angle of 10 deg and for active sail control. As stated earlier, it is seen that the fixed pitch sail has a much larger variation in offset angle than an actively controlled sail, which thus provides a much more consistent level of surface illumination for science data acquisition. Figure 66 shows the displacement of periherm and apoherm from their nominal values for fixed sail pitch and an active sail control. The active sail significantly reduces variation in orbit size and shape, with periherm varying by 40 km and apoherm by less than 100 km. Hence the active sail control would provide a much more constant surface resolution during science data acquisition. A final

consideration with the use of an active sail control system is that it should not adversely affect the orbit elements that were previously noted to vary little with a fixed sail control system. It was found that both the fixed sail pitch and active sail control system result in only very small variations in inclination and argument of pericentre.

Parameter	Constant factor on $\Lambda^n D$ score
Eccentricity	8
Semi-Major Axis	6
Inclination	20
Argument of Pericentre	4
Ascending Node	35
Radius of pericentre = Radius of apocentre	0

Table 8   Constants by which $\Lambda^n D$ scores are multiplied to gain final weight functions during forced Sun-synchronous orbits with active sail control.
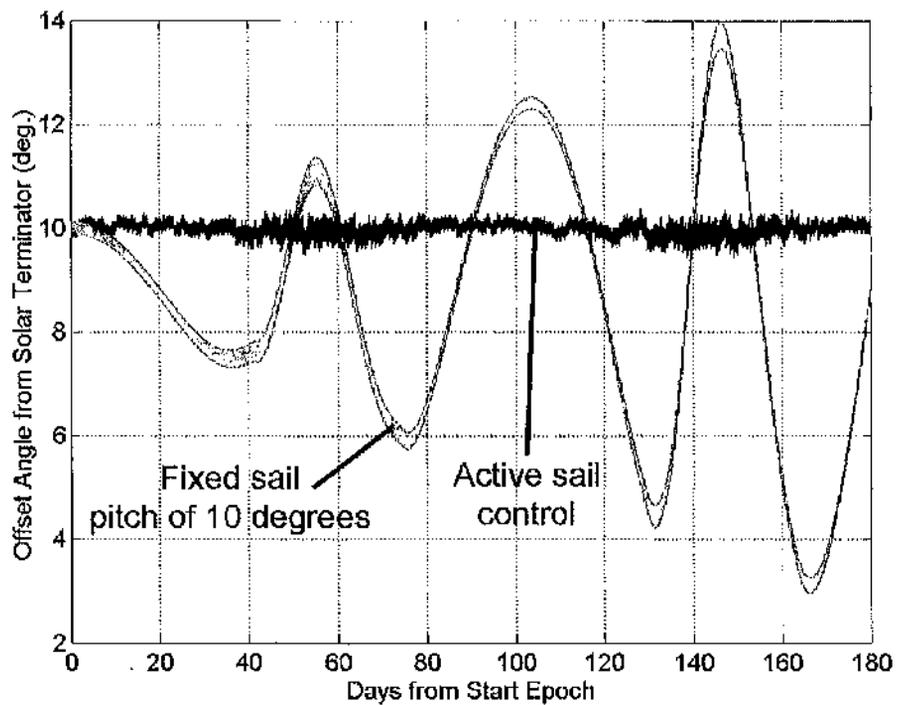
Figure 65   Orbiter/solar terminator offset angle, fixed and active sail control.
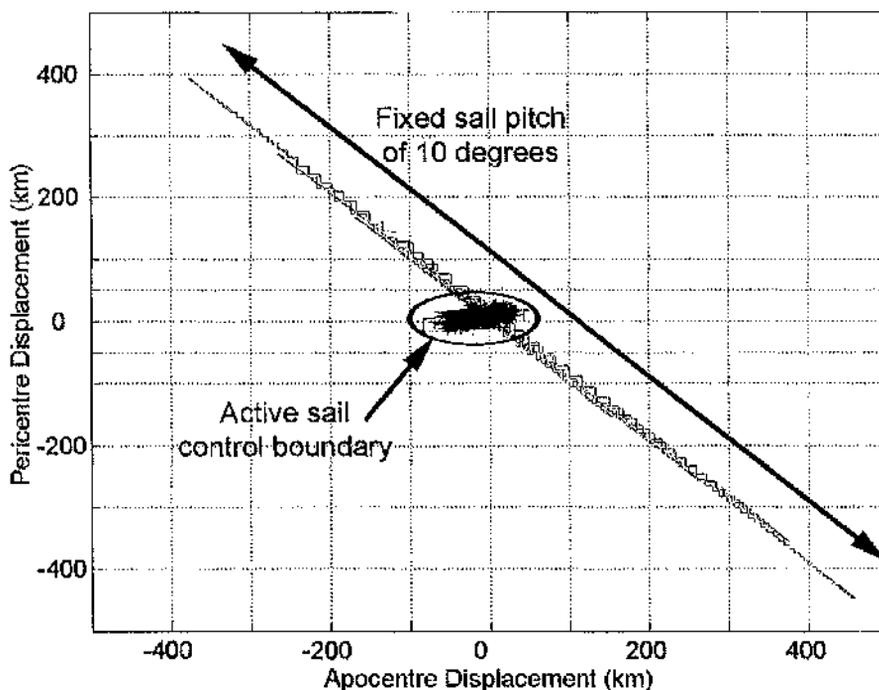
Figure 66 Displacement of periherm and apoherm from nominal values, fixed and active sail control.

5.4   GeoSail Orbit Control Using A^{n}D Blending

The use of active sail control for forced Sun-synchronous orbits at Mercury has been shown to significantly improve orbit uniformity, as illustrated in Figure 66. The short orbit period however necessitates many rapid sail slew manoeuvres. Conversely the GeoSail mission utilises a much larger orbit, with a nominal period of just over four days. The GeoSail mission concept is discussed in Section 1.5. The A^{n}D blending method for active sail control is considered within the GeoSail mission for a similar purpose to that for which it was considered within the Mercury forced Sun-synchronous orbiter mission. That is to say, while the GeoSail orbit is a more naturally stable scenario an active sail control system is considered to maintain a more precise orbit than would otherwise be attained by a fixed sail pitch scenario, with the corresponding oscillations in orbit elements. Additionally, it is considered whether the application of a fixed upper limit to the sail pitch angle can be implemented. The imposition of an upper pitch angle means that one can expect much smaller sail slew angles between discrete sail pitch settings. It also simplifies spacecraft design if the solar aspect angle is more constant. Setting the duration of each discrete sail pitch angle as one day, just under a quarter of the nominal orbit period, it was found that using a single set of constants on the A^{n}D scores good orbit control could be maintained down to an upper pitch angle of 15 deg. Reduction of the maximum pitch to 10 deg was found to be overly restrictive and the control system required an increase in sail performance. The required sail acceleration is determined by consideration of the orbit size and shape as defined in Reference 102; the required sail characteristic acceleration is 0.0999 mm s^{-2}.

Model 6 is used within the remainder of this section. Using the constants detailed in Table 9 the final weight functions utilised to generate a GeoSail orbit with active sail control are obtained. Note that the factor by which the argument of perigee is multiplied by is much larger than the other control laws. The value of each factor reflects the importance of the orbit element. Similar to the Mercury forced Sun-synchronous orbit one finds it is beneficial to target the primary orbital element, the argument of perigee, ahead of the nominal value. Thus, the argument of perigee is targeted 0.2 deg ahead of the Earth-Sun line. The maximum sail pitch angle allowed is 15 deg; thus if the blended locally optimal pitch angle is greater than 15 deg the pitch angle is set as 15 deg, with the clock angle allowed to take any value between 0 deg and 360 deg. Similar to the Mercury forced Sun-synchronous scenario, the sail control angles are set by consideration of only the current orbit elements and no forward-looking considerations are taken. Figure 67 shows the displacement of perigee and apogee from the nominal values of eleven and twenty-three Earth radii respectively, for the fixed sail pitch and active sail control scenarios. It is found that the active sail control scenario using A^{n}D blending to select the sail control angles reduces the variation in orbit perigee and apogee, thus providing a much more consistent orbit shape and size. The radius of perigee is seen in Figure 67 to be centred on a value of approximately 10.8 Earth radii, rather than the nominal value of 11 Earth radii. If desired the target perigee could be raised to 11.2 Earth radii, thus the orbit would shift towards the nominal GeoSail orbit in a similar manner to that used for accurate targeting of the argument of perigee. Note however that this is a magnetotail science mission and as such the primary science requirement is for accurate control of apogee and the argument of perigee thus no such adjustment was adopted.

Figure 68 shows the variation in the angle between the Earth-Sun line and the orbit major axis for a fixed and active sail. Note that the orbit major axis and Earth-Sun line vary by as much as 3 deg for a fixed sail pitch, while the A^{n}D blending method reduces this variation to less than 1 deg. Finally, Figure 69 shows the sail control angles generated by the A^{n}D blending method for the GeoSail scenario with active sail control. Typically the sail pitch angle is 15 deg with the clock angle rotating the sail thrust vector either left or right of the orbit major-axis. As such the maximum required sail slew angle between discrete sail pitch settings is 30 deg; however also note that on occasion the sail does not move for as much as three to five days. The technology requirement for sail slew capabilities in an active sail mission scenario are thus defined as 30 deg in 1.25 − 2.5 hrs, that is to say in 5 − 10 % of the duration of each discrete sail setting. The time for each slew manoeuvre is thus significantly less than the duration of each discrete set of control angles. GeoSail is a demonstration class mission and as part of an extended mission active sail control using A^{n}D blending could be demonstrated over and above the basic solar sail demonstration capabilities of the GeoSail mission.

Parameter	Constant factor on A^nD score
Radius of pericentre	35
Radius of apocentre	25
Semi-Major Axis	2
Argument of Pericentre	170
Eccentricity = Inclination = Ascending Node	0

Table 9 Constants by which A^nD scores are multiplied to gain final weight functions during GeoSail trajectory with active sail control.
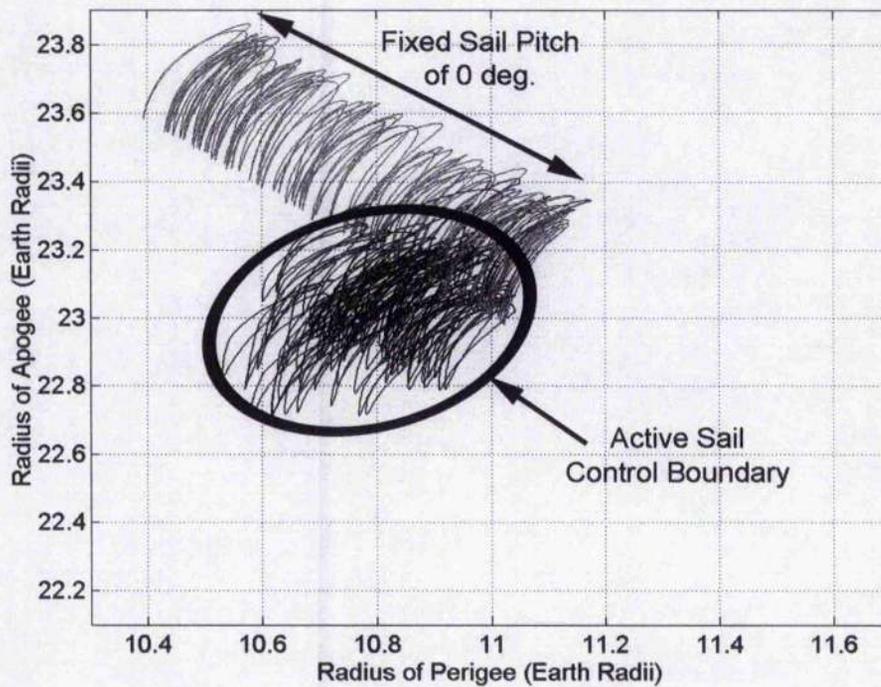

Figure 67 Displacement of perigee & apogee from nominal values, fixed and active sail control.
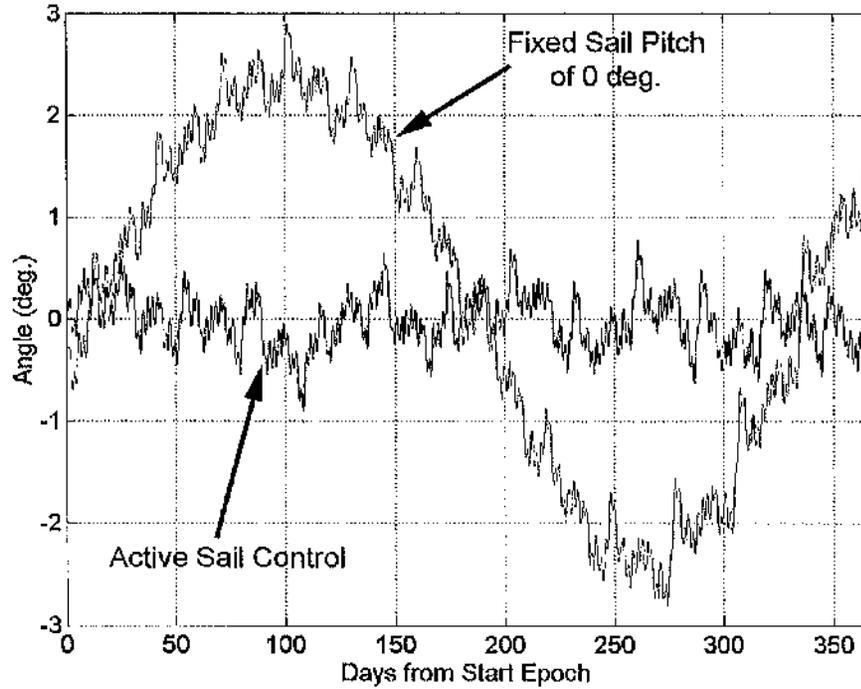
Figure 68 Variation in angle between Earth-Sun line and orbit major axis.
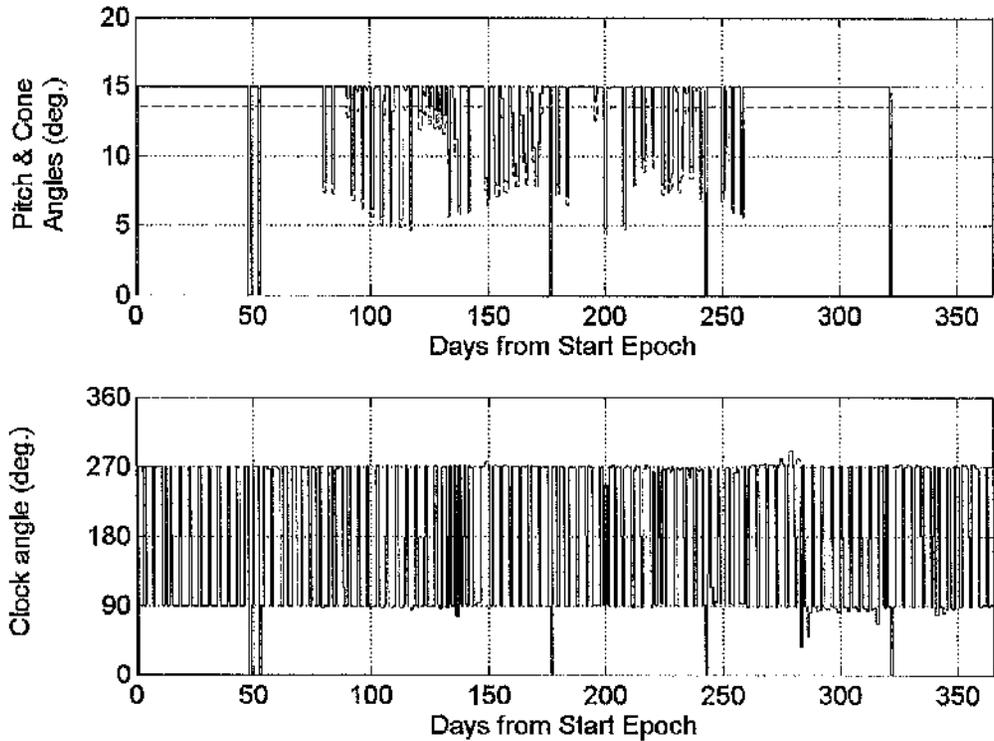
Figure 69 Active sail control angles selected using A²D blending for GeoSail.

6 Heliocentric Orbit Transfers Using a Solar Sail and Locally Optimal Control Laws

This section evolves the control method developed in Section 5 to allow the rapid generation of heliocentric trajectories for either, top-level analysis of mission concepts where only a near-optimal trajectory is required, or for generation of near-optimal trajectories as initial guesses for further optimisation by, for example, a Sequential Quadratic Programming (SQP) algorithm. The use of locally optimal control laws has previously been stated as a means of finding an initial guess for optimisation algorithms.[1, 35] However, little work has been presented as to the validity of such claims. A recent paper used the locally optimal energy gain/reduction control law to generate solar system escape trajectories which for high-performance sails were relatively close to those produced using complex optimisation procedures.[198] This work however struggled to produce near-optimal escape trajectories when the sail performance was lowered and the trajectory became more complex. It is also noted that for a high performance sail the solar system escape trajectory is relatively simple; hence the adopted method within Reference 198 would be expected to approximate the global optimal.

Initially the algorithms used to blend the control laws will be modified in-order to improve their efficiency. The new A^{n}D algorithms will then be used to analyse three different heliocentric solar sail mission scenarios, where the results will be compared with similar results generated using a SQP method and an evolutionary neurocontroller, InTrance.[137, 154, 199, 200]

6.1 A^{n}D Score Evolution

In Section 5 the A^{n}D blending method is used for planet-centred trajectories, thus when calculating the deficit score of each element Equation 122 and Equation 123 must be applied. As a result when blending many elements, as in Sections 5.3.2 and 5.4, the calculation of the deficit score can be a computationally intensive phase of each trajectory step. Consequently in Section 5.2 the deficit score is found through application of an ideal solar sail force model. However, within this section Equation 122 and Equation 123 need not be applied as the Sun-sail line reference frame is coincident with the *RTN* reference frame. The deficit score is thus evolved such as to be defined through the application of the same sail force model used to determine the sail thrust vector. The assumption of an ideal sail when deriving the control laws now no longer compromises the control law efficiency when a non-ideal sail force model is used.

6.2 Interstellar Heliopause Probe

The core of an Interstellar Heliopause Probe mission was outlined in Section 1.5.2. Recall that the IHP mission is a TRS; as such the mission requirements within this dissertation follow the TRS requirements. These state that the spacecraft should be delivered to latitude 7.5 deg and longitude 254.5 deg at 200 AU from the Sun, in the ecliptic coordinate frame, that is to say the nose of the heliosphere. The spacecraft should reach a solar distance of 200 AU in 25 years or less.[110, 111] Recall further from Section 1.5.2 that the

sail should be jettisoned by the spacecraft at 5 AU to eliminate any potential interference caused by the solar sail on the local space environment.[86, 87]

Solar radiation flux drops off as $(1/r^2)$, thus a solar sail becomes increasingly ineffective at large solar distances. Equally, solar sails become increasingly effective at low solar radii. It was thus realised by Sauer that a close approach of the Sun by a solar sail would allow sufficient velocity to be gained to rapidly escape the solar system. Such a manoeuvre is often termed a Solar Photonic Assist (SPA).[69, 114, 126, 127] The logic behind a SPA trajectory is that the perihelion velocity, for a fixed perihelion radius, increases as the aphelion radius is increased. Thus, the required energy change to gain escape velocity from a fixed perihelion radius is minimised for ever increasing aphelion radius. However, the gradient of increase in perihelion velocity is reduced as aphelion radius is increased, as shown in Equation 163,

$$v_p = \sqrt{\frac{2\mu}{r_p}\left(1 - \left[1 + \frac{r_a}{r_p}\right]^{-1}\right)}$$ Equation 163

It is thus clear that at perihelion passage the instantaneous aphelion radius should be maximised, hence minimising the required energy change for solar system escape and maximising the solar system hyperbolic excess velocity. However, if the radius of aphelion passage is too large the sail will spend a prolonged period within the inner solar system and the total trip time to 200 AU will be increased. Thus, an optimal aphelion passage radius must exist, above which the extra hyperbolic excess velocity gained through an increased aphelion radius will not compensate sufficiently for the extra time spent in the inner solar system. The optimal aphelion radius will increase as the target solar radius is increased. It has already been shown that the optimal aphelion radius is reduced as sail performance is increased.[114]

In Section 1.4.2 it is discussed how close a solar sail can approach the Sun and that the traditionally accepted minimum radius is approximately 0.25 AU, set through consideration of the thermal bounds of the sail film substrate. However, recall also from Section 1.4.2 that the real thermal limit is set by constraints on not only the sail film but also on the booms, other sail components and the thermal constraints of the spacecraft attached to the sail. A thermal limit of 0.25 AU will be adopted for most trajectories within this sub-section, however as A"D blending enables rapid generation of near-optimal trajectories the impact on trip time of varying the thermal limit will also be considered. Similarly, the effect of a non-ideal sail will be analysed through the use of Model 4, while a uniformly bright finite solar disc will be considered through the use of Model 5 and optical surface degradation considered through use of Model 7. Finally, removing the minimum radius bound and replacing it with a temperature boundary will be considered as has previously been performed by Dachwald in Reference 137 and 154. Note that some work has been conducted previously using this non-ideal sail force model for rendezvous trajectories and for Heliopause trajectories.[137, 154, 199 - 205] Thereby it was found that flight times are up to 5 % and on occasion up to 10 % longer for the non-ideal sail force model.

The first phase of a SPA trajectory, for high-performance sails, is the increase of orbit eccentricity. It was found that by using the eccentricity locally optimal control law the break-point between a single and dual-loop trajectory could be identified as shown in Figure 70, generated using Model 1 and Model 4 which are used unless otherwise stated within the remainder of this section on HIP trajectories. The trajectory was propagated from Earth departure until the first perihelion passage for a range of sail characteristic accelerations. It was found the minimum characteristic acceleration for a single loop trajectory with thermal limit 0.25 AU is 1.5 mm s^{-2}, for an ideal sail. A characteristic acceleration below 1.5 mm s^{-2} requires a dual SPA for an ideal sail with thermal limit 0.25 AU. For a non-ideal sail the break-point rises to 1.6 mm s^{-2}, as shown in Figure 70. It was found that even though a characteristic acceleration below the break-point can escape the solar system without touching the minimum radius, the trip time to 200 AU is reduced if a second revolution about the Sun is added to bring the final solar pass down to the minimum bound. In the same way, if the characteristic acceleration is reduced sufficiently a third revolution is required to attain the minimum allowed radius and to minimise the trip time to 200 AU.

6.2.1   Minimum Performance Single Loop Trajectory with Ideal Sail

Having established the minimum characteristic acceleration required for an ideal sail with thermal limit 0.25 AU to reach 200 AU with a single SPA is 1.5 mm s^{-2}, now consider a two-dimensional scenario; allowing rapid quantification of the problem in hand prior to introducing an inclination change to the trajectory design. Initially, the trajectory was generated using only the eccentricity control law. However, this results in a minimum radius of less than 0.25 AU. The trajectory was thus split into two phases, an eccentricity increase phase and an energy gain phase, with the first finishing after 659.6 days. The second phase uses the semi-major axis control law to raise orbit energy in a locally optimal manner, with the sail being jettisoned at
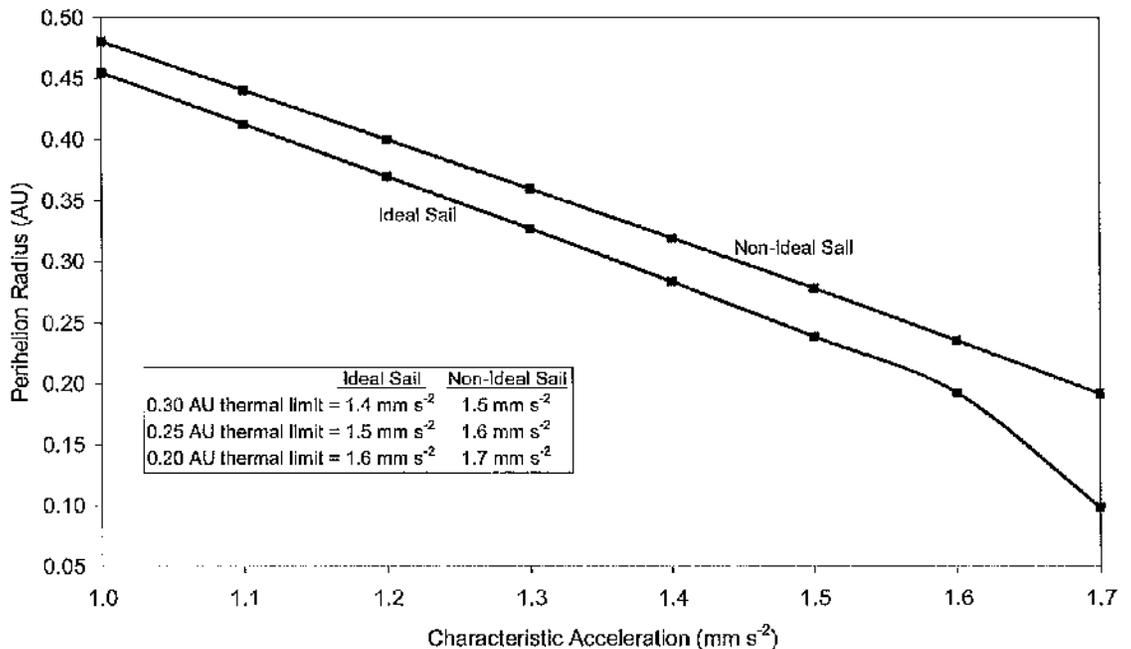
Figure 70   Required sail performace for single SPA to escape for range of minimum solar approach radii for ideal and non-ideal sail using the optical force model.

5 AU. When the controller is switched from eccentricity to semi-major axis the instantaneous radius of perihelion increases until perihelion passage. It is thus a simple matter of iteration to find the correct switch point to match the correct minimum radius. This initial trajectory was found to reach 200 AU in 22.96 years and is shown as the left most data point in Figure 71. The aphelion passage of this trajectory occurred at 2.50 AU.

To find the best-case 2D trajectory the aphelion passage radius must be varied. Using A^nD blending one can introduce the locally optimal pericentre control law to reduce the aphelion passage radius or the semi-major axis locally optimal control law to increase the aphelion passage radius. It was found that an increased radius was required. Phase 1 thus changes from the eccentricity control law only, to setting the eccentricity constant on the A^nD score as one and semi-major axis constant as the values shown in Figure 71. Note that Figure 71 is not constrained to the target longitude and is instead for a fixed Earth departure date of 03 January 2030, approximately at Earth perihelion passage. It was found that the most favourable 2D trajectory to 200 AU, with minimum radius 0.25 AU, using an ideal sail with characteristic acceleration 1.5 mm s^{-2} has an aphelion passage of approximately 2.83 AU and reaches 200 AU in 22.73 years. Note that the instantaneous aphelion value at perihelion passage is 10.7 AU. Figure 71 also shows the velocity of the spacecraft at 5 AU, the sail jettison point, where it is noted that the spacecraft is travelling at 10.50 AU yr^{-1} in the most favourable trajectory. One can increase this velocity to 10.51 AU yr^{-1} by increasing the aphelion passage to 2.90 AU. However, this increases the trip time to 200 AU. Similarly, one can reduce the velocity at 5 AU to 10.25 AU yr^{-1} by setting the aphelion passage at 2.5 AU. It is found that this trajectory is faster than the 2.83 AU trajectory to a solar radius of 125 AU, showing that the optimal aphelion passage radius increases in line with the target solar radius.
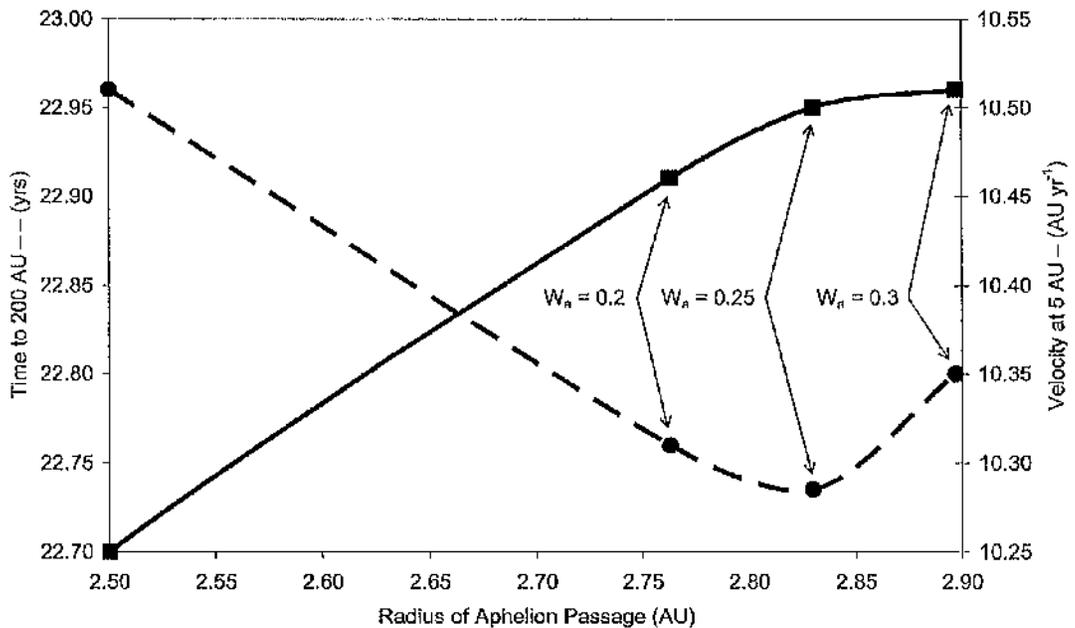
Figure 71 2D open azimuth trajectory trip time to 200 AU (– –) against radius of aphelion passage, with velocity of spacecraft at 5 AU (—) and constants used shown.

Using Figure 71 one can now consider the introduction of inclination to the trajectory design, recalling that the target point has latitude 7.5 deg in the ecliptic coordinate frame. It is immediately clear that the change of inclination is low and thus a high performance sail would be expected to easily impart such an angle change. This angle change should clearly be regulated and as such the best-case trip time was generally found if the inclination equalled its target value at the end of phase 1 of the trajectory, i.e. immediately prior to the beginning of the energy gain phase. This is a reasonably logical finding. Increasing the inclination during the energy gain phase results in a less than optimal energy gain as the thrust vector is pulled out of the plane of the velocity vector, thus it cannot be entirely directed in the same sense. Similarly, if the inclination is increased rapidly it results in a reduction of the instantaneous aphelion radius at perihelion passage, thus reducing the spacecraft velocity at 5 AU when the sail is jettisoned. Using a start epoch of 03 January 2030 an initial analysis of 3D trajectories to match the IHP requirements suggested that the optimal aphelion passage would increase to just over 3 AU. However, it was found that the velocity at 5 AU would likely remain very similar to the 2D case, with trip time also varying only slightly. A start epoch of 03 January gives an azimuth of order 230 deg at 200 AU. It is thus required that one rotates the orbit by approximately 25 deg to reach the nose of the heliosphere at 254.5 deg. The Earth rotates 25 deg in approximately 24.5 days, assuming a circular Earth orbit. Thus, the optimal Earth departure date is estimated as 27 − 28 January 2030. Using this estimate a launch window analysis was performed for a 1.5 mm s^{-2} ideal sail, with a thermal limit of 0.25 AU. It was found that for a given launch date the spacecraft azimuth at 200 AU could be varied by increasing or decreasing the aphelion passage radius, a larger azimuth being gained by increasing the aphelion passage radius. Thus, for a given launch date there is an optimal aphelion passage radius which will send the spacecraft to the correct azimuth at 200 AU and by varying the launch date the required aphelion passage radius and thus the trip time are varied, generating a launch window scan. It was found that the best launch date for a 1.5 mm s^{-2} ideal sail, with thermal limit 0.25 AU, was 26 January 2030. Trip time to 200 AU is 23.17 years, with a spacecraft velocity of 10.47 AU yr^{-1} at 5 AU. The best open azimuth 2D trip time found was just over 5 months (< 2 %) shorter than the 3D trajectory, with the velocity at 5 AU being very similar.

The launch date scan is shown in Figure 72 from 23 January 2030 until 01 February 2030, with a maximum azimuth error at 200 AU of ± 0.2 deg. Furthermore, it is noted that the inclination / latitude convergence is within 10^{-8} deg for all three-dimensional IHP trajectories within this dissertation. The launch date scan should repeat with a 1 year period, although small variations may occur due to the variation of the Earth's orbit with time. From Figure 72 it is noted that a delay from 26 January until 01 February results in an increase in trip time of 80 days, or 1 %.

The radius of aphelion passage is plotted against trip time and velocity at 5 AU in Figure 73, where it is seen that the most favourable radius of aphelion passage is approximately 3.14 AU. The data points in Figure 72 map onto the data points in Figure 73, with 23 January at the right-hand side and 01 February on the left-hand side. The constants applied to the A^{n}D scores are mapped onto the data points in Figure 73; the constant of eccentricity is not given as it was held constant at one. The constant of inclination has been rounded to two decimal places within Figure 73.
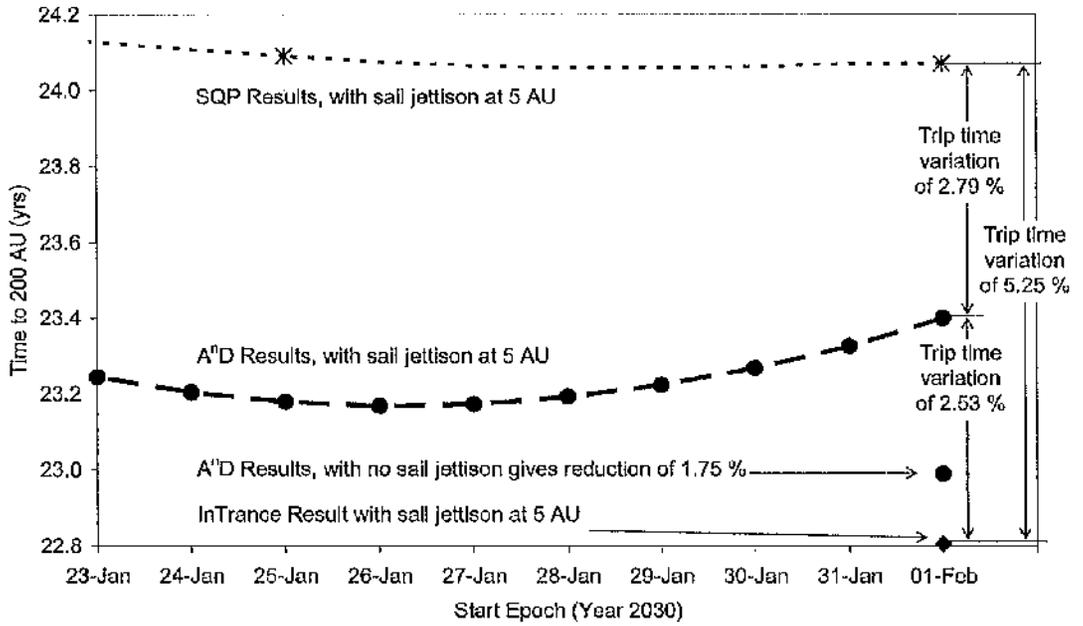
Figure 72 Launch date scan for an ideal sail with characteristic acceleration 1.5 mm s^{-2}.
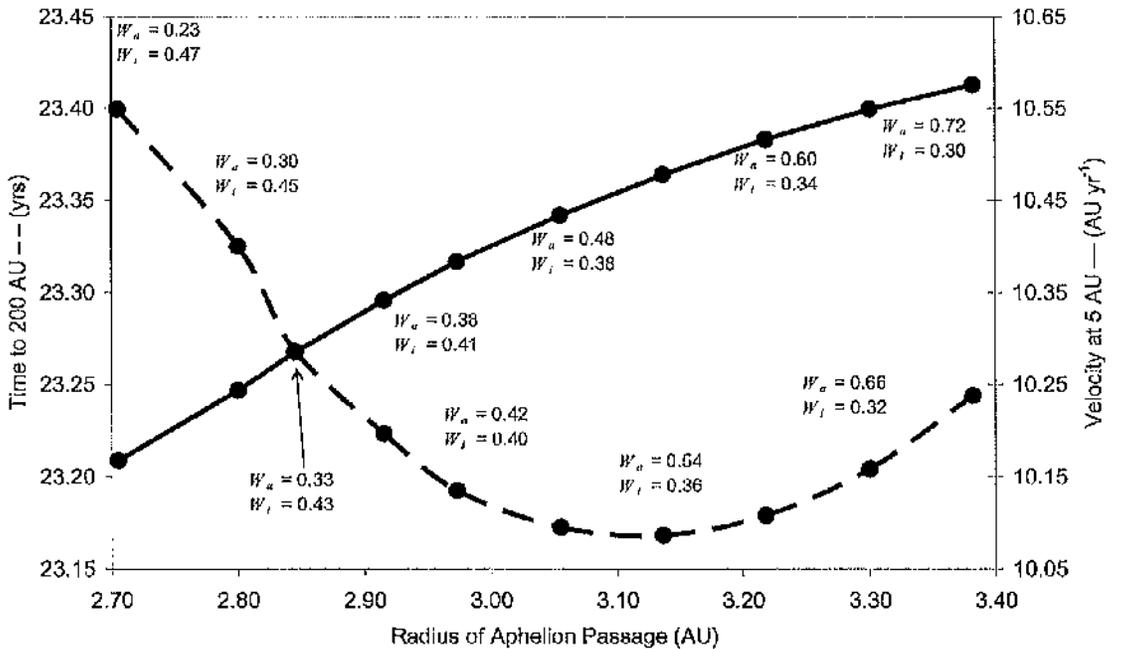
Figure 73 3D restricted azimuth trip time (—–) and velocity of spacecraft at 5 AU (- -) against radius of aphelion passage. Constants in phase 1 shown; eccentricity constant fixed at one.

The 26 January trajectory from Figure 72 is shown in Figure 74, with the corresponding sail control angles shown in Figure 75; note the cone angle equals the pitch angle as this solution is for an ideal sail. Note from Figure 75 that the maximum sail pitch angle is 81 deg. However, it is seen that this maximum occurs as a short discontinuity in an otherwise smooth profile. The discontinuity in sail pitch angle is due to the minimum radius bound and does not represent the near-optimal nature of this trajectory as will be seen in Section 6.2.5 when optical solar sail degradation is considered. The belief that optimal control angle profiles will always be smooth is an inaccuracy if restrictions such as minimum radius are placed on the trajectory. Previously generated solutions within the public domain have used short coast-arcs to avoid a radius below the minimum allowed value; Figure 75 strongly suggests that coast arcs are sub-optimal for this scenario of SPA trajectory and perhaps for other SPA trajectory scenarios. The size of the discontinuity illustrates how much of an effect the radius boundary is having on the trajectory. A large discontinuity means the trajectory would ideally pass much closer to the Sun, while a small discontinuity means the trajectory would ideally only pass a short distance closer to the Sun.

The constants applied to the A^nD scores for the 26 January trajectory in phase 1 (duration 977.4 days) were, eccentricity equal 1, semi-major axis equal 0.54 and inclination equal 0.35602, with the second phase using the locally optimal semi-major axis control law exclusively. The inclination constant is found iteratively, such as to match the inclination to the target value as close to the end of phase 1 as possible. The variation in the weight values through the trajectory are shown in Figure 76. It was found that the sail only just reached the minimum perihelion bound as shown in Figure 77, thus confirming that 1.5 mm s^{-2} is indeed the minimum ideal sail characteristic acceleration for a single SPA trajectory. From Figure 77 it is seen that the inclination reaches its target value significantly before the SPA, allowing the sail to concentrate on gaining orbit energy. If the sail is used beyond 5 AU only a small reduction in trip time is obtained. This is shown in Figure 72 for a launch on 01 February where a saving of 1.75 % is made. This corresponds with prior results.[114] Furthermore, modelling the Sun as a uniformly bright finite disc (using Model 5) altered the trip time to 200 AU by less than a fifth of a day.

To quantify the optimality of the A^nD blending method the best-case trajectory can be compared with independently generated trajectories using a SQP method and an evolutionary neurocontroller, InTrance.[137, 154, 199, 200] The same sail characteristic acceleration, thermal limit and sail force model was used in these methods, with the sail being jettisoned at 5 AU. Each scenario begins with an Earth departure C_3 of zero. The optimal duration of the SQP generated trajectory, using 201 control nodes, is shown in Figure 72, where it is seen that the optimal launch date was found to be 01 February, giving a trip time of 24.07 yrs, 2.79 % longer than the equivalent A^nD trajectory on 01 February. The SQP trajectory passes through aphelion at a radius of 2 AU, compared with the A^nD blending aphelion of 2.7 AU. Furthermore, the spacecraft has a velocity of only 9.7 AU yr^{-1} at sail jettison, compared with 10.16 AU yr^{-1} using A^nD blending. It is interesting to note that the maximum pitch of the SQP generated control profile is 90 deg and that the inclination does not reach its target value until after perihelion. The sub-optimal nature of the SQP trajectory is troubling, especially the post-perihelion change of inclination. However it is proposed by the
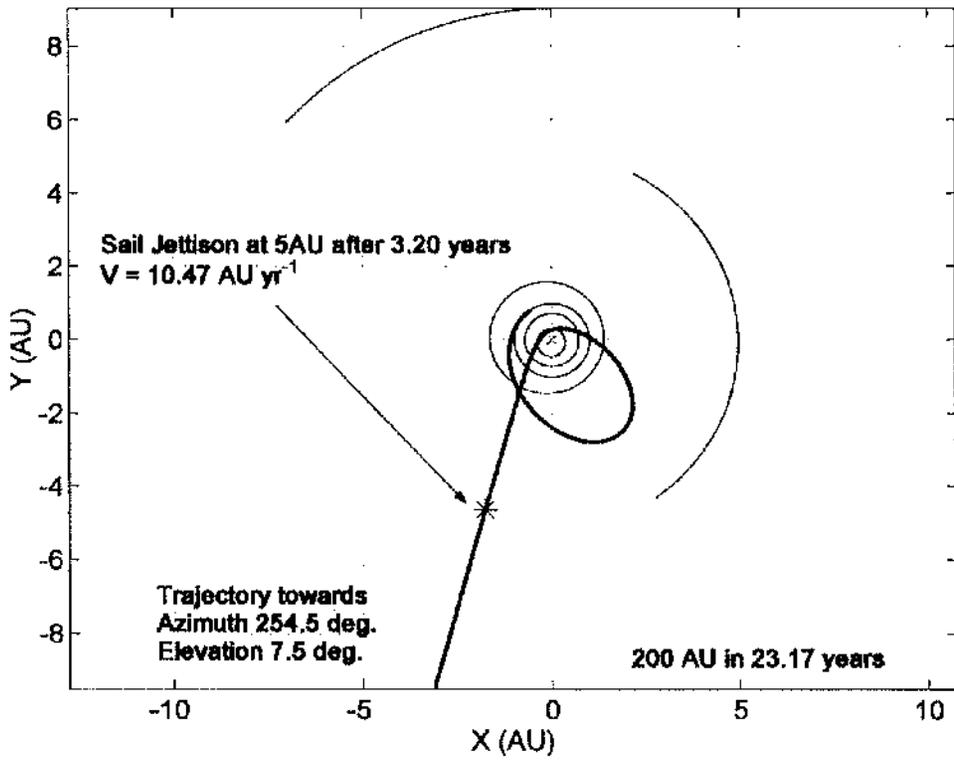
Figure 74 26 January 2030 trajectory to 200 AU using Model 1 with characteristic acceleration 1.5 mm s^{-2}
and thermal limit 0.25 AU.
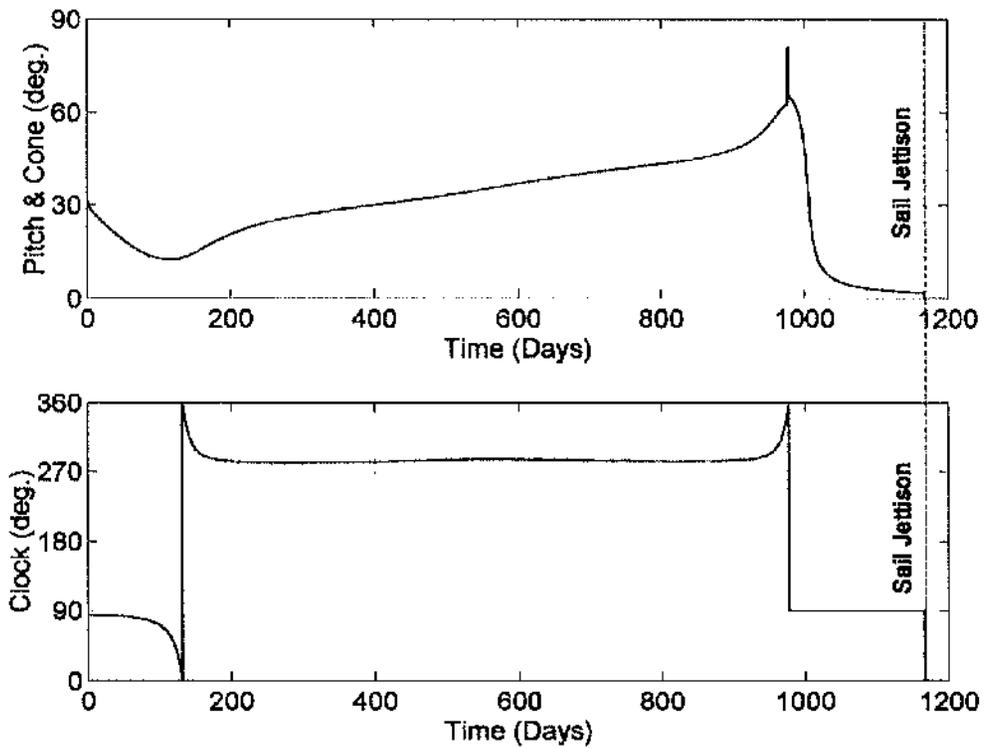

Figure 75 Control angles used in Figure 74.
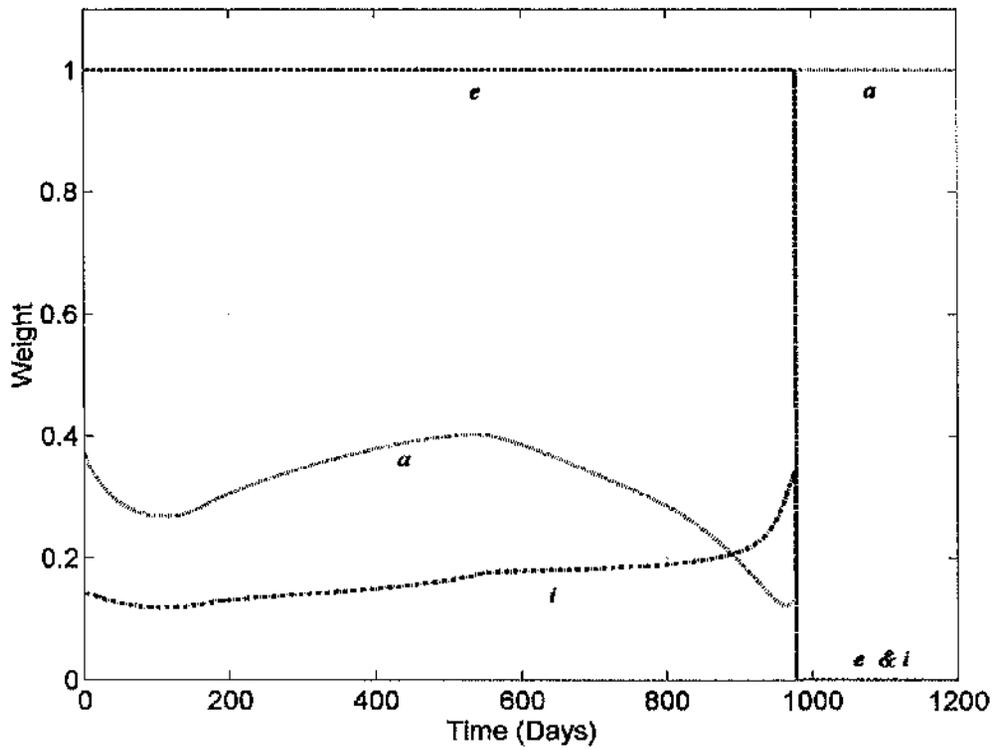
Figure 76 Weights used in Figure 74 and Figure 75.
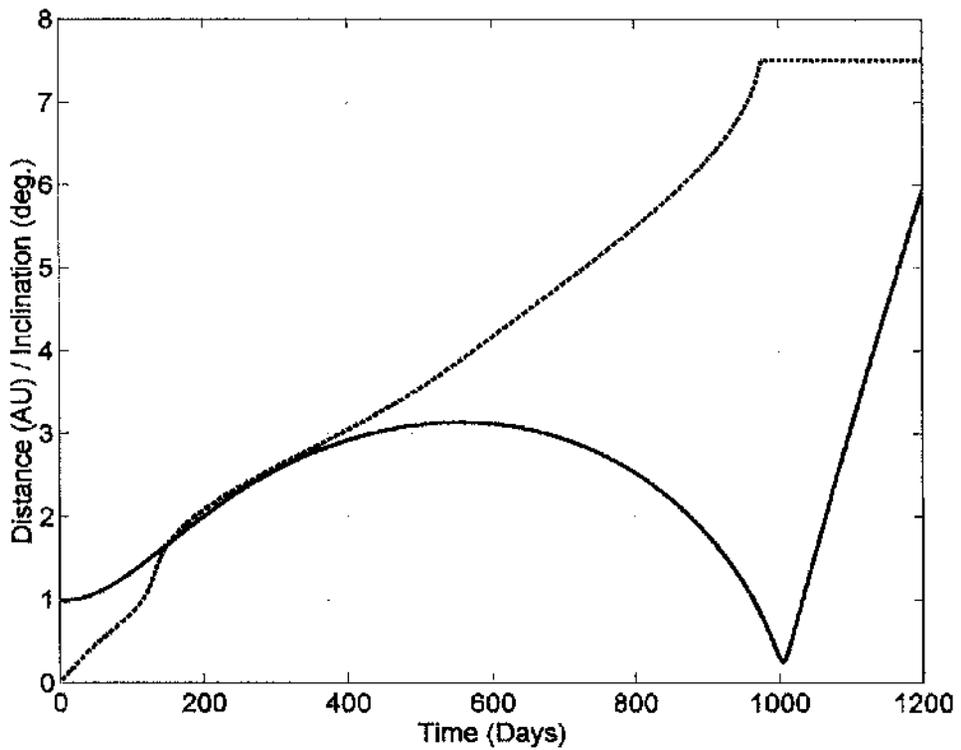
Figure 77 Radius (—) and inclination (--) of trajectory plot in Figure 74.

candidate that either this trajectory is a local optimal, which seems unlikely given the simple nature of the search space or more likely is due to the uniform spacing of the nodes, such that the rapid perihelion passage is not captured well. This however can only partially explain the sub-optimal nature of the SQP solution and not the post-perihelion inclination change. Thus, to quantify the optimality of the A^nD blending method only the InTrance results can be used. The evolutionary neurocontroller, InTrance, is discussed in detail in Reference 137, 154, 199 and 200. Note that this method does not require an initial guess to converge to an optimal solution and can thus be left to run unsupervised. To maintain consistency the InTrance trajectory start epoch was fixed as 01 February 2030, thus allowing the sub-optimal nature of both the SQP and A^nD results to be found. It is seen in Figure 72 that the InTrance optimal trip time to 200 AU is 22.81 years, which is 7 months shorter than the equivalent A^nD trajectory and over 15 months shorter than the equivalent SQP result. Furthermore, the InTrance solution has a velocity at 5 AU of 10.51 AU yr^{-1}, which is faster than the A^nD trajectory and partly explains the faster trip time. The InTrance trajectory has an aphelion passage of 2.90 AU, which is greater than the A^nD trajectory and further explains the improved trip time. One can thus conclude that the A^nD blending method has generated a trajectory to within 2.5 % of the best solution found, while the SQP method has generated a trajectory which is 5.25 % slower than the best solution. It is noted that the reduction of the trajectory from a 402-dimension SQP solution, i.e. 201 nodes of pitch and 201 nodes of clock, into a 2-dimension A^nD solution, that is to say finding the optimal weight of semi-major axis and inclination, with the eccentricity weight fixed, results in a significant reduction in computational effort. For example, if the solution is discretised every ten steps the A^nD method would have 10^2 solutions, while the SQP method has 10^{402} solutions, thus the computational effort is reduced by the order of 10^{400}. Note the InTrance method optimises over many more nodes than the SQP solution, thus InTrance is significantly more computationally intensive than either SQP or A^nD.

6.2.2   Minimum Performance Single Loop Trajectory with Non-Ideal Sail

Recall from Figure 70, it was estimated that a non-ideal sail, with a thermal limit of 0.25 AU, requires a characteristic acceleration of 1.6 mm s^{-2} to optimally reach 200 AU with a single SPA. This scenario was thus investigated to demonstrate the capability of A^nD blending to cope with a more complex sail force model. Figure 78 shows a plot of open azimuth trip time to 200 AU against radius of aphelion passage, with constant of semi-major axis and inclination used in phase 1 also shown. Start epoch is set at 03 January 2030. Once again, the constant of eccentricity was equal to 1 at all times during phase one. The trajectories were split into the same two phases as the previous scenario, with the inclination once again matching its target value at the end of phase one. The minimum time trajectory constants used in phase 1 were eccentricity equal to 1, semi-major axis equal to 0.30 and inclination equal to 0.44437, with the semi-major axis controller used exclusively in phase two. It is shown in Figure 78 that the minimum trip time found was 23.164 years, which is similar to the time found for an ideal sail with characteristic acceleration 1.5 mm s^{-2}, as was discussed above. It is interesting to note that the minimum time trajectory from Figure 78 has a set of very similar weight function plots as those shown in Figure 76.
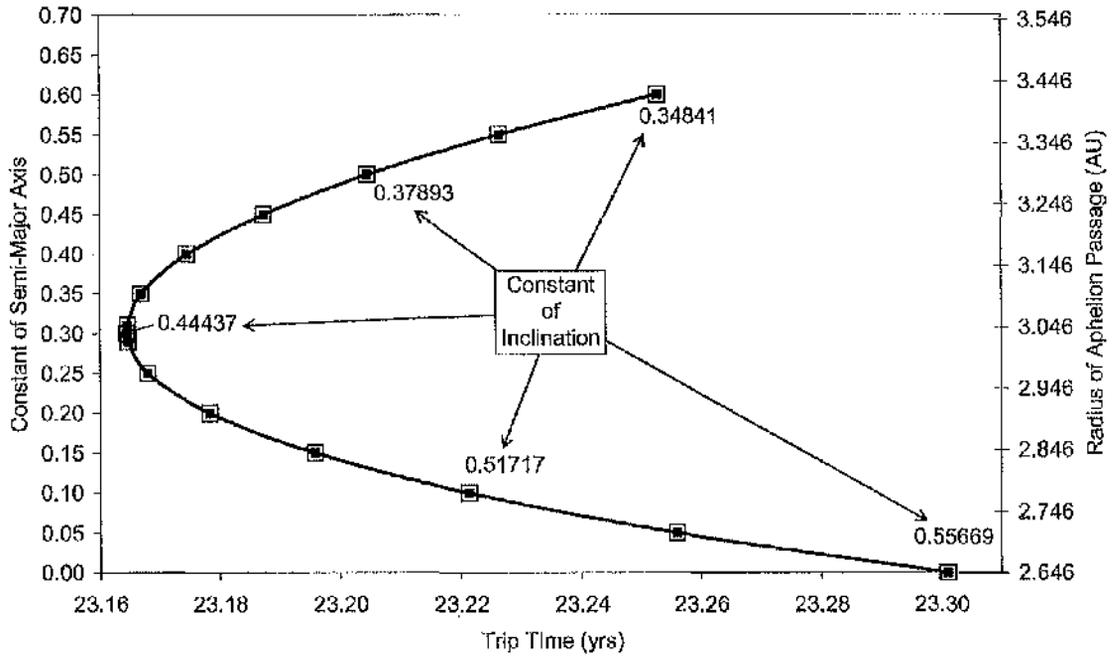
Figure 78 3D, open azimuth, trip time to 200 AU against radius of aphelion passage, and constant of semi-major axis controller in phase 1, with constant of inclination also shown.

The control angles for the minimum time trajectory from Figure 78 are shown in Figure 79. Note that the cone angle is no longer equal to the pitch angle as this scenario was calculated using Model 4. Once again it is noticed that the pitch angle plot is very smooth, with only a small discontinuity of 8 deg at the end of phase one, giving a maximum pitch angle of less than 67 deg. The very small size of the discontinuity shows how little an effect the minimum radius has on this trajectory. No cruise phase was found to be required for this trajectory once again suggesting that such phases are sub-optimal; contradicting previous assumptions.[69, 114, 126, 127, 201] Note that no attempt to verify the optimality of this trajectory was made using SQP methods or InTrance.

6.2.3   A Dual-Loop Scenario

Having established that the A^nD blending method is capable of generating trajectories with an ideal and non-ideal sail force model the characteristic acceleration is now reduced to 1 mm s^{-2}, allowing investigation of dual loop trajectories. It was found that low velocity solar system escape could be achieved with a single SPA with this level of sail performance. Thus, the first revolution of the trajectory must be used only to obtain a suitable position for the close solar pass which will provide a fast solar system escape. Note that it is exactly this scenario which previous attempts to use locally optimal methods to produce solar system escape trajectories have struggled with and typically produced very sub-optimal results.[198]

Launch was fixed at 03 January 2030 and no attempt was made to constrain the spacecraft azimuth at 200 AU as this was found in the ideal sail scenario above to alter trip time by less than 1.3 %; thus allowing
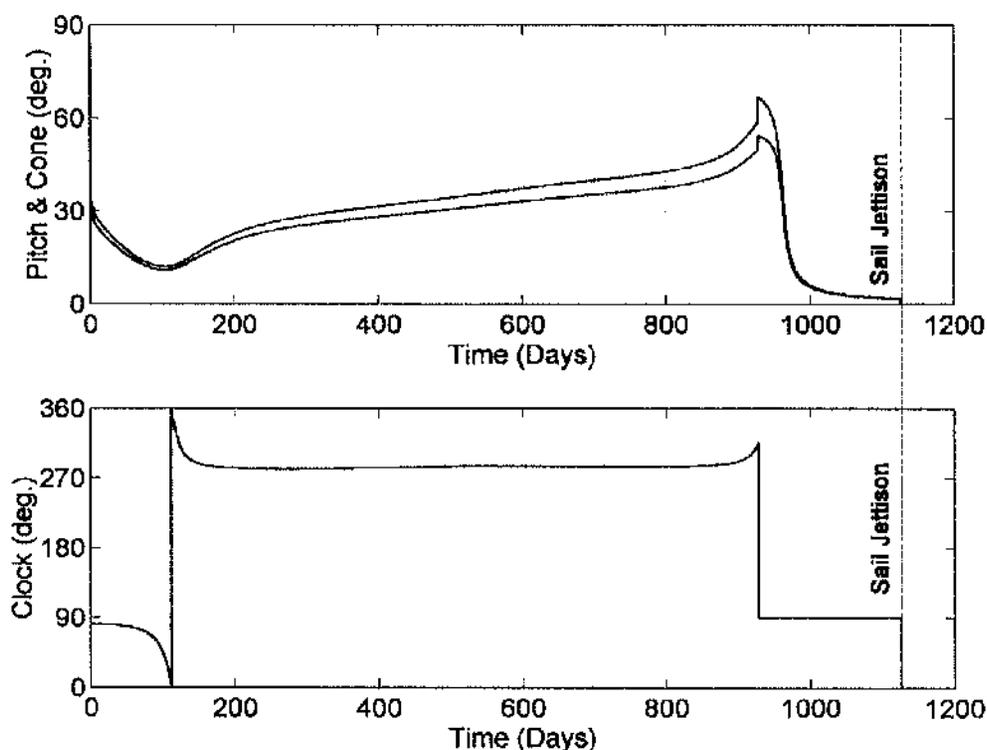
Figure 79   Control angles used for most favourable trajectory to 200 AU in Figure 78.

the near-optimal trajectory to be rapidly identified. Initially an ideal sail and a 2D trajectory are assumed, such that, for example, the launch vehicle has provided the required declination prior to separation of the payload stack and sail deployment. The first aphelion passage radius can be made large, however as the first perihelion cannot be much less than 0.5 AU (from Figure 70) this results in an even larger second aphelion passage if escape can be avoided and a long trip time until the close solar approach. This is the problem faced in previous attempts to use locally optimal methods to produce solar system escape trajectories.[198] The first aphelion passage is thus minimised by applying the locally optimal perihelion reduction control law, with passage occurring in 100 days at radius 1.05 AU. At this point the instantaneous perihelion is 0.65 AU. The second phase of the trajectory targets the second aphelion passage radius, much as during the first phase of the single revolution trajectories previously discussed. A third phase is thus used as the energy boost phase during the second perihelion passage, where once again only the semi-major axis controller is used. The second phase of the trajectory is thus the key phase, with the constant of eccentricity fixed at one and the constant of pericentre increased to reduce the second aphelion passage radius, or the constant of semi-major axis increased to increase it. Note that even if the instantaneous radius of perihelion is below the minimum allowed value the control law continues to reduce it as the third phase of the trajectory ensures the actual radius never drops below the minimum allowed value. A trade of radius of second aphelion passage versus time to 200 AU and velocity at 5 AU is seen in Figure 80. Using the spacecraft velocity at 5 AU it can be shown that making the assumption of constant velocity between 5 AU and 200 AU is a significant simplification which typically underestimates the total trip time by as much as 2.5 years, or up to 10 %.
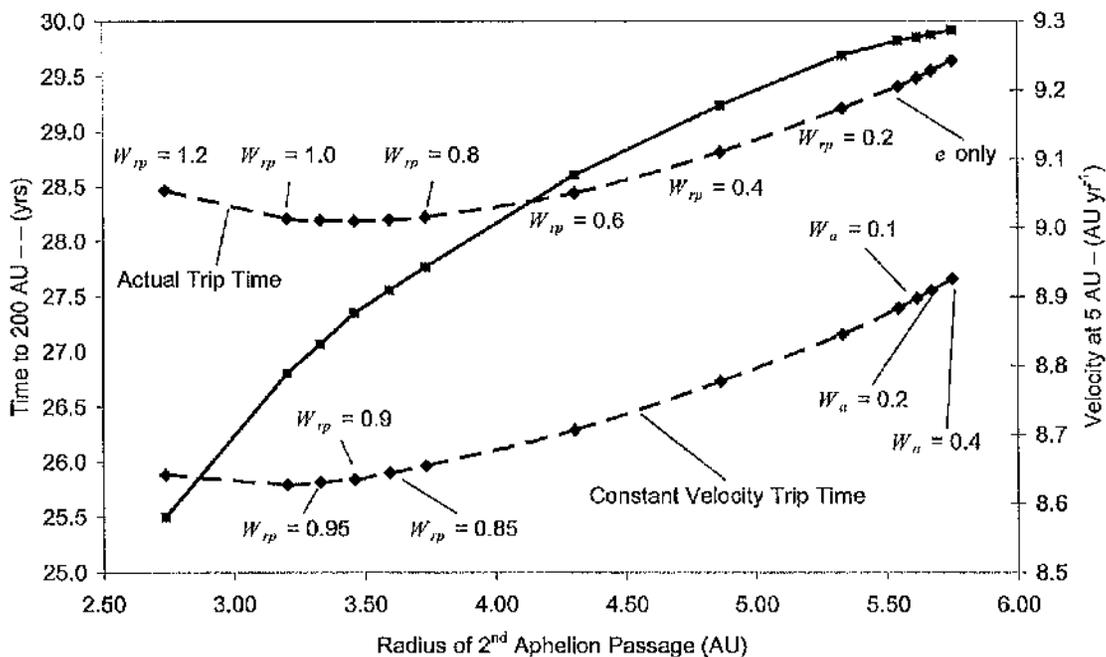
Figure 80 2D open azimuth trajectory trip time to 200 AU (– –) against radius of 2nd aphelion passage, with velocity of spacecraft at 5 AU (—) and constants used shown.

Furthermore, it is found that the most favourable radius of the second aphelion passage is different if this assumption is made. Thus, such an approximation is invalid at all times. Figure 80 shows that assuming constant velocity from 5 AU gives a best-case 200 AU trip time of 25.7 years, for an ideal sail with characteristic acceleration 1 mm s^{-2} and a thermal limit of 0.25 AU; the trip time is truly 28.2 years.

To quantify the optimality of the A^{n}D blending method for this 2D dual loop trajectory consider a 2D trajectory propagation from perihelion, at 0.25 AU to 200 AU using the locally optimal energy gain control law to a radius of 5 AU, where the sail is then jettisoned. It is found that if eccentricity is equal to one at 0.25 AU the time from sail jettison, at 5 AU to 200 AU is approximately 23.4 years, with a velocity of 9.2 AU yr^{-1} at 5 AU. Notice however in Figure 80 that the velocity at 5 AU peaks at almost 9.3 AU yr^{-1}. It is found in fact that the sail has already passed through the point of escape prior to the SPA at 0.25 AU for trajectories with velocity at 5 AU above 9.2 AU yr^{-1}. The trajectory represented by the right most point in Figure 80 has an eccentricity of 1.016 at closest solar approach. It was found that an eccentricity of 1.105 was required to reduce the trip time from sail jettison to 200 AU to just over 21 years, with a velocity at 5 AU of almost 10 AU yr^{-1}. At least such a trajectory would be needed to match the requirements defined for the IHP TRS using an ideal sail with characteristic acceleration 1 mm s^{-2} and a minimum radius of 0.25 AU. Thus, a trip time of 25 years with this level of performance seems impossible as the time required to position such a sail prior to the SPA would be much greater than 4 years. Note in Figure 80 the time to sail jettison varies from 3.15 to 6.66 years, increasing in line with the radius of the second aphelion passage. Using the 0.25 AU to 200 AU trajectory propagation findings, one concludes that the best-case trip time found, at 28.2 years, is likely near-optimal.

Introducing inclination change to the dual loop trajectories significantly increases the complexity of the analysis as it is difficult to logically quantify when it is optimal to have inclination attain its target value, such as to allow concentration on energy gain. Generally, it was found that the inclination need not have attained its target value prior to the beginning of the final phase, the locally optimal energy gain phase, though it should have prior to the second SPA. The first of the three phases once again uses the locally optimal pericentre reduction control law. However, it now also uses the inclination increase control law. It was found that by setting the constant of pericentre to one and varying the inclination constant the time to aphelion passage was minimised for constant of inclination equal to 0.23, for both an ideal and non-ideal sail. Similar to the 2D trajectory, the second phase begins at the first aphelion passage, at radius 1.05 AU and inclination 1.19 deg after 97.4 days, for an ideal sail and radius 1.06 AU, inclination 1.14 deg and 106.4 days for a non-ideal sail. The first phase is fixed independent of the minimum thermal bound, which will be varied to quantify its effect for an ideal and non-ideal sail. It was found that the best way to optimise the second phase was to set the semi-major axis or pericentre constant as fixed and then allow the inclination constant to vary such that following the third phase the time to 200 AU was minimised for that value of semi-major axis or pericentre constant. Following this the constant of semi-major axis or pericentre was varied slightly and then the process of inclination constant refinement repeated to find a new minimum trip time. The constant of semi-major axis or pericentre was altered depending on the value of the new minimum found, until the best-case trip time was determined for each minimum radius considered. Once again, the eccentricity constant was set to 1 during phase two. Such a process is more time consuming than the single loop trajectories previously presented, however the process remains very rapid. For example, the non-ideal scenario in Figure 78 and Figure 79 took approximately 2½ hours to produce all the data points presented, however the actual minimum time trajectory was found in less than 1 hour. By contrast the non-ideal scenario in Figure 81 and Figure 82 for thermal limit 0.25 AU required several hours to find. Recall that these trajectories require the user to manually alter the constants applied to the A^nD score and if this process was automated the time required to generate a solution would be reduced by perhaps an order of magnitude or more. Trip time to 200 AU with an ideal and non-ideal sail of characteristic acceleration 1 mm s^{-2} is shown in Figure 81, along with the corresponding velocity at 5 AU for each of the most favourable trajectories found. It is seen from Figure 81 that the most favourable 0.25 AU limited trajectory is only 11 days longer than the most favourable 2D trajectory found in Figure 80. Note also that a non-ideal sail typically adds in the region of 5 % to the ideal sail trip time, as was found in Reference 137. A minimum solar radius of less than 0.2 AU is required to reduce the non-ideal sail trip time to less than 25 years, as required for the IHP TRS. Figure 82 shows the radius of the 2nd aphelion passage corresponding to the trajectories shown in Figure 81, along with the constants used for each during phase two, noting that in addition to the constants given in Figure 82 the constant of eccentricity is fixed at 1. Similar to the 2D trajectories, if the instantaneous radius of perihelion is below the minimum allowed value the controller continues to reduce it, with the third phase of the trajectory ensuring the actual radius never drops below the minimum allowed value. Note that the gradient of the non-ideal curve in Figure 82 for the A^nD results is less than that of the ideal sail, while the InTrance results also show a crossover. The minimum time trajectory for a 0.25 AU thermally limited non-ideal sail is shown in Figure 83, where once again it is noted that this is an open azimuth trajectory. The corresponding sail control angles are seen in Figure 84, where one notes that the pitch angle plot is relatively smooth except for a significant discontinuity at the end of the second phase,

where the pitch angle jumps by almost 30 deg to 71 deg, the highest pitch angle required during the trajectory. No cruise phase was found to be required for this trajectory, again suggesting that such phases are sub-optimal. The weights generated using A^nD blending and the constants from Figure 82 are shown in Figure 85. Note the spike in inclination weight prior to the first perihelion passage (day 261) where the inclination rapidlly increases (see Figure 86), illustrating the opportunistic nature of A^nD blending.

Finally, Figure 86 confirms that the inclination has once again reached its final value prior to the second SPA and that the first perihelion passage occurs at 0.5 AU, which corresponds well with the estimate given in Figure 70. Though not shown in Figure 81 the minimum trip times generated using SQP with an ideal sail, at 0.20 AU and 0.25 AU thermal limit, were found to be 6 % and 3 % longer respectively than the exactly equivalent A^nD result. Figure 81 and Figure 82 show the near-equivalent results generated using InTrance. Note at this stage that the method of constraint definition used within InTrance does not allow for an elevation constraint to be set without an azimuth constraint. Rather, within InTrance the azimuth and elevation are constrained and the launch date remains open, where it was fixed within the A^nD analysis. This distinction is a small but bothersome factor as it means that the InTrance results in Figure 81 and Figure 82 cannot be stated as exactly equivalent, only near-equivalent. Recall from earlier that the change from a fixed launch date to open launch date resulted in an increase in trip time of approximately 1.3 %, using A^nD. Thus, so long as the InTrance solution is less than 1.5 % different from the A^nD solution it can be assumed the A^nD solution is near-optimal. It was found that the trip times in Figure 81 generated using InTrance are mostly slower than the near-equivalent A^nD solution. However, all are within 1 % except for the non-ideal sail scenario with minimum radius 0.20 AU, which is 1.21 years or 4.4 % slower. No explanation for this variation was found or is offered, however it is clear that the A^nD solution must be near-optimal.
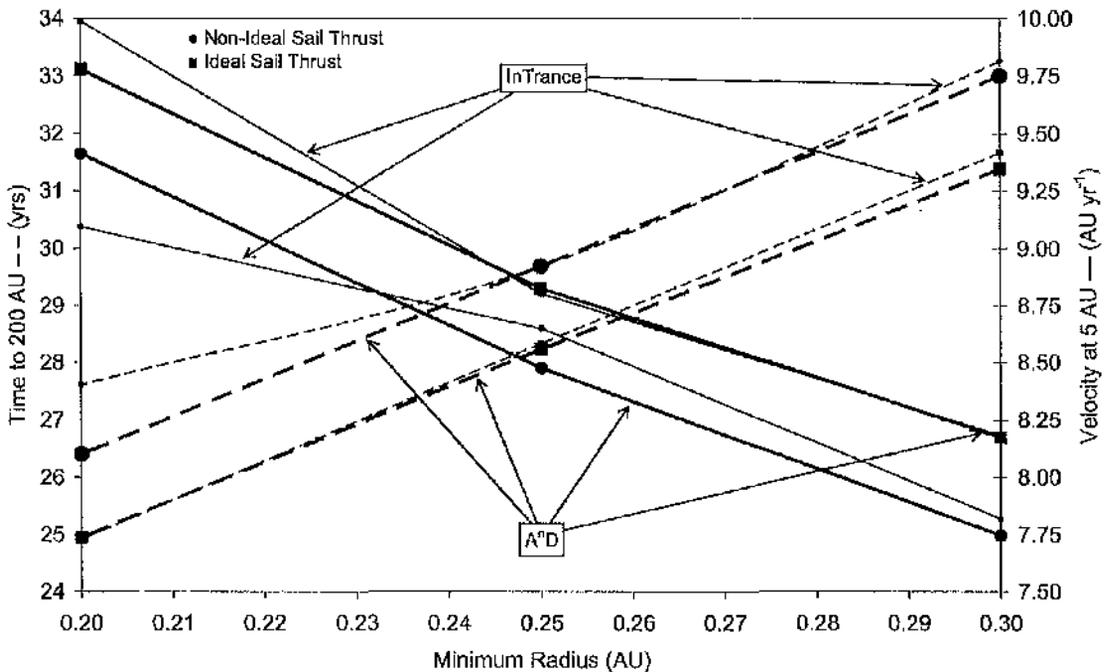
Figure 81   Trip time to 200 AU (– –) against minimum radius and velocity at 5 AU (—) for ideal and non-ideal sails with characteristic acceleration 1 mm s⁻².
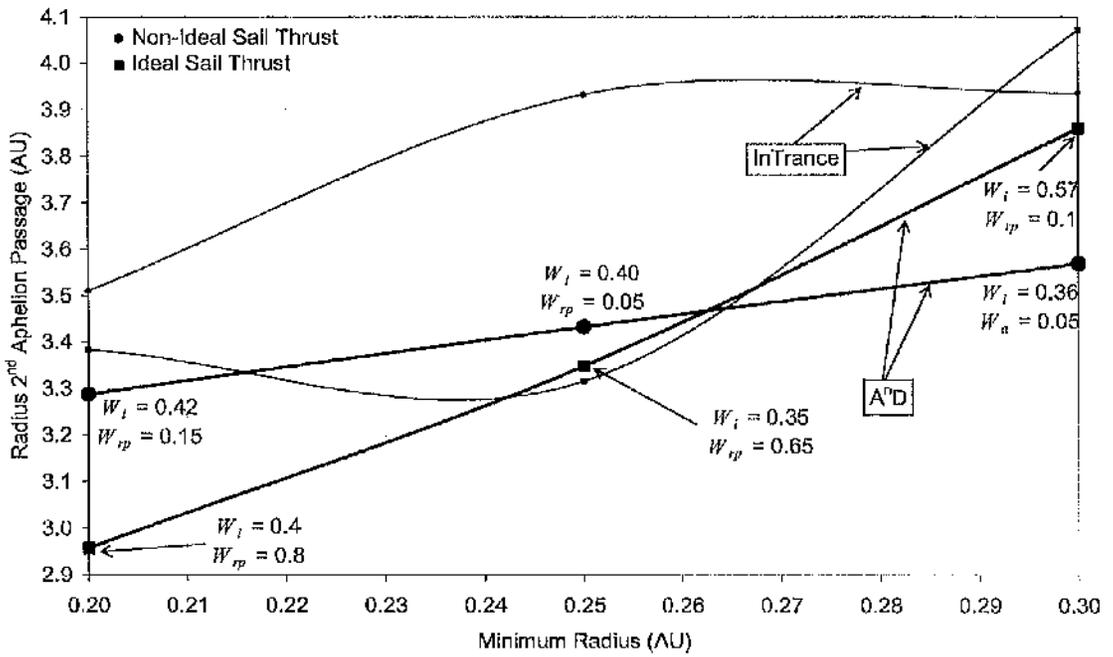
Figure 82 Radius of 2nd aphelion passage against minimum radius for ideal and non-ideal sails with characteristic acceleration 1 mm s^{-2}. $W_e = 1$ in phase 2, other constants used in phase 2 shown.
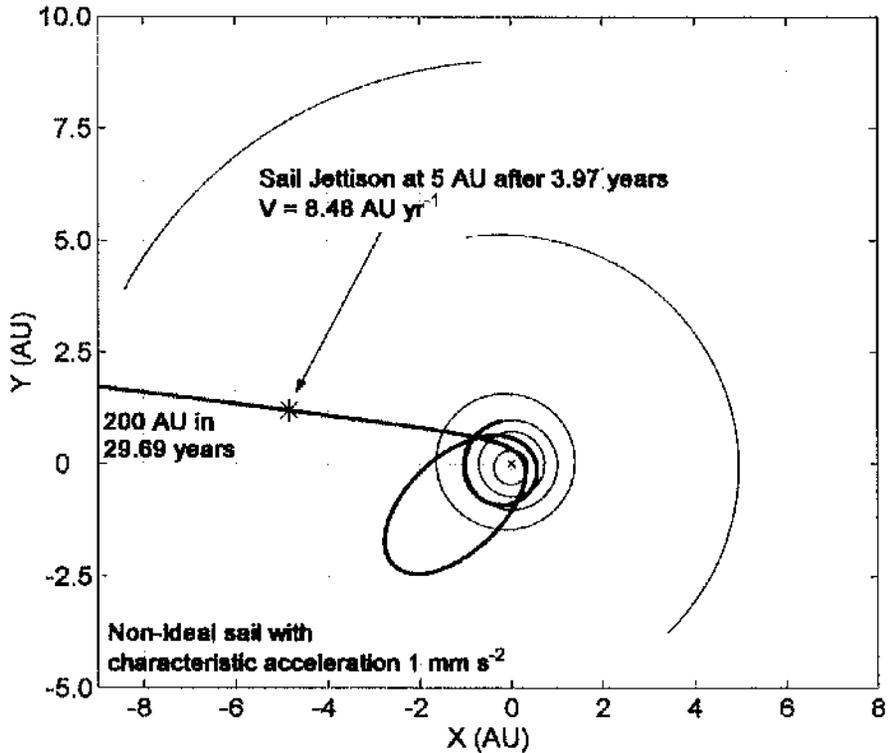
Figure 83 Open azimuth trajectory to 200 AU with thermal limit 0.25 AU.
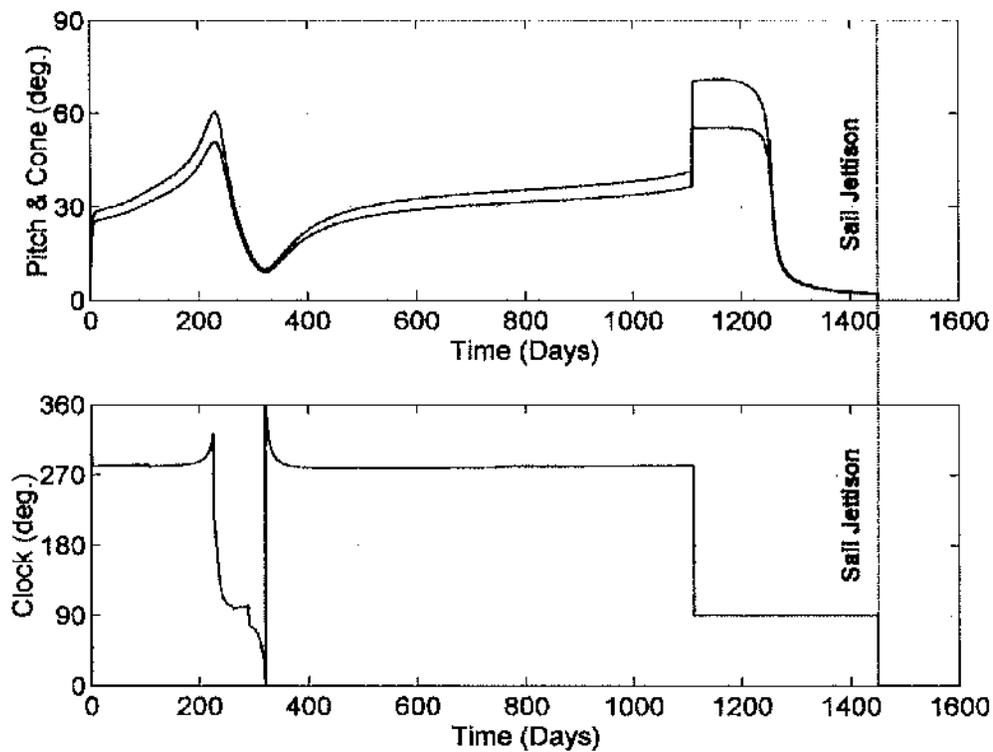
Figure 84 Control angles used in Figure 83.
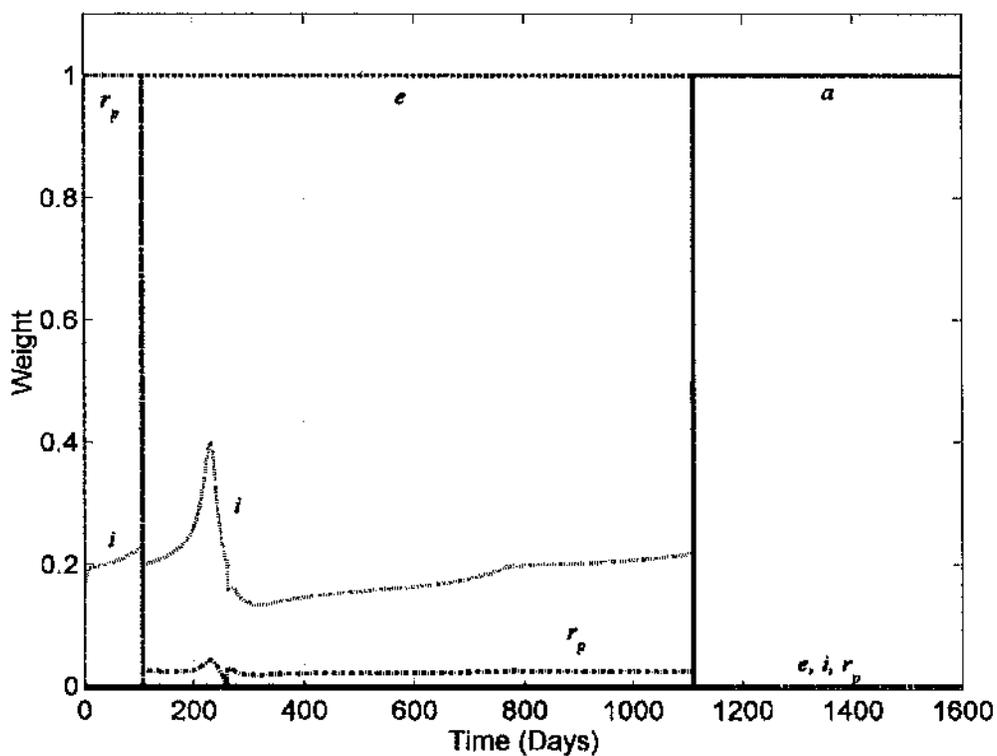
Figure 85 Weights used in Figure 83 and Figure 84.
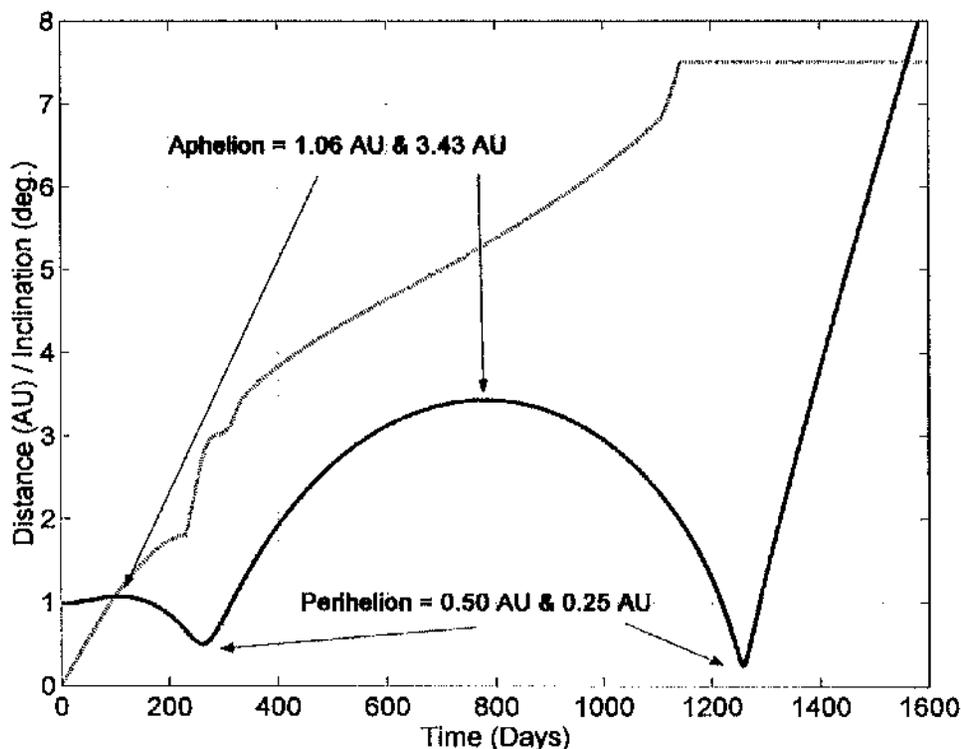
Figure 86 Radius (——) and inclination (···) of trajectory plot in Figure 83.

6.2.4   Temperature Limited Trajectories

Trajectories to 200 AU which use the sail film temperature to constrain the trajectory rather than a minimum radius are considered. This approach was first employed by Koblik, et al for the generation of solar sail transfer trajectories to near-Sun orbits.[203 - 205] This approach is in some respects more realistic since the sail temperature is a function of the solar aspect angle, however it does assume that the sail booms and other systems can survive such a close solar approach. Note that this technique could in-principle be used with a more advanced set of temperature constraints based a complete thermal model of the spacecraft. Three scenarios were recently considered using InTrance in Reference 137, with sail film temperature limited to 200, 240 and 280 deg Celsius; these results are reproduced in Figure 87. In Reference 137 the sail is not jettisoned at 5 AU; rather it is used all the way to 200 AU. Furthermore, the trajectories have an unconstrained azimuth and elevation at 200 AU. Once again these variations make direct comparison difficult between A^{n}D and InTrance. However, it is noted that an open azimuth/elevation InTrance trajectory with a characteristic acceleration of 1.5 mm s^{-2} and using an ideal sail all the way to 200 AU has a trip time of 22.248 years, for a thermal limit of 0.25 AU. Approximately 2.9 % faster than the equivalent constrained start date, elevation, open azimuth trajectory generated using A^{n}D. It is of interest to note that the completely unconstrained InTrance solution has final inclination 12.93 deg and does not remain within the ecliptic plane as one may expect. One concludes that if the temperature limited trajectories are within 3 % of the InTrance trip times they are near-optimal. It is found that the 200 °C temperature limit A^{n}D trajectory is actually very marginally faster (0.04 %) than the InTrance trajectory, while the A^{n}D trajectory is marginally slower (0.25 %) than InTrance trajectory for the 240 °C temperature limited scenario. The A^{n}D blending 280 °C

limited result is 1.26 % slower than the InTrance result, however this is significantly within the error margin discussed above. Figure 87 shows the sail's escape velocity at 200 AU $\left(v_e = \sqrt{v^2 - 2\mu/r} \right)$, where it is seen that both A"D and InTrance provide very similar solutions.

6.2.5   Trajectories with Optical Surface Degradation

Having demonstrated the ability of the A"D blending method to cope with a non-ideal sail force model and produce near-optimal results for both single and dual loop trajectories, the effects of optical surface degradation, as discussed in Section 1.3.4, are now considered using Model 7. A sail characteristic acceleration of 1.75 mm s^{-2} was selected for this analysis, which gives a zero degradation trip time to 200 AU of 21.74 years. Note that the trajectories within this section have a fixed minimum radius of 0.25 AU, a Half-Life Solar Radiation Dose of 0.5 W$_c$ yr and a fixed start epoch of 03 January 2030. The trajectories within this section have not been validated using either InTrance or SQP methods, however extrapolating the results from prior sections one can assume the results to be very near-optimal.

Figure 88 shows the trip time to 200 AU for $d = 0$ to 0.30. Note that up to $d = 0.2$ the trip time increase due to degradation is exactly linear. At $d = 0.25$ and 0.3 the trip time does not fit the expected linear trip time relationship due to the sail continuing to increase orbit inclination beyond the point where effort would be better spent gaining orbit energy. Furthermore, at $d = 0.30$ the sail is unable to reach the minimum radius boundary and thus unable to fully take advantage of the close solar pass. It is thus clear that increasing the
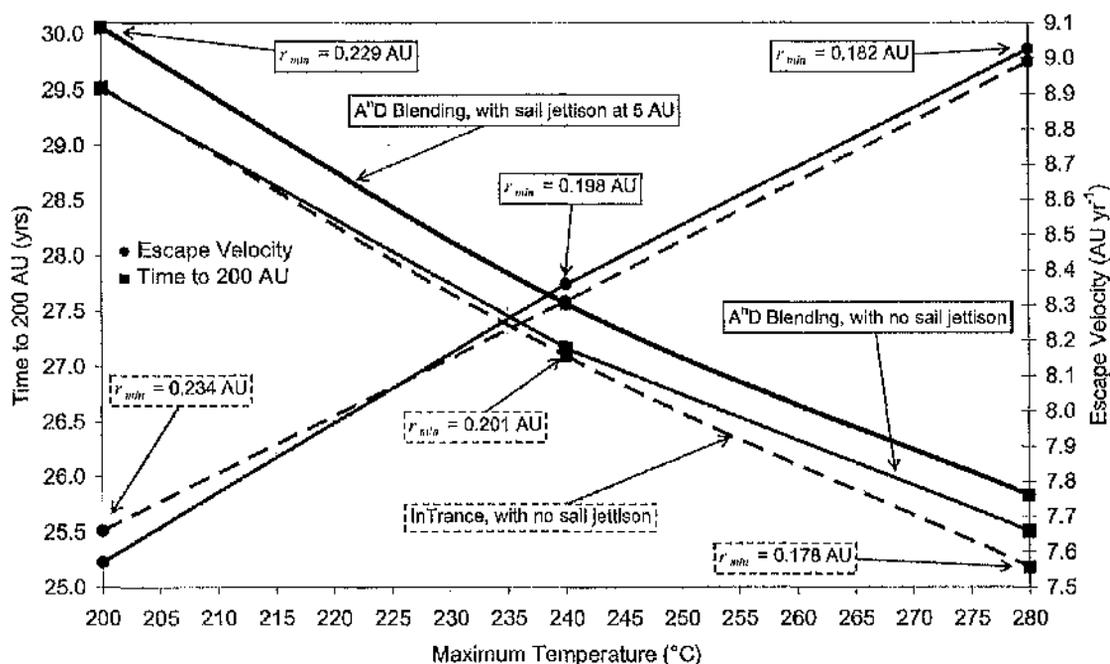

Figure 87   Trip time to 200 AU as a function of maximum sail film temperature with characteristic acceleration 1.0 mm s^{-2}. A"D (—) and InTrance (— —) solutions shown.

degradation limit results in a change in the structure of the trajectory such that the minimum radius begins to increase from 0.25 AU to the point where it will become necessary for the sail to perform a dual loop trajectory rather than a single loop. The value of d at which this trend begins will vary according to the characteristic acceleration; at 1.75 mm s^{-2} the critical value of d is 0.3, while at 1.6 mm s^{-2} the critical value is of d is 0.05. Figure 89 shows the radius of aphelion passage increases by approximately 1 AU as the degradation limit is increased from zero to 30 %. This represents a significant change in trajectory structure and hence a significantly different set of design requirements for both the telecommunications and thermal systems on the spacecraft while attached to the sail, requiring carefully consideration during the initial spacecraft design phases. Note the radius of aphelion passage gradient is at no time linear, although it is approximately constant. Figure 90 shows the velocity of the spacecraft at sail jettison is decreasing in an approximately linear manner, however the time to sail jettison much more closely matches the shape of the radius of aphelion passage relationship. It is thus notable that given all these relationships are only approximately linear that the trip time to 200 AU is exactly linear for $0 \le d \le 0.2$. Figure 91 shows the value of optical coefficients at sail jettison and the total SRD variation with degradation limit. Note from Figure 91 that the total SRD for $d = 0.3$ is not as large as expected due to the increased value of the minimum solar approach radius yet the degradation of coefficients remains almost constant due to the increased trip time to 5 AU.

Figure 88 – Figure 91 quantify the general trends associated with optical surface degradation; these can be further examined however by analysis of the orbit parameters throughout each best-case trajectory found. Figure 92 shows the most favourable trajectory plot for each degradation limit. Note that a change in degradation limit for a fixed start epoch significantly changes the final spacecraft azimuth at 200 AU which would significantly impact mission science objectives. In Figure 92 it is seen that the SPA occurs at the same physical location independent of the degradation limit as the argument of pericentre is not varied from trajectory to trajectory. However, the time of each SPA is seen to vary by as much as 1¼ years in Figure 93. It is also seen in Figure 93 that the inclination has attained its final value significantly before the SPA for $0 \le d \le 0.2$. At $d = 0.25$ the inclination reaches 7.5 deg just prior to the SPA, while at $d = 0.3$ this does not occur until after the SPA.

Due to the low pointing accuracy of a solar sail it was identified previously that an X-band, or lower, low or medium gain antenna is required for Earth – spacecraft communications when attached to the sail.[111] Thus, the maximum slant range of this telecommunications system is a key mission driver. Figure 94 shows the slant range of the spacecraft up to sail jettison. It is seen that the slant range during the second aphelion passage is altered by as much as 1.5 AU. It is also seen that between $d = 0.1$ and $d = 0.2$ the slant range of the second aphelion passage becomes larger than during the first aphelion passage. The slant range at sail jettison is seen in Figure 94 to be maximum for the $d = 0$ case and to vary by as much as 1.75 AU with sail degradation.

The sail control angles used in each best case trajectory are illustrated in Figure 95 where it is noted that the maximum pitch angle of each trajectory is similar, yet the maximum cone angle decrease in-line with the degradation limit increase. Furthermore, the size of the discontinuity within each trajectories control angle

profile in Figure 95 decreases as d increases, until at $d = 0.3$ there is no discontinuity as the trajectory is no longer limited by the minimum radius value. Thus, the size of the discontinuity is directly related to the impact the minimum radius setting has on the trajectory. Figure 96 shows the variation of the optical coefficients over the first 4 years of trajectories shown Figure 92, where it is seen that the bulk of the degradation occurs during the close solar pass. It is thus logical to assume that multiple close solar passes would have an adverse effect on the quality of the optical surface and should be avoided when designing such trajectories. Figure 97 further underlines the effect of each SPA on the optical surfaces by clearly showing that the bulk of the Total SRD occurs during the SPA.
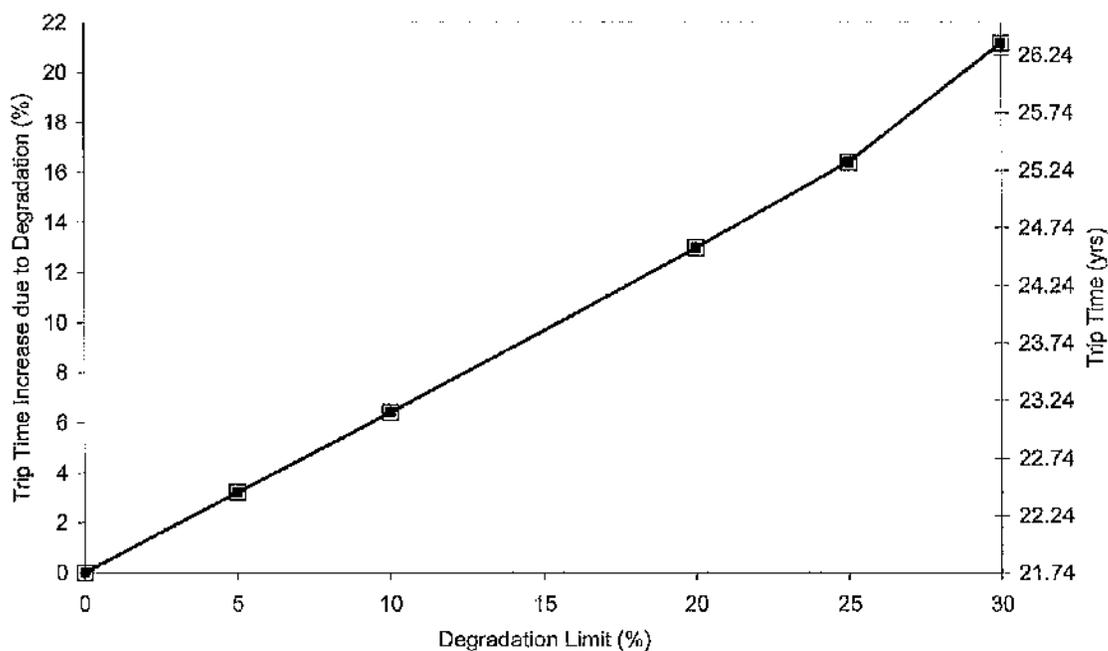
Figure 88 Trip time to 200 AU for a range of degradation limits.
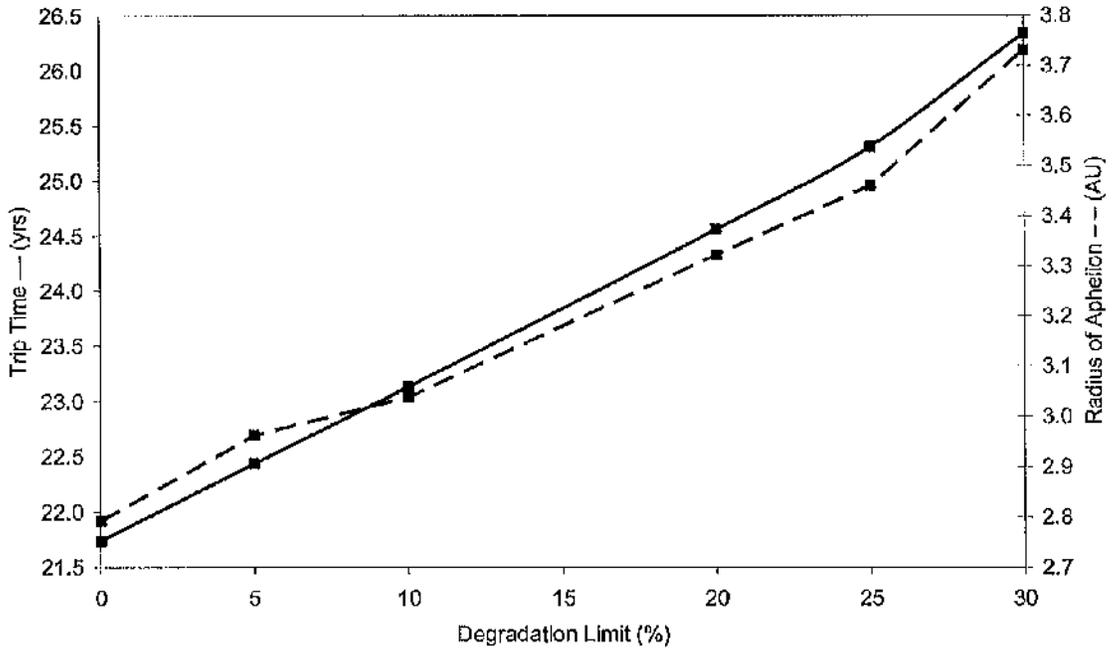
Figure 89 Trip time to 200 AU (——) and radius of aphelion passage (— —) for a range of degradation limits.
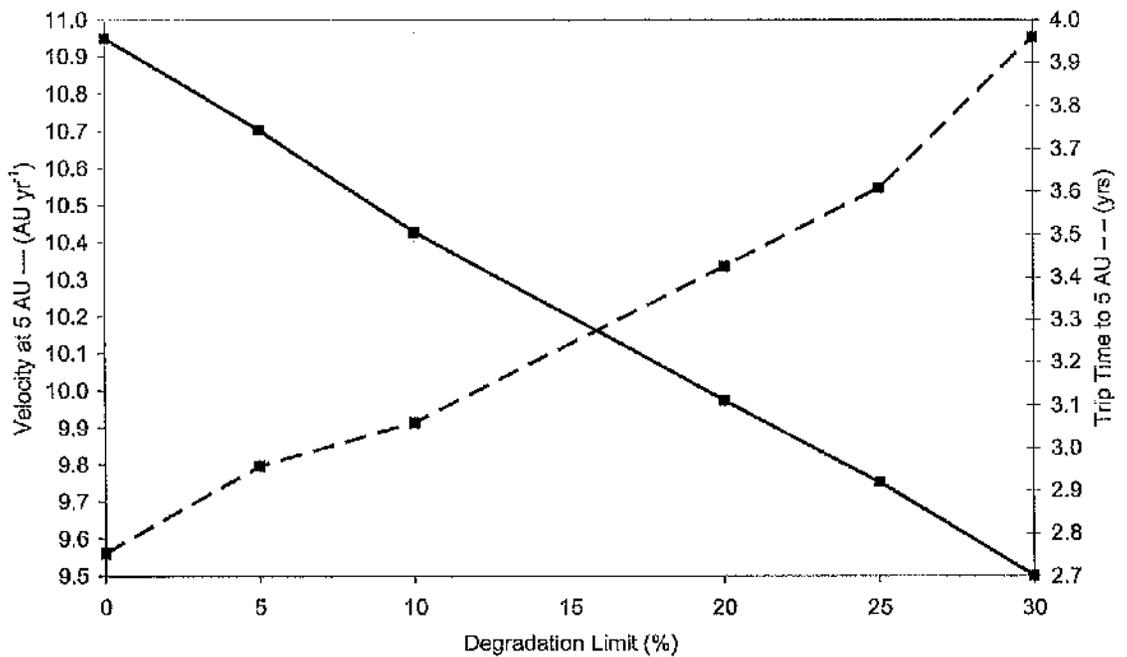
Figure 90 Velocity at 5 AU (——) and trip time to 5 AU (— —), the sail jettison point, for a range of degradation limits.
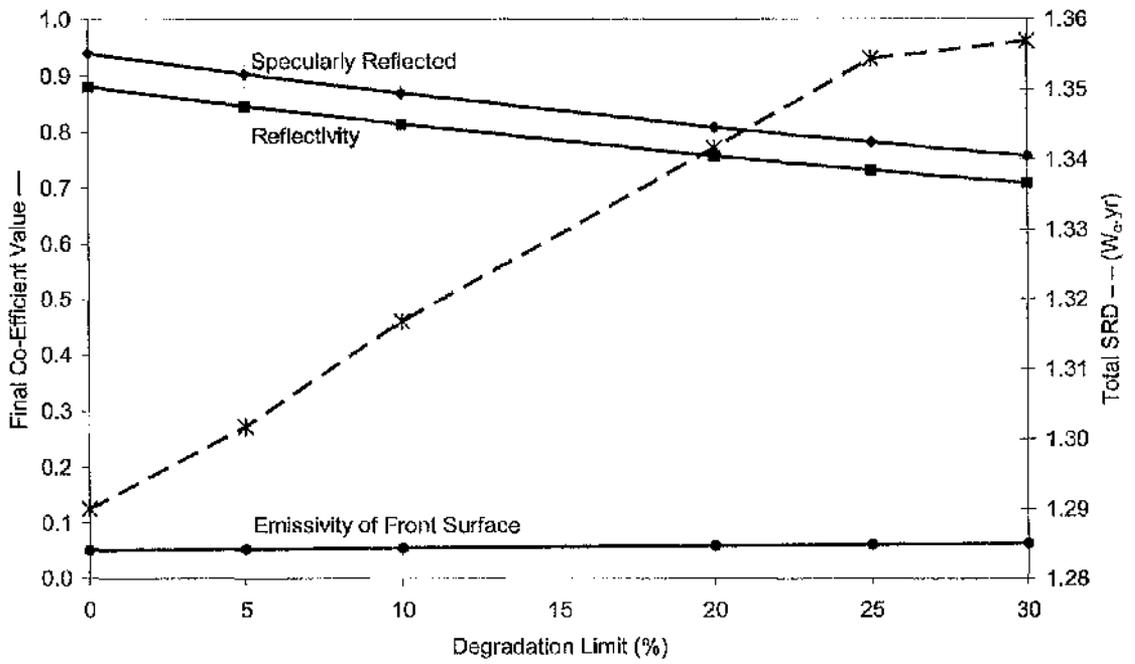
Figure 91 Final value of optical coefficients (——) and total SRD (— —) for a range of degradation limits.
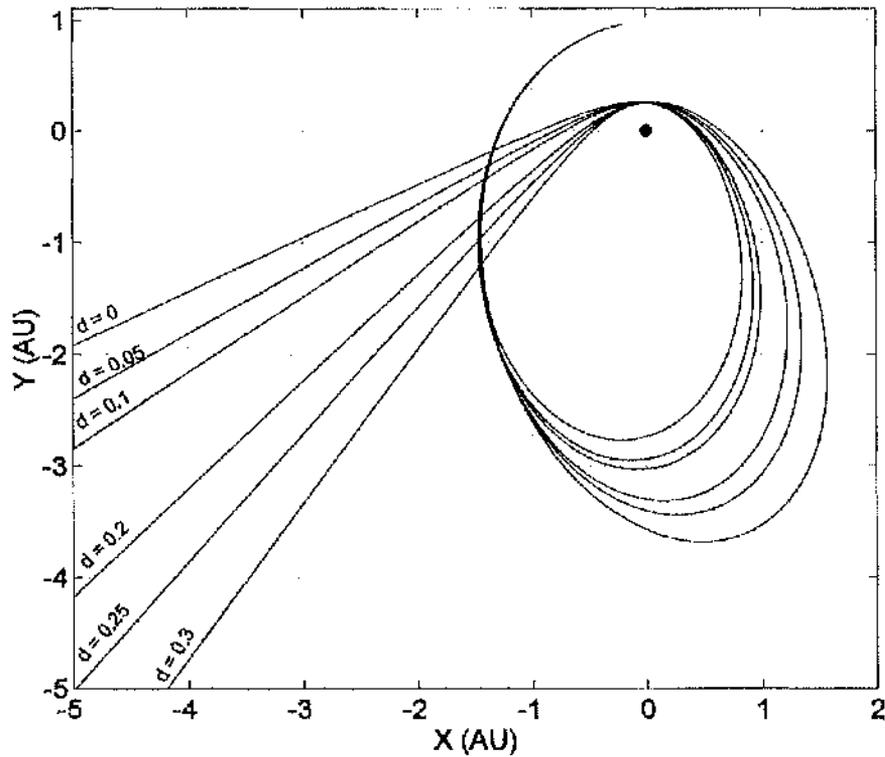
Figure 92 Inner solar system trajectory plots, start epoch 03 January 2030, for a range of degradation limits.
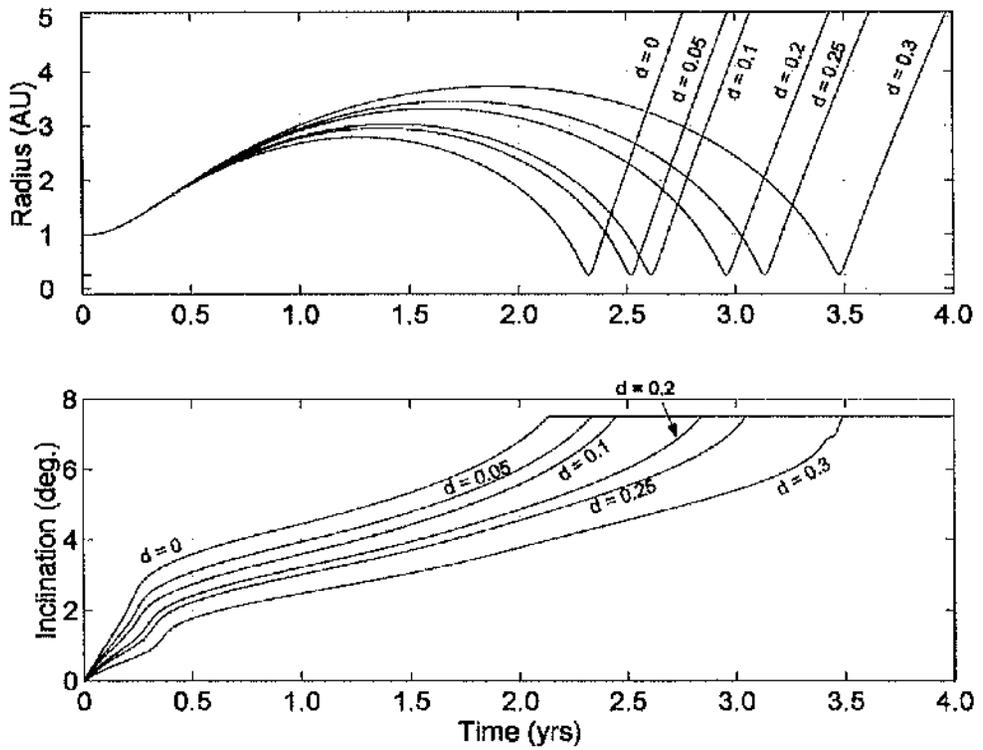
Figure 93 Radius and inclination variation over first 4 years of trajectories shown in Figure 92.
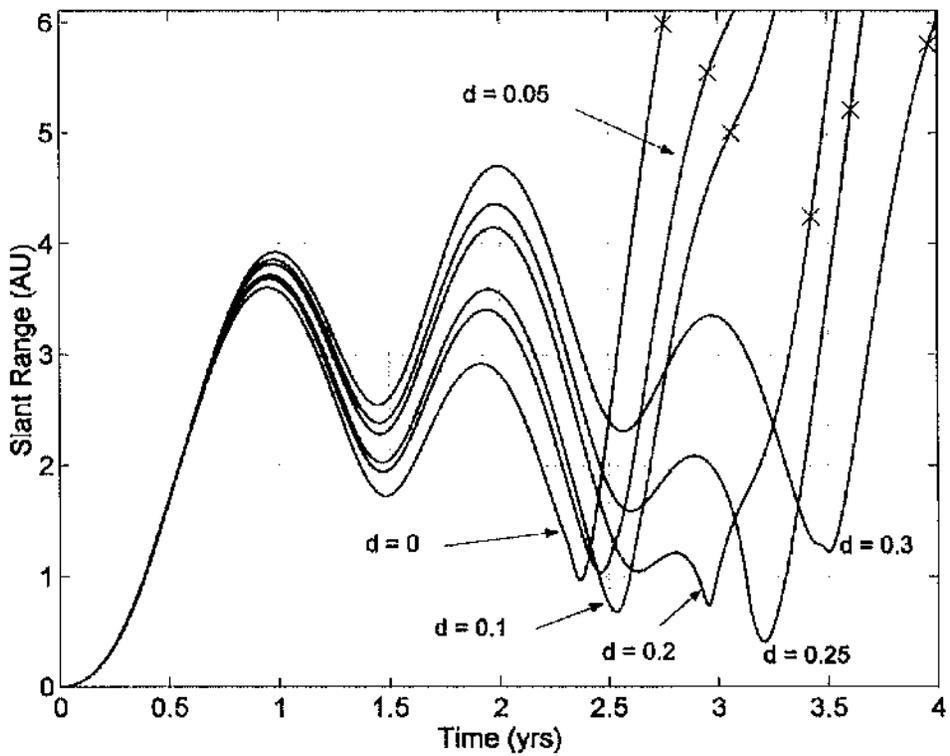
Figure 94 Slant range over first 4 years of trajectories shown in Figure 92, with sail jettison shown (×).
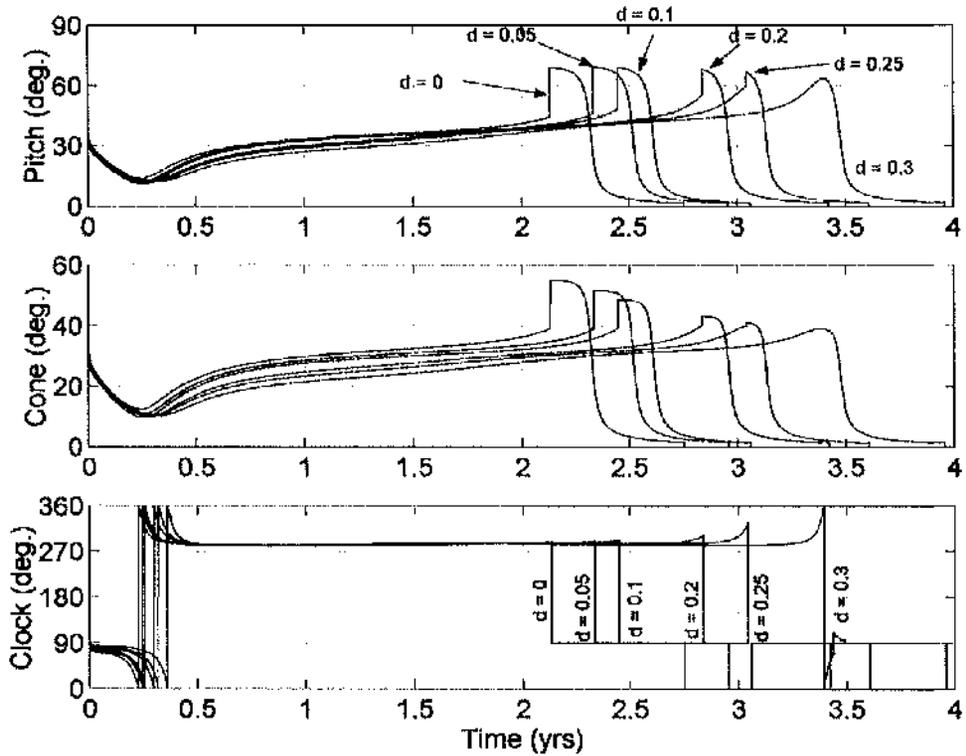
Figure 95 Control angles used for each trajectory shown in Figure 92.
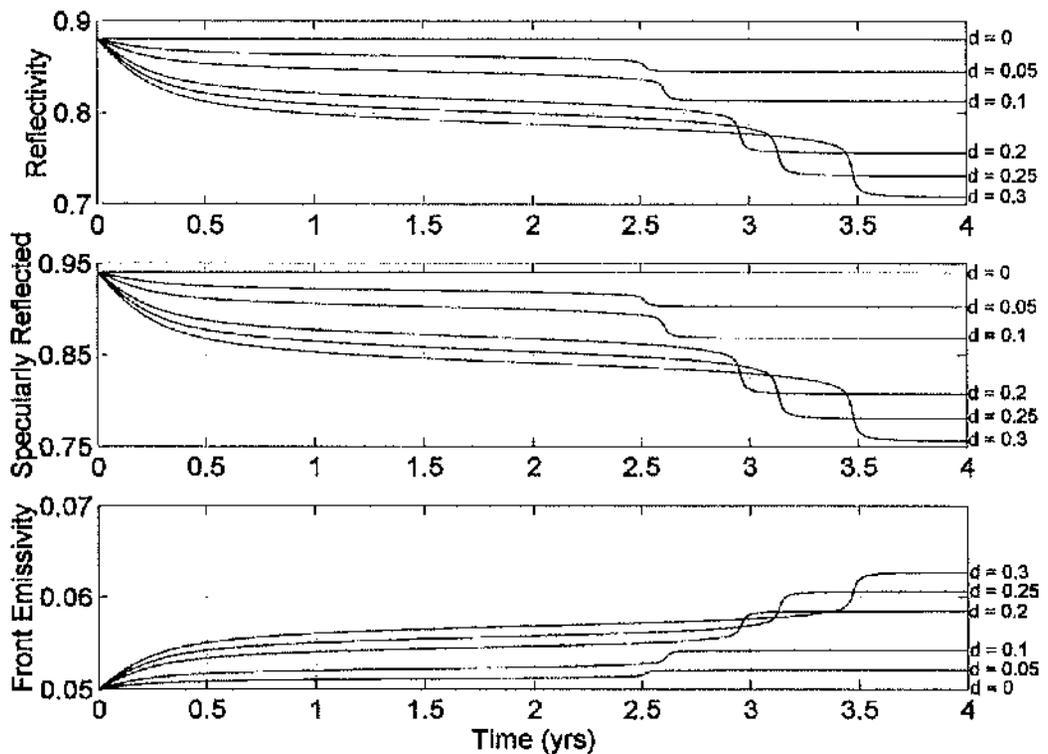
Figure 96 Variation of optical coefficients over first 4 years of trajectories shown in Figure 92.
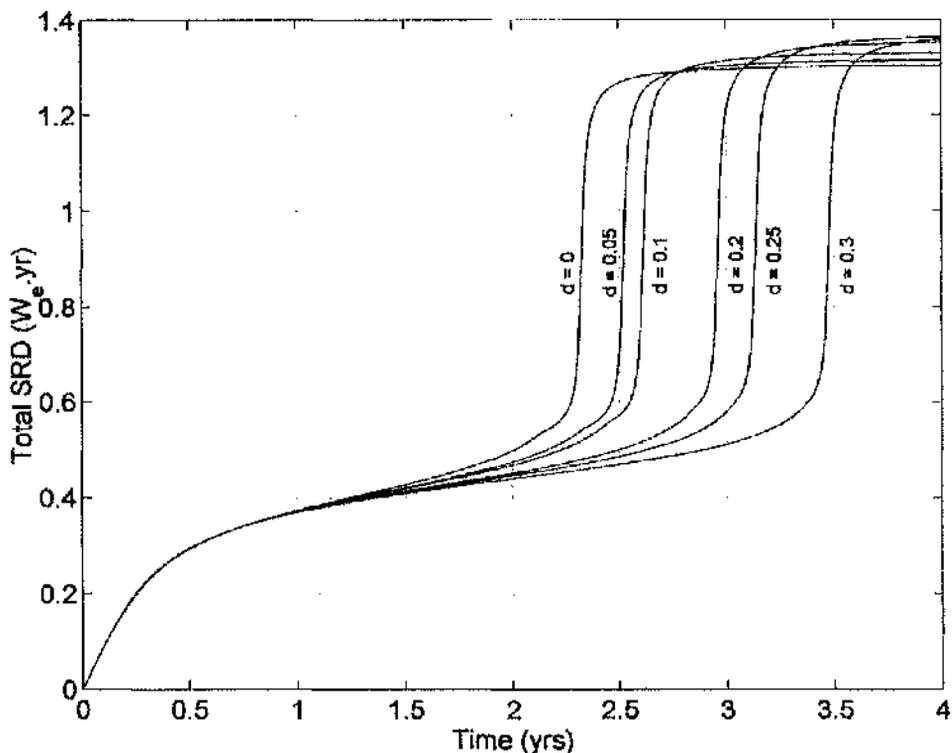
Figure 97 Variation of total SRD over first 4 years of trajectories shown in Figure 92.

6.3   Earth – Mercury Transfer

The generation of an Earth – Mercury transfer trajectory allows an excellent case study for A^{n}D blending, with ample trajectories available within the literature for comparison. Due to the proximity of Mercury to the Sun the required sail characteristic acceleration is much less than that used for 200 AU missions. However, as a result of the low acceleration an Earth – Mercury transfer becomes a multi-revolution transfer which increases the computational difficulty, especially for traditional optimisation techniques. Maintaining consistency with Section 5.3 a characteristic acceleration of 0.25 mm s^{-2} is assumed for the Earth – Mercury transfer, using Model 1. It was found that using A^{n}D blending the transfer should be split into two phases, the first concentrating on lowering the semi-major axis, the second on increasing orbit eccentricity and inclination to match those of Mercury. The first phase is 753.3 days in duration, with the second requiring a further 298.3 days. The constants used on the A^{n}D scores are detailed in Table 10 and illustrated in Figure 98, where it is noted that during the first phase despite inclination having the lowest constant in Table 10 at the crossing of the minor-axis it dominates the eccentricity weight. It is also of interest to note that the plot of weight values is now much more complex than Figure 76 and Figure 85 which illustrate the trajectories to 200 AU. Note that if an element is not listed in Table 10 then the weight is set to zero and the element is not considered within the trajectory design. Using the constants given in Table 10 the resultant orbit radius, semi-major axis, perihelion and aphelion are plotted in Figure 99, where it is noted that the semi-major axis rate of change is approximately linear throughout the 1051.6 day transfer and that the semi-major axis and eccentricity obtain good convergence with the target values of Mercury. Note this is not a rendezvous trajectory, rather an orbit-to-orbit transfer which can be used to rapidly obtain the approximate minimum

rendezvous transfer time. A plot of the transfer is shown in Figure 100, where the change of semi-major axis, eccentricity and inclination is shown. From Figure 100 it is seen that the orbit inclination remains low until the semi-major axis has been reduced, thus allowing the inclination to be increased more rapidly due to the shortened orbit period. The sail control angles generated by the A^{n}D blending algorithms are shown in Figure 101, where the oscillatory nature of the control angle profile is seen. Figure 101 also shows the control angle history for a 51 node SQP generated Earth – Mercury transfer.[133 - 135] The SQP generated transfer duration is 1041 days, that is to say, 10 days or less than 1 % shorter than the transfer generated by A^{n}D blending. Note the SQP result was the best found in the literature and is thus used as the reference point. The SQP generated transfer is a Mercury rendezvous trajectory, note however that the control angles generated using A^{n}D blending are very similar to those generated using SQP and as such would provide an excellent initial guess towards such a local optimisation method.

Parameter	Constant in Phase 1	Constant in Phase 2
Eccentricity	0.290	0.455
Semi-Major Axis	1.000	0.000
Radius of Pericentre	0.000	0.865
Inclination	0.250	0.530

Table 10   Constants used during Earth – Mercury Transfer.
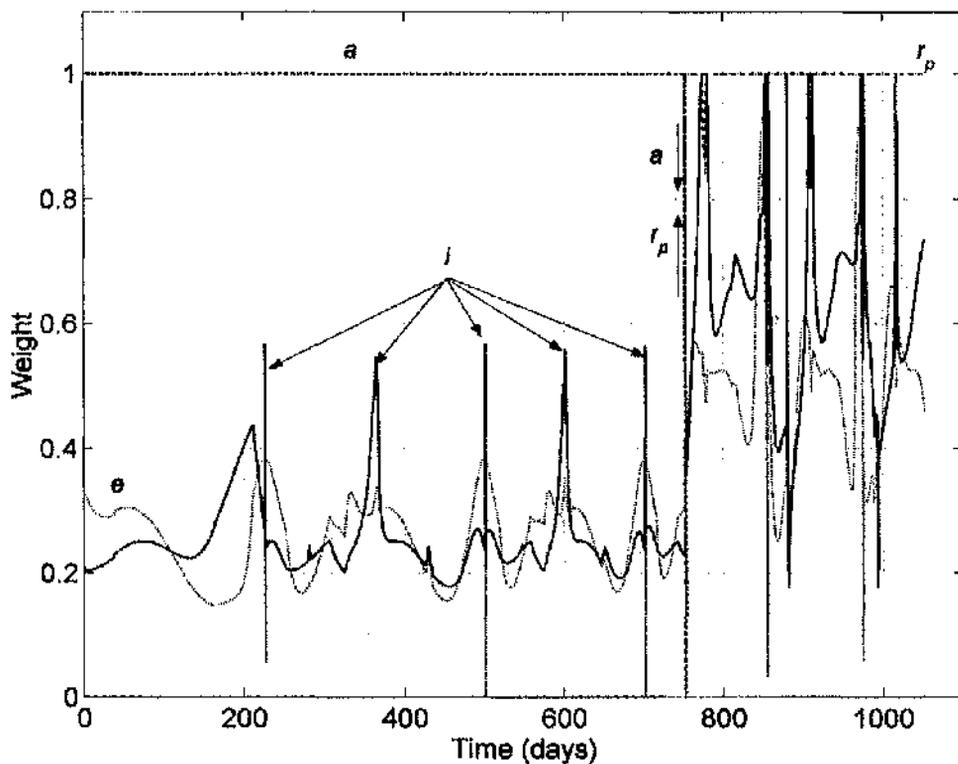
Figure 98   Weights used during Earth – Mercury Transfer.
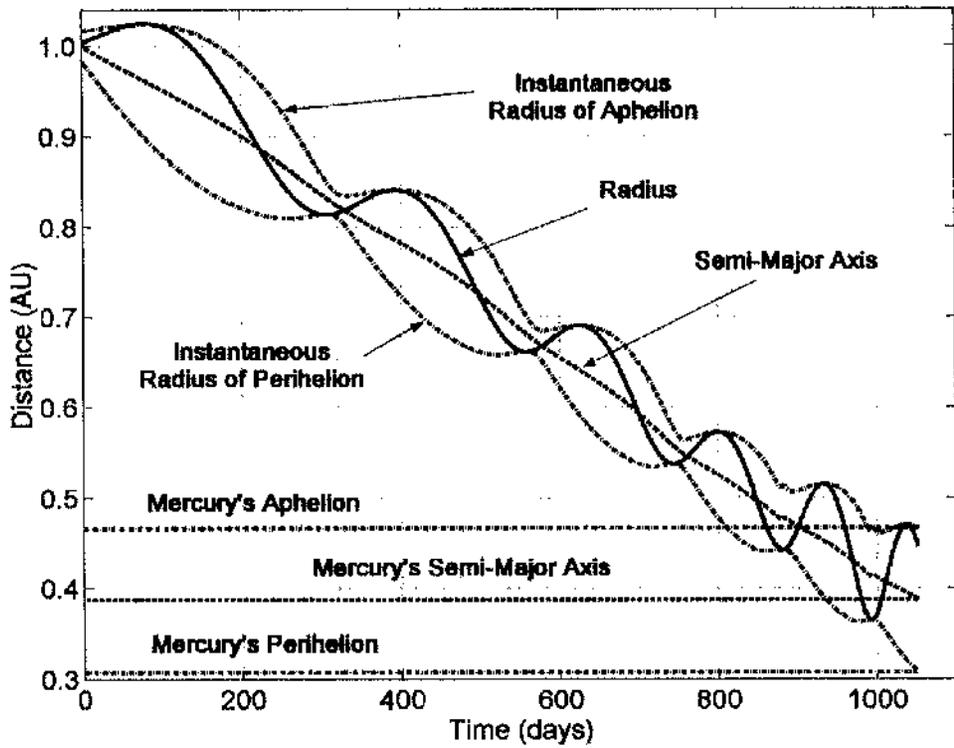
Figure 99 Orbit radius and semi-major axis plot for an Earth – Mercury transfer.
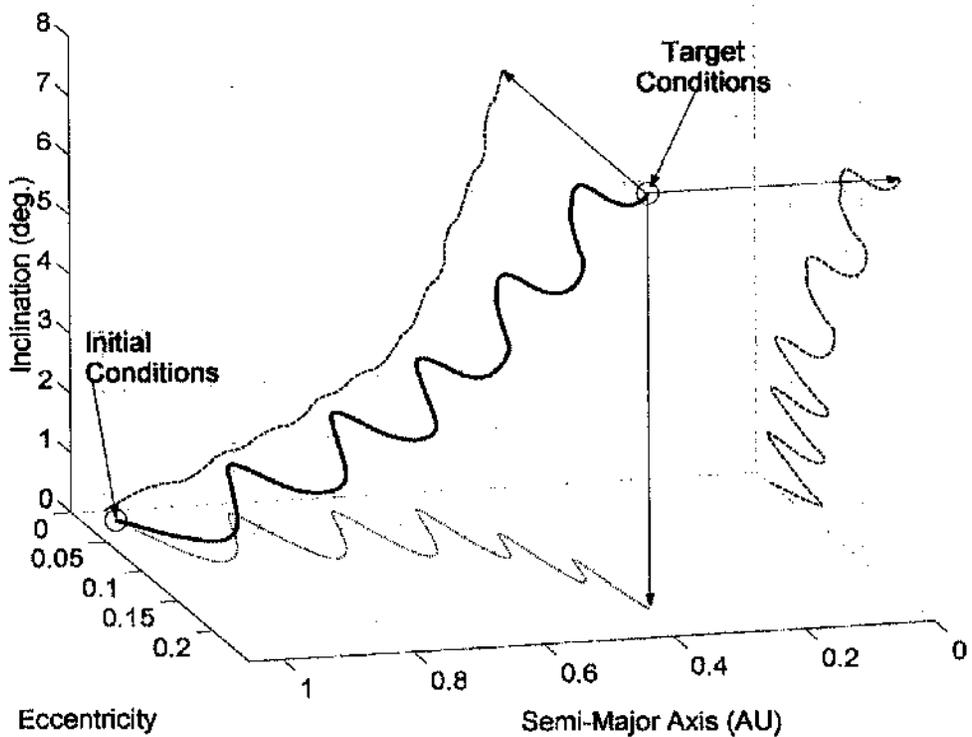
Figure 100 Diagram of a, e and i during Earth – Mercury transfer, with projection into axis also shown.
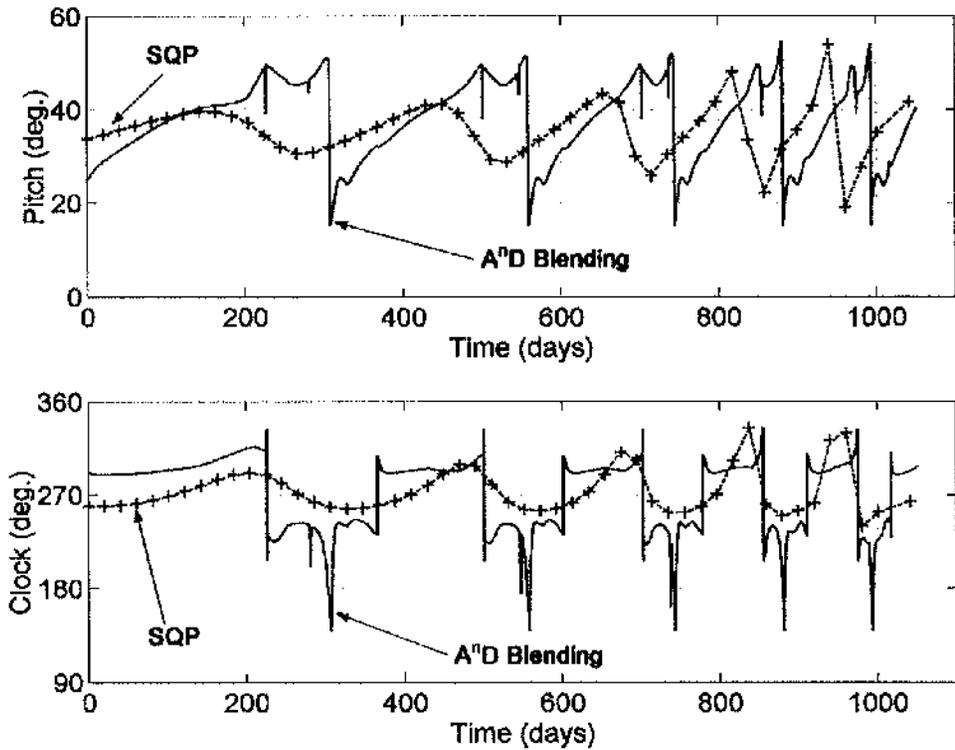
Figure 101   Control angles for the Earth – Mercury transfer, A^nD blending (—) and SQP (–+–) generated angles both shown.

6.4   Solar Polar Orbiter Transfer with a Gravity-Assist

The core of a Solar Polar Orbiter (SPO) mission was outlined in Section 1.5.2, from where it is recalled that the SPO mission is another of the ESA TRS. The target solar polar orbit is defined by the direction of the solar poles. The solar equator is inclined at 7.25 deg to the ecliptic plane, with a right ascension of ascending node angle of 75.8 deg plus 1.4 deg for every century from J2000. Thus, the desired polar orbit is inclined at 82.75 deg with an ascending node of 255.8 deg + 0.014 deg yr^{-1} from J2000, within a standard ecliptic plane reference frame.[81] Analysis of Sunspot motion has revealed that the direction of the solar poles is less well defined than indicated above, however these values are acceptable as mean values for the purpose of this analysis.[206] Spacecraft orbit phasing with respect to Earth must be carefully considered such as to position the SPO near to the solar limb as seen from Earth, allowing observation of the corona along the Sun-Earth line. Maintaining this alignment eliminates solar conjunctions and hence loss of telemetry. It is thus necessary that the spacecraft orbit is in resonance with Earth's orbit about the Sun. Potential target solar orbits are defined as circular with radius $N^{(-2/3)}$ AU, for integer values of N. The baseline mission uses $N = 3$, giving a target solar radius of 0.48 AU. Using A^nD blending a solar sail, gravity assist combination transfer trajectory is considered.

Multiple gravity assists within the inner solar system tend to be prolonged in duration and can be limited in launch window frequency, especially if considering non-resonant combinations. It is therefore anticipated that any benefit over a non-gravity assist scenario will occur through use of a single gravity assist

manoeuvre, probably at Venus; allowing for a perihelion inside the Venusian orbit. Use of a Mars or Earth fly-by would result in a high aphelion, which is detrimental to solar sailing. Furthermore, it is envisaged that sail deployment will commence only after the final gravity assist due to guidance difficulties with such a large structure and inaccurate pointing control. The delayed deployment of the sail will avoid the need for accurate sail navigation and control during the gravity assist. However, it does necessitate an autonomous sail deployment at a slant range of approximately 0.6 AU. Using a Soyuz Fregat 2-1b launch from Kourou one can consider a vast range of potential solar orbits post-Venus fly-by. Note however that the orbit aphelion is restricted to 0.73 – 0.74 AU, set by the position of Venus. It is possible to acquire a post-Venus orbit perihelion of 0.48 AU. However, it is found that when trying to circularise the orbit, prior to the inclination raising phase of the trajectory, maintaining a minimum solar approach radius of 0.48 AU significantly extends the mission duration. It was thus found that increasing the post-Venus orbit perihelion reduced mission duration. At a perihelion of 0.48 AU the maximum plane-change due to the Venusian fly-by with a Soyuz vehicle is approximately 16 deg. However, as one raises perihelion one can also increase the plane change magnitude. It was found that the maximum possible plane change magnitude at Venus fly-by for a Soyuz Fregat 2-1b launch from Kourou was 18.5 deg. However, this reduces the mission launch window size to one day. It was found that for a sail characteristic acceleration of 0.40 mm s^{-2}, the adopted performance level for this analysis, the post-Venusian perihelion orbit should be approximately 0.5 – 0.6 AU, depending on the fly-by conditions. Additionally, it is noted that in this range of perihelion values the orbit inclination can be maintained at 18 deg, with a suitably long launch window. The Venusian fly-by provides the sail deployment orbit with an ascending node angle fixed by the position and nature of the gravity assist manoeuvre. The ascending node angle can be altered by altering the post-Venusian orbit perihelion, thus in order to ensure a suitable ascending node angle one must vary the radius of perihelion. It is thus highly unlikely that the optimal perihelion for a single Venus fly-by scenario will be the same as during preceding or subsequent launch windows and may even change slightly within a launch window. It is however of interest that the ascending node angle was always found to be within close proximity of the required value for a true solar polar orbit and as such a launch window occurs every Venus opportunity.

An initial orbit of 0.73 AU × 0.52 AU × 18 deg was identified for a launch opportunity in December/January of 2016/2017 which approximately aligned the post fly-by orbit ascending node angle with the required solar polar orbit fly-angle. The launch window is shown in Figure 102. It was noticed that during the orbit inclination cranking phase of the trajectory the ascending node angle tends to drift forward by approximately 2 – 3 deg, thus rather than aim the post fly-by ascending node angle exactly to match the required solar polar value it was found that trip time was minimised by aiming the ascending node angle slightly low. The launch window for this opportunity lasts from 27 December 2016 until 08 January 2017; with a maximum allowed launch mass of 650 kg and a maximum launch C$_3$ of 35.5 km^2 s^{-2}, as shown in Figure 102. Selecting the first available launch opportunity within the window a 2883 km Venus fly-by is performed 142 days later, on 18 May 2017, placing the un-deployed solar sail on a 0.73 AU × 0.52 AU × 18 deg orbit. The post fly-by ascending node angle is 253.34 deg, which is approximately 2.5 deg below the required solar polar value.

Following sail deployment the primary trajectory goal is to circularise the orbit at 0.48 AU. During orbit circularisation it was found that the orbit inclination can be increased slightly with no adverse effect on the

circularisation goal. During circularisation the trajectory is split into two phases, with the constants used on the A^nD scores given in Table 11 resulting in a circular orbit of 0.48 AU with inclination 22.32 deg after 195 days; 337 days after launch. Phase one reduces the semi-major axis of the post-Venus orbit and lasts for 137 days, while the second phase has duration 58 days and circularises the orbit. The combination of the A^nD score and constants in Table 11 are shown in Figure 103 with the resulting A^nD generated sail control angles shown in Figure 104. It is noted that the control angle profile is relatively smooth except for the step change at day 137, which corresponds to the end of phase one. Following the circularisation of the orbit at 0.48 AU and inclination 22.32 deg the locally optimal inclination increase control law is used to "crank" the orbit.

The increase in orbit inclination throughout the entire sail trajectory is shown in Figure 105. It is noted that the inclination is slowly increased through to day 195, where the exclusive use of the locally optimal increase of inclination controller begins a much more rapid change. Figure 105 also shows the variation of ascending
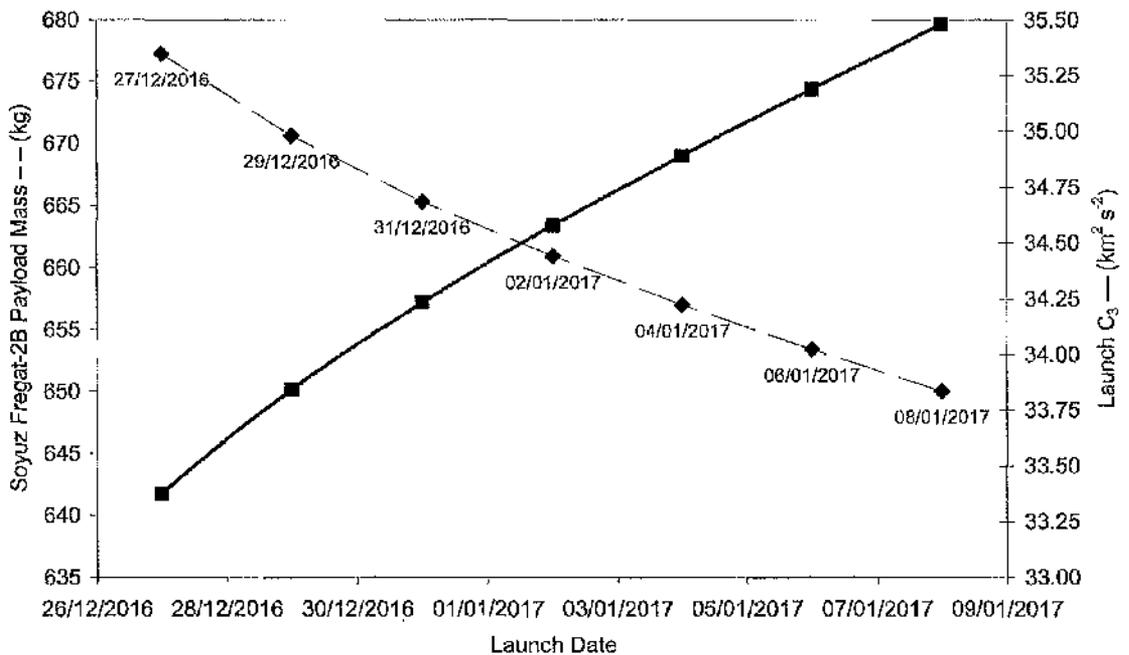
Figure 102 SPO launch window opportunities during December 2016 – January 2017 (– –) and corresponding launch C_3 (—).

Parameter	Constant in Phase 1	Constant in Phase 2
Eccentricity	0.00	1.0
Semi-Major Axis	1.00	0.55
Inclination	0.25	0.57

Table 11 Constants used during circularisation of orbit.

node angle following sail deployment. Recall that the ascending node has a target value of 255.8 deg + 0.014 deg yr^{-1} from J2000, noting that this trajectory is arriving at the polar orbit in approximately the middle of 2021 the target ascending node angle is thus approximately 256.1 deg. Figure 105 shows four potential sail jettison points, where the orbit inclination, in the ecliptic reference, is greater than 60 deg and the orbit element target conditions have been met. The four points correspond to heliographic inclinations 72.8 deg, 79.0 deg, 85.2 deg and 91.4 deg. However, it is found that the first two points are unsuitable as following sail jettison the Earth would not have a continuous, uninterrupted view of the SPO spacecraft as the spacecraft – Earth – Sun angle periodically drops below 10 deg. The third and fourth points do provide continuous, uninterrupted communications capability and are thus suitable termination points. The first suitable sail jettison point occurs 1245.6 days after sail deployment, or 4.16 years after launch. The second suitable sail jettison point occurs 4.52 years after launch. If a true and exact solar polar orbit is required further optimisation of the final 10 – 15 deg of this second point would be required to identify the correctly phased transfer orbit. However, from A^{n}D blending one can estimate the transfer time as 4.5 years from launch to an exact solar polar orbit, a reduction of approximately half a year from a trajectory without a Venus gravity assist.[81] Assuming passage directly over the solar pole is not required, as seems likely, and a heliographic inclination of approximately 85 deg is suitable, one can thus estimate the SPO transfer time, using a Venus gravity assist as 4.2 years, with an ideal sail (i.e. Model 1) with characteristic acceleration 0.4 mm s^{-2}. The SPO trajectory from sail deployment through to the second viable sail jettison point is shown in Figure 106.
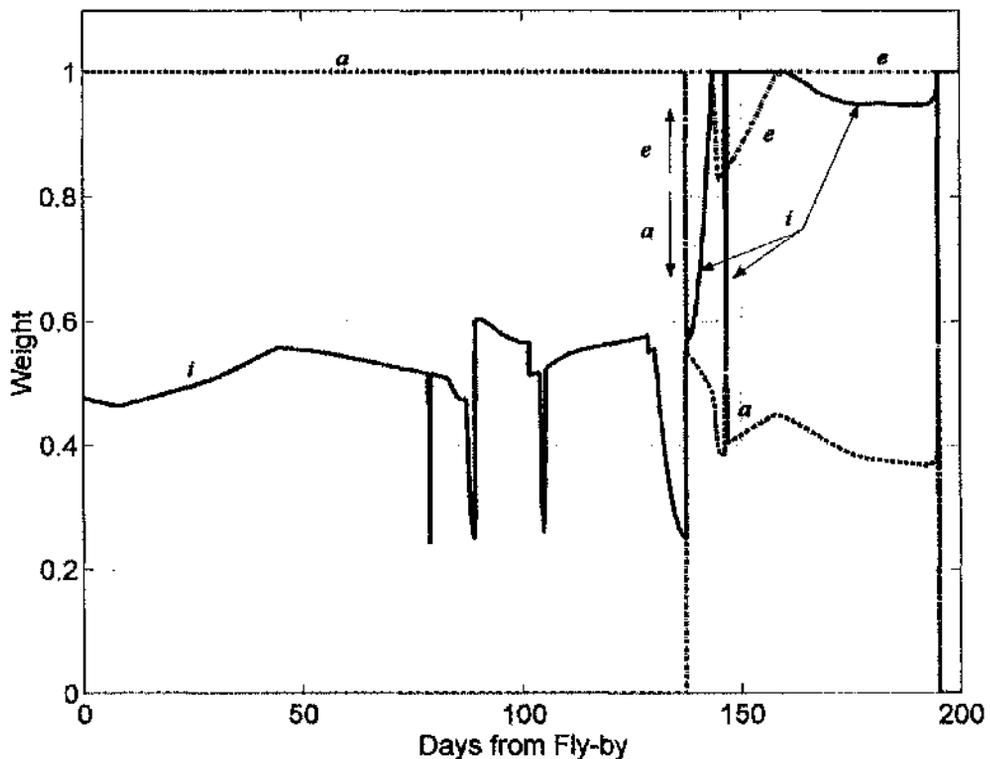
Figure 103 Weights used during circularisation of orbit.
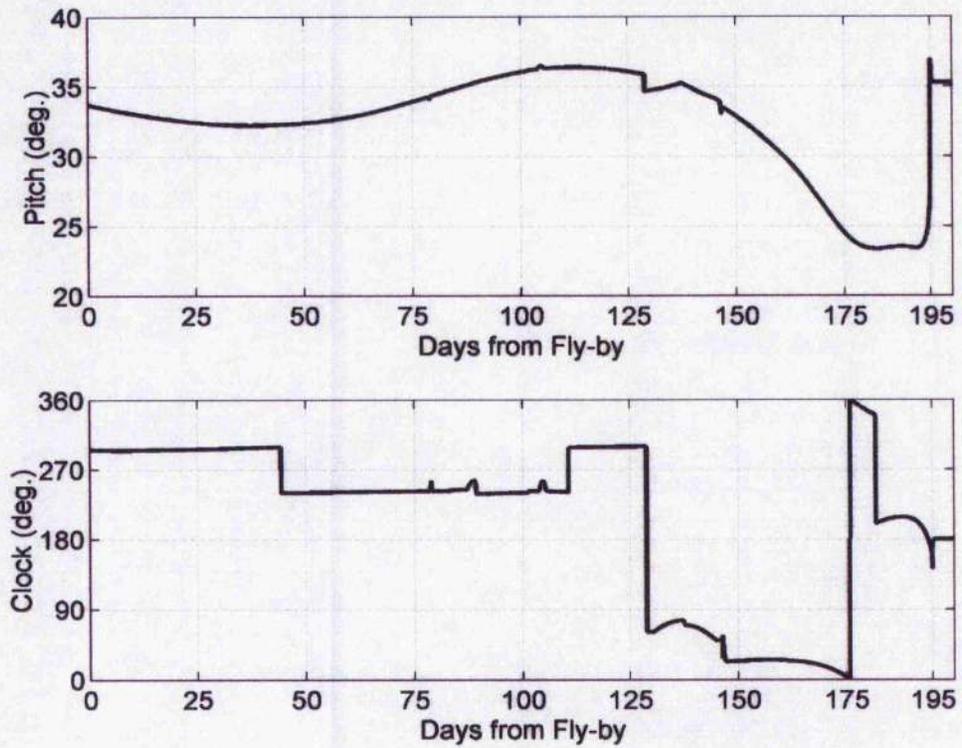
Figure 104 Sail control angles resulting from Figure 103.
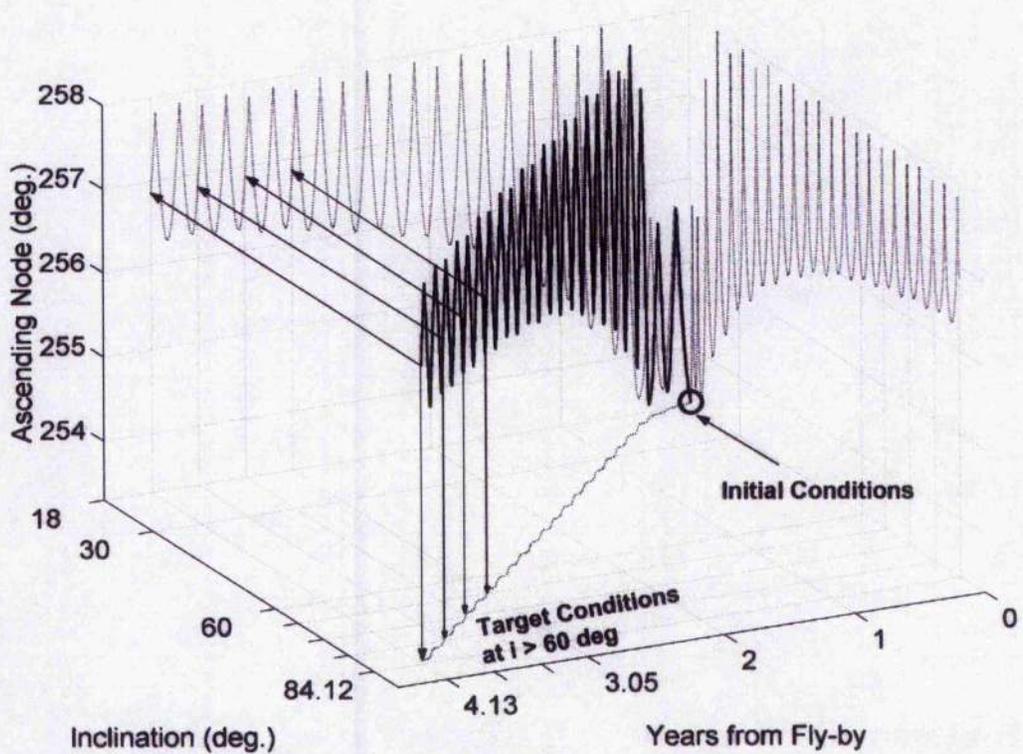
Figure 105 Inclination (ecliptic reference) and ascending node angle verses time from sail deployment through to jettison, with projection of data in axis also shown.
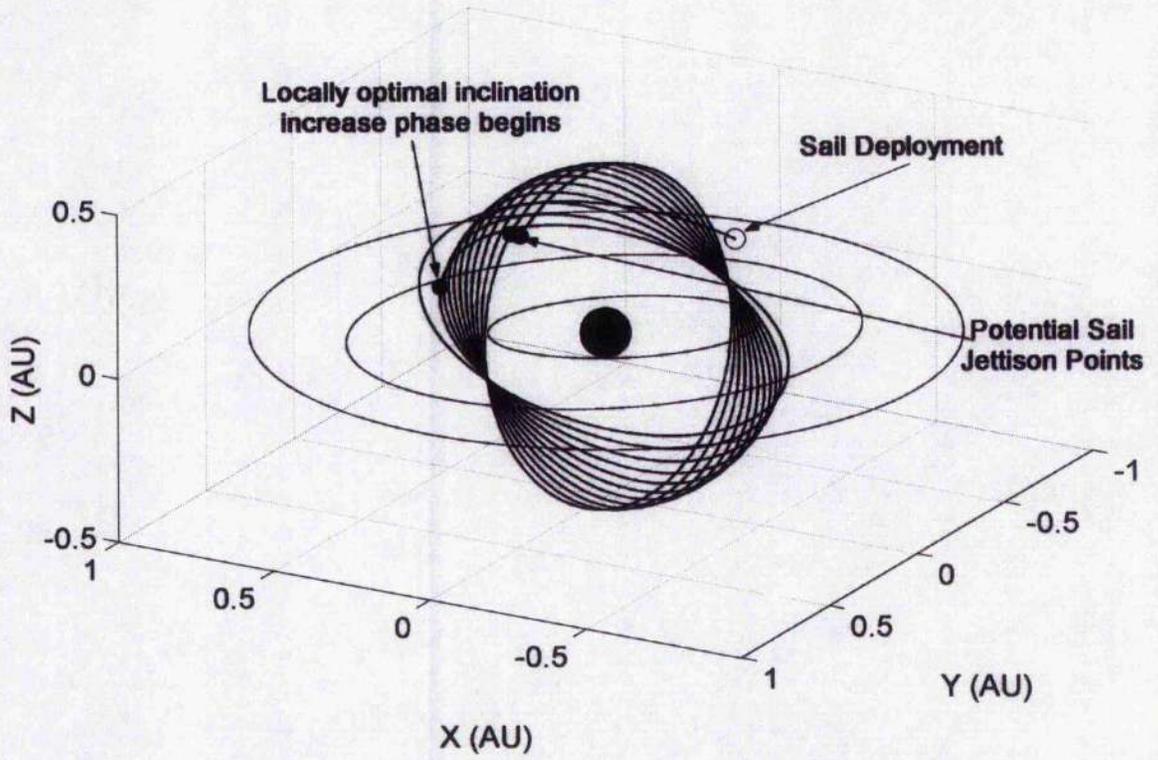
Figure 106 SPO trajectory from sail deployment following Venus gravity assist.

7   Conclusions and Further Work

The work objectives of this dissertation were outlined in Section 1.6, where six questions and issues were listed as requiring investigation for completion of the thesis. The six points are reproduced below along with the relevant findings and conclusions.

1. Perform an analytical investigation of planetary escape using solar sail propulsion such as to explain previously noted but unexplained anomalies; for example, the variation in Earth escape time as a function of the Earth's position about the Sun.

   It has been shown in Sections 4.1 and 4.2 that variation of orbital elements, where the rate of change is a function only of the radial and transverse sail acceleration, is optimally induced by a solar sail operating within the ecliptic plane. It was further demonstrated that Earth shadow does not alter this optimal configuration despite a drop in sail propulsive efficiency. The derivation and corroboration of the ecliptic plane as the optimal orbit orientation explains a prior anomaly identified within the literature but not previously explained.

2. Investigate, for the first time, the variation in solar sail escape time from Mercury as a function of Mercury's position about the Sun.

   An apparent anomaly over the optimal time of the Hermian year to begin an escape trajectory was investigated and clarified in Section 4.3. It was found that low performance sails have a minimum escape duration when the trajectory begins at Mercury's aphelion, with the sail passing through the point of escape close to perihelion. As the sail performance is increased the optimal start date to minimise the escape duration migrates from aphelion passage towards perihelion passage. Furthermore, to minimise the Julian Day at escape the optimal strategy was found to never require a period of delay in the start epoch.

3. Generate an analytical solar sail trajectory design method which allows the generation of near-optimal realistic planetary escape trajectories for the first time (trajectories which do not, for example, have negative altitude phases).

   A method of blending locally optimal control laws was presented in Section 4.4 which maintains the near-optimal nature of the locally optimal energy gain controller, while also insuring a safe minimum altitude through use of a pericentre control law. The algorithms presented are explicitly independent of time and have been shown capable of adapting to different perturbations from those included within the original design scenario. Thus, the control algorithms are potentially suitable as an autonomous on-board controller.

4. Perform a thorough investigation of solar sail Earth escape trajectories which do not pass through the Earth's shadow cone, while using the methods developed for point 3.

> The required sail characteristic acceleration to escape from a polar orbit without Earth occultation of the sail/Sun line was investigated in Section 4.5. It was shown that the required sail characteristic acceleration increases exponentially as initial altitude is decreased. It was also seen that the time until escape corresponding to the minimum sail acceleration requirement was largely independent of initial altitude, with an approximate duration of 141.5 days.

5. Develop an analytical solar sail trajectory design method which can rapidly produce complex planet-centred orbit transfers and station-keeping algorithms for the first time.

> A new method of assessing the relative importance of orbit elements during solar sail transfer and station-keeping manoeuvres has been introduced in Section 5.2, allowing rapid generation of trajectories by blending locally optimal control laws. The Accessibility and Deficit, A^nD, blending method considers both an orbital element's variation from its target value and how well that orbital element will use the sail thrust prior to calculating a score for the element. A^nD blending directs the blended locally optimal force vector such that it avoids prolonged periods of high sail pitch settings, which are an inefficient use of the sail, thus increasing sail efficiency.

> The A^nD blending method has been demonstrated for generation of a complex orbit transfer at Mercury in Section 5.3.1, where the primary constraint of the transfer was thermal rather than time. The transfer trajectory rotates argument of pericentre by 180 deg, while continually re-targeting towards a new ascending node final value, which is initially rotated 180 deg from the trajectory's starting ascending node value. The trajectory is verified shadow free and all orbit elements converge well with the target values. The use of A^nD blending allows the generation of such a transfer trajectory in a much more rapid fashion than would be possible with conventional optimisation techniques.

> Furthermore, A^nD blending has been demonstrated to act as an excellent solar sail station-keeping algorithm in Sections 5.3.2 and 5.4; capable of adjusting to significant variations in orbit perturbation magnitude. Similar to Section 4, the control method demonstrated is capable of providing the sail control angles in real-time, based solely on the current spacecraft state-vectors. Thus, making A^nD blending suitable as a potential autonomous on-board sail control system.

6. Develop an analytical heliocentric solar sail trajectory design method which can rapidly produce near-optimal solutions (trajectories within $1 - 3$ % of optimal).

The A^nD blending method has been demonstrated for generation of sail trajectories to 200 AU in Section 6.2, where it was repeatedly shown that A^nD generated trajectories were very similar in duration to the optimal transfer time found using an evolutionary neurocontroller. Furthermore, the A^nD generated trajectories were consistently more efficient than sequential quadratic programming, SQP, generated trajectories. A^nD blending has been shown able to design a solar sail trajectory using a non-ideal sail force model and to restrain the trajectory based on either a minimum radius or maximum temperature. The A^nD blending results clearly demonstrate that a non-ideal sail will require a characteristic acceleration of approximately 1.5 mm s^{-2} to reach 200 AU in 25 years, assuming no optical surface degradation.

An Earth – Mercury trajectory was presented in Section 6.3 to demonstrate the capability of A^nD blending when attempting to find planet – to – planet transfer trajectories. It was found that the A^nD blending generated trajectory duration was within 1 % of the SQP generated trajectory and that the control angle profiles were closely matched. Thus, A^nD blending can provide a very good rapid assessment of such a mission scenario, or provide an excellent initial guess for further optimisation as part of a detailed mission analysis.

It is concluded that A^nD blending is a highly efficient method for the rapid generation of heliocentric trajectories, reducing the trajectory optimisation problem from finding the cone and clock angle control history, to finding a set of constants. Thus, in the Earth-Mercury transfer trajectory where the SQP method had to optimise 102 data points (51 nodes of pitch and 51 nodes of clock) the A^nD method reduces the problem to only 6 data points, split evenly over two trajectory phases; a significant reduction in computational effort. Furthermore, the 200 AU trajectories simplify the equivalent SQP problem from 402 data points, to 4 data points for the dual loop trajectories and only 2 data points for the single loop trajectories; a reduction in computational effort of many orders of magnitude. It is highly likely that further effort to optimise the constants would reduce the sub-optimal nature of the solutions. However, the current manual nature of the optimisation process hinders this. A^nD blending allows swift and accurate mission analysis while also providing an excellent initial guess to other optimisation methods.

The work documented within this dissertation could be extended in many ways. Two principle areas of future work are:

- Automation of the selection of the constants which are applied to the A^nD scores.

- Extension of the A^nD blending method to other low-thrust propulsion methods, such as electric propulsion.

By automating the selection of constants which are applied to the A^{n}D scores one could expect to obtain a much more rapid trajectory generation method, as discussed in Section 6.2.3. Automation is however not a simple matter, as the automatic selection of trajectory phases could prove troublesome. Assuming this can be solved, either through automation or the user optimising several individual phases per trajectory, then the sub-optimal nature of the results would likely be reduced. The method of automatically selecting the constants which are applied to the A^{n}D scores requires study and it would be rash to propose a suitable method without detailed study of the options. It is logical however that application of a global optimiser would be a best-case solution, as this would remove the requirement on the user to provide an initial guess of the constants. Furthermore, if a global optimiser such as an evolutionary neurocontroller was applied then it may be possible for the neurocontroller to learn where is best to split a trajectory into each of the required phases, thus through experience the system will become increasingly automated.

Extension of A^{n}D blending to other low-thrust propulsion methods would require careful consideration, as the propellant used must now also be considered. Thus, thrust-arcs must be introduced through analysis of the variational equations of motion to ensure that the engines only thrust when they will have a suitably large impact on the orbit elements. Similarly, the engines should not be rapidly turned on and off. This is analogous to the current Accessibility component of the A^{n}D score, however development would be required. It is considered by the candidate that development of A^{n}D blending to reaction low-thrust methods should initially be performed independent of the automation of the solar sail method. In other words the initial reaction low-thrust trajectories should be generated by a manual process, as performed within this dissertation for solar sailing. Thus, ensuring the methods simplicity is maintained prior to automation, rather than relying on computational effort to replace simplicity.

8 References

[1] McInnes, C.R., "Solar Sailing: Technology, Dynamics and Mission Applications", Springer-Praxis, Chichester, 1999.

[2] Maxwell, J. C., "Electricity and Magnetism", Oxford University Press, 1873.

[3] NASA/JPL Technical Memorandum 33-734, Volume 2, "Mariner Venus Mercury 1973 Project Final Report, Extended Mission-Mercury II and III Encounters", 01 December 1975.

[4] NASA/JPL Technical Memorandum 33-734, Volume 1, "Mariner Venus Mercury 1973 Project Final Report, Venue and Mercury 1 Encounters", 15 September 1976.

[5] Shirley, D., "The Mariner 10 Mission to Venus and Mercury", IAF-02-Q.4.1.01, Proceedings of 53rd International Astronautical Congress, Houston, Texas, 10-19 October 2002.

[6] Lebedew, P., "The Physical Causes of the Deviation from Newton's Law of Gravitation", Astrophysical Journal, Vol. 10, pp 155-161, 1902.

[7] Nichols, E.F., Hull, G.F., "A Preliminary Communication on the Pressure of Heat and Light Radiation", Physics Review (Series I), Vol. 13, pp. 307-320, 1901.

[8] Nichols, E.F., Hull, G.F., "The Pressure Due to Radiation. (Second Paper.)", Physics Review (Series I), Vol. 17, pp. 26-50 and pp 91-104, 1903.

[9] Ciolkowski, K.E., "Extension of Man into Outer Space", 1921.)Also, Tsiolkovsky, K.E., "Symposium Jet Propulsion", No. 2, United Scientific and Technical Presses, 1936)

[10] Tsander, K., "From a Scientific Heritage", NASA Technical Translation TTF-541, 1967, a translation of "Iz Nauchnogo Naslediya", Nauca Press, Moscow, 1924.

[11] Ciolkowski, K.E., "Exploration of the Universe with Reaction Machines", The Science Review, No. 5, St. Petersburg, 1903.

[12] Oberth, H., „Die Rakete zu den Planetenräumen", Oldenbourg, München, pp. 86-88, 1923. In German.

[13] Oberth, H., „Wege zur Raumschiffahrt", Oldenbourg, München, pp. 353-371, 1929. In German.

[14] Wiley, C., [pseudonym Sanders, R.] "Clipper Ships of Space", Astounding Science Fiction, p. 135, May 1951.

[15] Garwin, R.L., "Solar Sailing – A Practical Method of Propulsion Within the Solar System", Jet Propulsion, Vol. 28, pp. 188-190, March 1958.

[16] Cotter, T.P., "Solar Sailing", Sandia Research Colloquium, SCR-78, April 1959.

[17] Tsu, T.C., "Interplanetary Travel by Solar Sail", American Rocket Society Journal, Vol. 29, pp. 422-427, June 1959.

[18] London, H.S., "Some Exact Solutions of the Equations of Motion of a Solar Sail with a Constant Setting", American Rocket Society Journal, Vol. 30, pp. 198-200, February 1960.

[19] Sands, N., "Escape from Planetary Gravitational Fields by Using Solar Sails", American Rocket Society Journal, Vol. 31, pp. 527-531, April 1961.

[20] Gordon, B.J., "A Study of Interplanetary Travel by Solar Sail", M.S. Thesis, Air Force University, USAF, Wright-Patterson Air Force Base, Ohio, August 1961.

[21] Fimple, W.R., "Generalized Three-Dimensional Trajectory Analysis of Planetary Escape by Solar Sail", American Rocket Society Journal, Vol. 32, pp. 883-887, June 1962.

[22] Birnbaum, M., "Minimum Time Earth / Mars trajectory for Solar Sail", GGC/EE/68-3 AD-836734, Air Force Institute of Technology, USA, 1968.

[23] Norem, R., "Interstellar Travel, A Round Trip Propulsion System with Relativistic Velocity Capabilities", AAS Paper 69-388, June 1969.

[24] Bacon, R.H., "Logarithmic Spiral: An Ideal Trajectory for the Interplanetary Vehicle with Engines of Low Sustained Thrust", Journal Physics, Vol. 27, pp. 164-165, 03-1959.

[25] Fekete, T.A., Sackett, L. L., von Flotow, A.H., "Trajectory Design for Solar Sailing from Low-Earth Orbit to the Moon", AAS 92-184, Advances in Astronautical Sciences, Vol. 79, Pt. 3, pp. 1083-1094, 1992.

[26] Eguchi, S., Ishii, N., Matsuo, H., "Guidance Strategies for Solar Sail to the Moon", AAS 93-653, Advances in Astronautical Sciences, Vol. 85, Pt. 2, pp. 1419-1433, 1993.

[27] Morgan, T.O., "The Inclination Change for Solar Sails and Low Earth Orbit", AAS 79-104, Advances in Astronautical Sciences, 1979.

[28] Irving, J.H., "Space Technology", John Wiley & Sons Inc., New York, 1959.

[29] Lawden, D.F., "Optimal Escape from a Circular Orbit", Astronautica Acta, Vol. 4, pp. 218-234, 1958.

[30] Green, A.J., "Optimal Escape Trajectories From a High Earth Orbit by Use of Solar Radiation Pressure", T-652, Master of Science Thesis, Massachusetts Institute of Technology, 1977.

[31] Sackett, L.L., "Optimal Solar Sail Planetocentric Trajectories", R-1113, The Charles Stark Draper Laboratory, Inc., JPL-NASA Contract NAS 7-100 Final Report, September 1977.

[32] Sackett, L.L. & Edelbaum, T.N., "Optimal Solar Sail Spiral to Escape", Advances in Astronautical Sciences, AAS/AIAA Astrodynamics Conference, A78 31-901, 1978.

[33] Coverstone-Carroll, V. & Prussing, J.E., "A Technique for Earth Escape Using a Solar Sail", Advances in Astronautical Sciences, AAS 99-333, Astrodynamics Specialist Conference, Westin Prince Alyeska Resort, Girdwood, Alaska, August 16-19 1999.

[34] Coverstone, V. L. & Prussing, J.E., "Technique for Earth Escape from Geosynchronous Transfer Orbit Using a Solar Sail", Journal of Guidance, Control and Dynamics, Vol. 26, No. 4, 2003.

[35] Leipold, M., "Solar Sail Mission Design", Doctoral Thesis, Technische Universität München, DLR-Forschungsbericht 2000-22.

[36] Pagel, G., „Extremale Steuerstrategien für Sonnensegler am Beispiel von Bahntransferproblemen zum Erdmond", Doctoral Thesis, Technischen Universität Berlin, May 2002. In German.

[37] MacNeal, R.H., "Comparison of the Solar Sail with Electric Propulsion Systems", NASA-CR-1986, February 1972.

[38] Lippman, M., E., "In-Space Fabrication of Thin-Film Structures", NASA-CR-1969, Astro Research Corporation ARC-R-410, 1972

[39] MIT Student Project, "Concepts for the Space Disposal of Nuclear Waste, Appendix B, Propulsion by Reflection of Light from the Sun", Aeronautics and Astronautics Project, Spring Term 1972.

[40] Grinevitskaia, L.K., Polaikhova, E.N., "Derivation of an Approximate Solution to the Equation of Geocentric Motion of a Space Vehicle with a Solar Sail", Vestnik, Mekhanika, Astronomiia, Leningrad University (Now Saint-Petersburg State University), April 1973. In Russian.

[41] Cotter, T.P., "An Encomium on Solar Sailing", Informal Report LA-5231-MS, Los Alamos Scientific Laboratory, May 1973.

[42] Wright, J.L., "Solar Sailing: Evaluation of Concept and Potential", Battelle Memorial Institute Report, no. BMI-NLVP-TM-74-3, November 1974.

[43] Uphoff, C.W., "Solar Sailing: Potential for Future Missions", JPL Report 760-164, 25 January 1977. Revision of internal memorandum, dated 10 June 1975.

[44] Wright, J., Warmke, J., "Solar Sailing Mission Applications", Paper No. 76-808, AIAA/AAS Astrodynamics Conference, San Diego, August 1976.

[45] MacNeal, R.H., "The Heliogyro, An Interplanetary Flying Machine", NASA Contractor's Report, CR 84460, June 1967.

[46] Hedgepath, J.M., Benton, M.D., "Analysis of Planetary Flyby Using the Heliogyro Solar Sailor", Astro Research Corporation ARC-R-296, August 1968.

[47] MacNeal, R.H., Hedgepath, J.M., Schuerch, H. U., "Heliogyro Solar Sailor Summary Report", NASA-CR-1329, June 1969.

[48] MacNeal R.H., "Structural Dynamics of the Heliogyro", NASA-CR-1745A, 1971.

[49] Jacobson, R.A., Thornton, C.L., "Elements of Solar Sail Spacecraft Navigation with Application to a 1986 Halley's Comet Rendezvous Mission", AAS/AIAA Astrodynamics Conference, Jackson Hole, WY, September 1977.

[50] Friedman, L., Carroll, W., Goldstein, R., Jacobson, R., Kievit, J., Landel, R., Layman, W., Marsh, E., Ploszaj, R., Rowe, W., Ruff., W., Stevens, J., Stimpson, L., Trubert, M., Varsi, G., Wright, J., MacNeal, R., "Solar Sailing - The Concept made Realistic", AIAA 78-82, 16th AIAA Aerospace Sciences Meeting, Huntsville, January 1978.

[51] Sauer, D.G., "Comparison of Solar Sail and Ion Drive Trajectories for a Halley's Comet Rendezvous Mission", AAS/AIAA Astrodynamics Conference, Jackson Hole, WY, September 1977.

[52] Logsdon, J.M., "Missing Halley's Comet: The Politics of Big Science", ISIS, Vol. 80, No. 302, pp 254-280, 1989.

[53] Freeland, R.E., Bilyeu, G.D., Veal, G.R., Steiner, M.D., Carson, D.E., "Large Inflatable Deployable Antenna Flight Experiment Results", IAF-97-1.3.01, Proceedings of 48th International Astronautical Congress, Torino, October 1997.

[54] Leipold, M., Eiden, M., Garner, C., Herbeck, L., Kassing, D., Niederstadt, T., Krüger, T., Pagel, G., Rezazad, M., Rozemeijer, H., Sebolt, W., Schöppinger, C., Sickinger, C., Unckenbold, W., "Solar Sail Technology Development and Demonstration", 4th IAA International Conference on Low-Cost Planetary Missions, Laurel, Maryland, May 2000.

[55] Sebolt, W., Leipold, M., Rezazad, M., Herbeck, L., Unckenbold, W., Kassing, D., Eiden, M., "Ground-Based Demonstration of Solar Sail Technology", IAF-00-S.6.11, Proceedings of 51st International Astronautical Congress, Rio de Janeiro, Brasil.

[56] Herbeck, L., Sickinger, C., Herrmann, A., „Ultraleichte, Entfaltbare Maststrukturen aus CFK", DGLR Annual Meeting, Berlin, September 1999. In German.

[57] Lichodziejewski, D., Derbès, B., West, J., Reinert, R., Belvin, K., Pappa, R., "Bringing an Effective Solar Sail Design Towards TRL 6", AIAA 2003-4659, AIAA Joint Propulsion Conference, July 2003.

[58] Murphy, D.M., Murphey, T.W., Gierow, P.A., "Scalable Solar-Sail Subsystem Design Concept", Journal of Spacecraft and Rockets, Vol. 40, No. 4, 2003.

[59] Murphy, D.M., Macy, B.D., Gaspar, J.L., "Demonstration of a 10-m Solar Sail System", AIAA-200-1576, 40th Joint Propulsion Conference, Fort Lauderdale, Florida, July 2004.

[60] Garbe, G., Montgomery IV, E.E., "An Overview of NASA's Solar Sail Propulsion Project", AIAA 200-5274, Proceedings of 39th AIAA Joint Propulsion Conference, Huntsville, July 2003.

[61] Saburo, M., Osamu, M., Koji, N., Masafumi, I., Kuniyuki, O., Hideyuki, Y., "Initial investigation on Deployment and Control of Tether-Controlled Spinning Membrane", Proceedings of 13th Workshop on Astrodynamics and Flight Mechanics, Sagamihara, pp. 124-129, July 2003.

[62] Shinsuke, T., Junjiro, O., "Deployment Motion of Spin Type Solar Sail", Proceedings of 13th Workshop on Astrodynamics and Flight Mechanics, Sagamihara, pp. 109-112, July 2003.

[63] Takehiko, O., "A Study on the Spin Deployment Dynamics of Solar Sail", Proceedings of 13th Workshop on Astrodynamics and Flight Mechanics, Sagamihara, pp. 113-116, July 2003.

[64] Yuichi, T., Jun'ichiro, K., Takafumi, M., " Deployment Experiments Using S310 Sounding Rocket and It's Dynamic Analysis", Proceedings of 13th Workshop on Astrodynamics and Flight Mechanics, Sagamihara, pp. 104-108, July 2003.

[65] Yusuke, N., Yuichi, T., Osamu, M., Jun'ichiro, K., "The Deployment Experiment of Solar Sail with a Sounding Rocket", IAC-04-A.5.10, Proceedings of 55th International Astronautical Congress, Vancouver, October 2004.

[66] Shinsuke, T., Yuichi, T., Osamu, M., Aoki, Y., "Deployment Experiment Result of Solar Sail using Sounding Rocket", IAC-04-I.7.09, Proceedings of 55th International Astronautical Congress, Vancouver, October 2004.

[67] Lewis, G.N., Letter to the editor of Nature magazine, Vol. 118, Part 2, pp. 874-875, December 1926.

[68] Griffiths, D. J. "The Photon (1900-1924)." Section 1.2 "Introduction to Elementary Particles", New York: Wiley, pp. 14-17, 1987.

[69] Wright, J.L., "Space Sailing", Gordon and Breach Science Publishers, Philadelphia, 1992.

[70] Rios-Reyes, L., Scheeres, D.J., "Generalized Model for Solar Sails", Journal of Spacecraft and Rockets, Vol. 42, No. 1, pp.182-185, 2005.

[71] Rogan, J., Gloyer, P., Pedlikin, J., Veal, G., Derbes, B., "Encounter 2001: Sailing to the Stars", SSC01-112, Proceedings of the 15th Annual/USU Conference On Small Satellites, Logan, Utah, August 2001.

[72] Solar Sail Degradation Model Working Group , "A Parametric Model for Optical Solar Sail Degradation", Available from URL: http://www.spacesailing.net/SSDMWG/ [citied 03 May 2005]

[73] Dachwald, B., Seboldt, W., Macdonald, M., Mengali, G., Quarta, A.A., M'Innes, C.R., Rios-Reyes, L., Scheeres, D.J., Wie, B., Görlich, M., Lura, F., Diedrich, B., Baturkin, V., Coverstone, V.L., Leipold, M., Garbe, G.P., "Potential Solar Sail Degradation Effects on Trajectory and Attitude Control", AIAA Paper 05-6172, Proceedings of AIAA/AAS Guidance, Navigation, and Control Conference, San Francisco, August 2005.

[74] Dachwald, B., Baturkin, V., Coverstone, V. L., Diedrich, B., Garbe, G. P., Görlich, M., Leipold, M., Lura, F., Macdonald, M., M'Innes, C. R., Mengali, G., Quarta, A.A., Rios-Reyes, L., Scheeres, D.J., Seboldt, W., Wie, B., "Potential Effects of Optical Solar Sail Degradation on Interplanetary Trajectory Design", AIAA Paper 05-413, Proceedings of AAS Astrodynamics Specialist Conference, Lake Tahoe, California, August 2005.

[75] Hollerman, W., Albarado, T., Lentz, M., Edwards, D., Hubbs, W., Semmel, C., "Ionizing Radiation Exposure Measurements for Candidate Solar Sails", Proceedings of the 22nd Space Simulation Conference, , Ellicott City, USA, 2002.

[76] Edwards, D., Hubbs, W., Stanaland, T., Hollerman, A., Altstatt, R., "Characterization of Space Environmental Effects on Candidate Solar Sail Material", in: E. Taylor (Ed.): Photonics for Space Environments VIII, Proceedings of SPIE Vol. 4823, pp. 67-74, 2002.

[77] Lura, F., Hagelschuer, D., Glotov, A.I., Tschaly, Y., "Experiments in the Test Facility KOBE for the Investigation of Degradation Effects on Thin Foil Samples for a Solar Sail Mission Concerning the Simultaneous Influence of Space Environment Properties", Proceedings of 22nd Space Simulation Conference, Ellicott City, USA, 2002.

[78] Edwards, D.L., Hubbs, W.S., Gray, P.A., Wertz, G.E., Hoppe, D.T., Nehls, M.K., Semmel, C.L., Albarado, T.L., Hollerman, A., "Characterization of Candidate Solar Sail Materials Subjected to Electron Radiation", Proceedings of 9th International Symposium on Materials in a Space Environment, Noordwijk, 2003.

[79] Edwards, L., Hubbs, W.S., Wertz, G.E., Hoppe, D.T., Nehls, M.K., Hollerman, W.A., Gray, P.A., Semmel, C.L., "Electron Radiation Effects on Candidate Solar Sail Material", High Performance Polymers, Vol. 16, pp. 277-288, 2004.

[80] McInnes, C.R., Brown, J.C., "Solar Sail Dynamics with an Extended Source of Radiation Pressure", Celestial Mechanics and Dynamical Astronomy, Vol. 9, pp 249-264, 1990.

[81] Macdonald, M., Hughes, G., McInnes, C. R., Lyngvi, A., Falkner, P., Atzei, A., "The Solar Polar Orbiter: A Technology Reference Study", Submitted to Journal Spacecraft & Rockets, March 2005.

[82] Macdonald M., McInnes C. R., "A Near-Term Roadmap for Solar Sailing", IAC-04-U.1.09, Proceedings of 55th International Astronautical Congress, Vancouver, October 2004.

[83] Wertz, J.R., Larson, W.J. (Eds.), "Space Mission Analysis and Design", Kluwer Academic Publishers Group, Dordrecht, 1999.

[84] McInnes, C.R., Macdonald, M., Hughes, G., "High Performance and Multi-Functional Solar Sail Technology", Technical Note 5, Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, December 2003.

[85] McInnes, C.R., Macdonald, M., Hughes, G., "Technology Roadmap", Technical Note 13(b), Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, December 2004.

[85] Garrett, H., Wang, J., "Simulations of Solar Wind Plasma Flow Around A Simple Solar Sail", Proceedings of the 8th Spacecraft Charging Technology Conference, Ed. R. Suggs, MSFC, Huntsville, Alabama, October 2003.

[87] Garrett, H.B., Personal Communiqué, email communications, May 2004.

[88] Ariane Structure for Auxiliary Payloads, on Ariane 5 Launch Manual. Available at; http://www.arianespace.com/site/images/ASAP5-manual.pdf [citied 10 February 2005].

[89] Soyuz user's Manual, Available at; http://www.arianespace.com/site/images/soyuz_users_manual_190401.pdf [citied 10 February 05].

[90] Lichodziejewski, D., Cassapakis, C., "Inflatable Power Antenna Technology", AIAA 99-1074, 37th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January, 1999.

[91] West, J., NMP ST-5 Solar Sail Project Concept Study Report, July 1999.

[92] Wie, B., "Sail Flatness, Attitude, and Orbit Control Issues for an ST-7 Solar Sail Spacecraft", NASA SSTWG FY01, Solar Sail Technical Interchange Meeting, NASA Goddard Space Flight Centre, 2001.

[93] West, J.L., "NOAA/NASA/DOD Geostorm Warning Mission", JPL Internal Document, D-13986, October 1996.

[94] Request for Information (RFI) – Geostorm Warning Mission Documentation, November 1997.

[95] West, J.L., "Solar Sail Vehicle System Design For The Geostorm Warning Mission", AIAA-2000-5326, Proceedings of Structures, Structural Dynamics, and Materials Conference, Atlanta, April 2000.

[96] Chen-wan, L.Y., "Solar Sail Geostorm Warning Mission Design", AAS 04-107, Proceedings of 14th AAS/AIAA Space Flight Mechanics Conference, Maui, Hawaii, February 2004.

[97] Sauer, Jr., C.G., "The L1 Diamond Affair", AAS 04-278, Proceedings of 14th AAS/AIAA Space Flight Mechanics Conference, Maui, Hawaii, February 2004.

[98] Price, H.W., Ayon, P., Buehler, M., Garner, C., Klose, G., Mettler, E., Nakazono, B., Sprague, G., "Design for a Solar Sail Demonstration Mission", Proceedings of Space Technology and Applications International Forum (STAIF 2001), Albuquerque, February 2001.

[99] Leipold, M., Garner, C.E., Freeland, R., Herrmann, A., Noca, M., Pagel, G., Sebolt, W., Sprague, G., Unckenbold, W., "ODISSEE – A Proposal for Demonstration of a Solar Sail in Earth Orbit", Acta Astronautica, Vol. 45, Nos. 4-9, pp.557-566, 1999.

[100] Dachwald, B., Seboldt, W., Richter, L., "Multiple Rendezvous and Sample Return Missions to Near-Earth Objects Using Solar Sailcraft", Proceedings of 5th IAA Conference on Low-Cost Planetary Missions, ESTEC, Noordwijk, September 2003.

[101] Macdonald, M., McInnes, C., "GeoSail: An Enhanced Magnetosphere Mission, Using a Small Low Cost Solar Sail", IAF-00-W.1.06, Proceedings of 51st International Astronautical Congress, Rio de Janeiro, October 2000.

[102] McInnes, C., Macdonald, M., Angelopous, V., Alexander, V., "GEOSAIL: Exploring the Geomagnetic Tail Using a Small Solar Sail", Journal of Spacecraft and Rockets, Vol. 38, No. 4, 2001.

[103] Alexander D., McInnes C. R., Angelopoulos V., Sandman A. W., Macdonald M., "GeoSail: A Novel Solar Sail Mission Concept for Geospace", Proceedings of Space Technology and Applications International Forum (STAIF 2002), Albuquerque, February 2002.

[104] 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, Proceedings of 53rd International Astronautical Congress, Houston, October 2002

[105] 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.

[106] McInnes, C.R., Macdonald, M., Hughes, G., "GeoSail Mission", Technical Note 12, Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, November 2004.

[107] McInnes, C.R., Macdonald, M., Hughes, G., "Jupiter Microsat Explorer Mission", Technical Note 11, Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, September 2004.

[108] Lyngvi, A., Flakner, P., Renton, D., v.d.Berg, M.L., Peacock, A., "Technology Reference Studies", IAC-04-U.1.06, Proceedings of 55th International Astronautical Congress, Vancouver, October 2004.

[109] McInnes, C.R., Macdonald, M., Hughes, G., "Solar Polar Orbiter", Technical Note 10, Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, June 2004.

[110] Lyngvi, A., Falkner, P., Kemble, S., Leipold, M., Peacock, T., "The Interstellar Heliopause Probe", Proceedings of the 55th International Astronautical Congress, Vancouver, October 2004.

[111] McInnes, C.R., Macdonald, M., Hughes, G., "Interstellar Heliopause Probe", Technical Note 9, Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, March 2004.

[112] "A Solar Polar Sail Mission," NASA Report, 2 February 1998. Available at URL: http://spacephysics.jpl.nasa.gov/spacephysics/SolarPolarSail/ [cited 12 June 2004].

[113] Goldstein, B., Buffington, A., Cummings, A.C., Fisher, R., Jackson, B.V., Liewer, P.C., Mewaldt, R.A., Neugebauer, M., "A Solar Polar Sail Mission: Report of a Study to Put a Scientific Spacecraft in a Circular Polar Orbit about the Sun", NASA Jet Propulsion Laboratory, MTG: SPIE International Symposium on Optical Science, Engineering and Instrumentation , July 1998.

[114] Sauer, C. G., Jr., "Solar Sail Trajectories for Solar-Polar and Interstellar Probe Missions", AAS 99-336, Advances in the Astronautical Sciences, Astrodynamics, Vol. 103, No. 1, pp 547 – 562, 1999.

[115] Forward, R. L., "Radioisotope Sails for Deep-Space Propulsion and Power," Journal of The British Interplanetary Society, Vol. 49, pp. 147-149, 1996.

[116] Gruntman, M., McNutt, Jr., R.L., Gold, R.E., Krimigis, S.M., Roelof, E.C., Gloeckler, G., Koehn, P.L., Kurth, W.S., Oleson, S.R., "Innovative Interstellar Explorer", Proceedings of the 55th International Astronautical Congress, Vancouver, October 2004.

[117] Kluever, C. A., "Heliospheric Boundary Exploration Using Ion Propulsion Spacecraft," Journal of Spacecraft and Rockets, Vol. 34, No. 3, May–June 1997.

[118] Leipold, M., Fichtner, H., Heber, B., Groepper, P., Lascar, S., Burger, F., Eiden, M., Niederstadt, T., Sickinger, C., Herbeck, L., Dachwald, B., Seboldt, W., "Heliopause Explorer – A Sailcraft Mission to the Outer Boundaries of the Solar System," Proceedings of Fifth IAA International Conference on Low-Cost Planetary Missions, ESA Special Publication SP-542, September 2003.

[119] McNutt, R. L., Jr. (P.I.), "A Realistic Interstellar Explorer," NIAC CP 98-01, Phase I Final Report, NASA Institute for Advanced Concepts, May 1999.

[120] Mewaldt, R. A., Kangas, J., Kerridge, S. J., and Neugebauer, M., "A Small Interstellar Probe to the Heliospheric Boundary and Interstellar Space," Acta Astronautica, Vol. 35, Suppl., pp. 267-276, 1995.

[121] Noble, R. J., "Radioisotope Electric Propulsion of Sciencecraft to the Outer Solar System and Near-Interstellar Space," Acta Astronautica, Vol. 44, Nos. 2-4, pp. 193-199, 1999.

[122] Wallace, R. A., "Precursor Missions to Interstellar Exploration," Paper No. 114, IEEE Aerospace Conference, Big Sky, Montana, March 1999.

[123] Wallace, R. A., Ayon, J. A., and Sprague, G. A., "Interstellar Probe Mission/System Concept," Paper No. 53, IEEE Aerospace Conference, Big Sky, Montana, March 2000.

[124] Zurbuchen1, T. H., Prashant, P., Gallimore, A., Scheeres, D., Murphy, N., Zank, G., Malhotra, R., Funsten, H. and the NASA Interstellar Probe Vision Mission Team, "Interstellar Probe: Breakthrough Science Enabled by Nuclear Propulsion", Proceedings of the 55th International Astronautical Congress, Vancouver, October 2004.

[125] Chen-wan L., Yen, "Comparing Solar Sail and Solar Electric Propulsions for Propulsive Effectiveness In Deep Space Missions (AAS 01-214)", Advances in the Astronautical Sciences, Spaceflight Mechanics, Vol. 108, No. 2, pp 1587 – 1600, 2001.

[126] Leipold, M., Wagner, O., "Solar Photonic Assist Trajectory Design for Solar Sail Missions to the Outer Solar System (AAS 98-386)", Advances in the Astronautical Sciences, Spaceflight Dynamics, Vol. 100, No. 2, pp 1035 – 1046, 1998.

[127] McInnes, C.R., "Delivering Fast and Capable Missions to the Outer Solar System" Advances in Space Research, Vol. 34, Issue 1, pp. 184-191, 2004.

[128] French, J.R., Wright, J., "Solar Sail Missions to Mercury", Journal of the British Interplanetary Society, Vol. 40, pp 534-550, 1987.

[129] Leipold M., Borg, E., Lingner, S., Pabsch, A., Sachs, R., Seboldt, W., "Mercury Orbiter with a Solar Sail Spacecraft," Acta Astronautica, Vol. 35, Suppl., pp. 635-644, 1995.

[130] Leipold M., Borg, E., Lingner, S., Pabsch, A., Sachs, R., Seboldt, W., "Mercury Sun-Synchronous Polar Orbiter with a Solar Sail," Acta Astronautica, Vol. 39, No. 1-4, pp. 143-151, 1996.

[131] Leipold, M.E., Wagner, O., "Mercury Sun-Synchronous Polar Orbits Using Solar Sail Propulsion", J. Guidance, Control and Dynamics, Vol. 19, No. 6, pp 1337-1341, 1996.

[132] McInnes C. R., Hughes G., Macdonald M., "Low Cost Mercury Orbiter and Mercury Sample Return Missions Using Solar Sail Propulsion", The Aeronautical Journal, Paper No. 2790, pp. 469-478, August 2003.

[133] Hughes G., Macdonald M., McInnes C. R., Atzei, A., "Terrestrial Planet Sample Return Using Solar Sail Propulsion", Proceedings of 5th IAA International Conference on Low-Cost Planetary Missions, ESA Special Publication SP-542, pp. 377-384, September 2003.

[134] Hughes, G., Macdonald M., McInnes C. R., "Analysis of a Solar Sail Mercury Sample Return Mission", IAC-04-Q.2.B.08, Proceedings of 55th International Astronautical Congress, Vancouver, Canada, October 2004.

[135] Hughes G., Macdonald M., McInnes C. R., Atzei, A., Falkner, P., "Sample Return from Mercury and other Terrestrial Planets Using Solar Sail Propulsion", Submitted to Journal of Spacecraft and Rockets, February 2005.

[136] Leipold, M., "To the Sun and Pluto with Solar Sails and Micro-Sciencecraft", IAA-L98-1004, 3rd IAA International Conference on Low-Cost Planetary Missions, Pasadena, April 1998.

[137] Dachwald, B., "Optimal Solar Sail Trajectories for Missions to the Outer Solar System", Journal of Guidance, Control and Dynamics, Vol. 28, No. 6, pp. 1187-1193, 2005.

[138] McInnes, C.R., "Solar Sail Halo Trajectories: Dynamics and Applications", IAF-91-334, Proceedings of the 42nd International Astronautical Congress, Montreal, October 1991.

[139] Forward, R.L., "Statite: A Spacecraft That Does Not Orbit", Journal of Spacecraft and Rockets, Vol. 28, No. 5, pp 606-611, 1991.

[140] Carroll, K.A., "POLARES Solar Sail Feasibility Study", TM 39-306, Dynacon Enterprises Ltd, Ontario, October 1993.

[141] McInnes, C.R., "Advanced Trajectories for Solar Sail Spacecraft", International Conference Space Missions and Astrodynamics III, Turin, June 1994.

[142] McInnes, C.R., "An Examination of the Constant Polar Orbit: Discussion Document", DR-9809, Department of Aerospace Engineering, University of Glasgow, Glasgow, April 1998.

[143] 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.

[144] Mulligan, P., "Solar Sails for the Operational Space Community", Space Views, August 1998. Available from http://www.spaceviews.com/1998/08/article3a.html .

[145] Wie, B., "Dynamic Modelling and Attitude/Orbit Control System Design for Solar Sail Spacecraft", JPL Contract no. 1228156, Draft Interim Report, 10 July 2001.

[146] Murphy, D., Wie, B., "Robust Thrust Control Authority for a Scalable Sailcraft", AAS 04-285, Proceedings of the 14th AAS/AIAA Space Flight Mechanics Conference, Maui, Hawaii, February 2004.

[147] McInnes, C.R., Macdonald, M., Hughes, G., "Technology Roadmap", Technical Note 13(a), Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, December 2004.

[148] McInnes, C.R., Macdonald, M., Hughes, G., "Pluto-Kuiper Belt and Oort Cloud Exploration Missions", Technical Note 8, Technical Assistance in the Study of Science Payloads Transported Through Solar Sailing, ESA Contract Number: ESTEC 16534/02/NL/NR, December 2003.

[149] Sauer, C.G., "Optimum Solar Sail Interplanetary Trajectories", AIAA-76-792, AAS/AIAA Astrodynamics Conference, August 1976.

[150] Van der Ha, J. C., "The Attainability of the Heavenly Bodies with the Aid of a Solar Sail," German Society for Aeronautics and Astronautics, DGLR Paper 80-012, March 1980.

[151] Otten, M., McInnes, C.R., "Near Minimum-Time Trajectories for Solar Sails", Journal of Guidance, Control and Dynamics, Vol. 24, No. 3, 2001, pp. 632 – 634.

[152] Hughes, G., McInnes, C.R., "Small-Body Encounters Using Solar Sail Propulsion", Journal of Spacecraft and Rockets, Vol. 41, No. 1, pp. 140 – 150, 2004.

[153] Hughes, G.W., McInnes, C.R., "Solar Sail Hybrid Trajectory Optimization for Non-Keplerian Orbit Transfers", Journal of Guidance, Control and Dynamics, Vol. 25, No. 3, pp.602-604, 2002.

[154] Dachwald, B.,"Optimierung des Lernverhaltens Neuronaler Netze mit Hilfe Genetischer Algorithmen", Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Institut für Meßtechnik, Diploma Thesis, July 1993, in German.

[155] Dachwald, B., "Low-Thrust Trajectory Optimisation and Interplanetary Mission Analysis Using Evolutionary Neurocontrol ", Doctoral Thesis, Universität der Bundeswehr München, Fakultät für Luft- und Raumfahrttechnik, Institut für Raumfahrttechnik, April 2004.

[156] Ilgen, M.R., "Low thrust OTV guidance using Lyapunov optimal feedback control techniques", AAS Paper 93-680, Proceedings of AAS/AIAA Astrodynamics Specialist Conference, Victoria, Canada, August 1993.

[157] Chang, D.E., Chichka, D.F., Marsden, J.E., "Lyapunov functions for elliptic orbit transfer", Proceedings of AAS/AIAA Astrodynamics Specialist Conference, Québec City, July/August 2001.

[158] Petropoulos, A.E., "Low-Thrust Orbit Transfers Using Candidate Lyapunov Functions with a Mechanism for Coasting", AIAA 2004-5089, Proceedings of AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island, August 2004.

[159] Kluever, C. A., "Simple Guidance Scheme for Low-Thrust Orbit Transfers", Journal of Guidance, Control and Dynamics, Vol. 21, No. 6, pp.1015-1017, 1998.

[160] Schoenmaekers, J., Pulido, J., Jehn, R., "SMART-1 Mission Analysis: Moon Option", ESOC report S1-ESC-RP-5001, Issue 1, 22 September 1998.

[161] Gefert, L. P. and Hack, K. J., "Low-Thrust Control Law Development for Transfer from Low Earth Orbits to High Energy Elliptical Parking Orbits", AAS Paper 99-410, Proceedings of AAS/AIAA Astrodynamics Specialist Conference, Girdwood, Alaska, August 1999.

[162] Petropoulos, A.E., "Simple Control Laws for Low-Thrust Orbit Transfers", AAS Paper 03-630, Proceedings of AAS/AIAA Astrodynamics Specialist Conference, Big Sky, Montana, August 2003.

[163] Macdonald M., M°Innes C. R., "Realistic Earth Escape Strategies for Solar Sailing", Journal of Guidance, Control, and Dynamics, Vol. 28, No. 2, pp. 315-323, 2005.

[164] Macdonald, M., M°Innes, C.R., "Analytic Control Laws for Near-Optimal Geocentric Solar Sail Transfers", AAS 01-473, Advances in the Astronautical Sciences, Vol. 109, No. 3, pp. 2393-2413, 2001.

[165] Macdonald M., M°Innes C. R., "Solar Sail Capture Trajectories at Mercury", AIAA-2002-4990, Proceedings of AIAA/AAS Astrodynamics Specialist Conference, Monterey, August 2002.

[166] Macdonald M., M°Innes C. R., "Analytical Control Laws for Planet-Centred Solar Sailing", Journal of Guidance, Control, and Dynamics, Vol. 28, No. TBC, pp. TBC, 2005.

[167] Macdonald M., M°Innes C. R., "Analytical Control Laws for Heliocentric Solar Sail Orbit Transfers", Submitted to Journal of Spacecraft and Rockets, April 2005.

[168] Roy, A.E., "Orbital Motion", Institute of Physics Publishing, Bristol, 1998.

[169] Battin, R.H., "An Introduction to the Mathematics and Methods of Astrodynamics", AIAA Educational Series, New York, 1987.

[170] Green, A.J., "Optimal Escape Trajectories From a High Earth Orbit by Use of Solar Radiation Pressure", T-652, Master of Science Thesis, Massachusetts Institute of Technology, 1977.

[171] Brouwer, D., "On The Accumulation of Errors in Numerical Integration", The Astronomical Journal, Vol. 46, No. 16, pp. 149-153, 1937.

[172] Cowell, P.H., Crommelin, A.C.D., "Jupiter, satellite VIII, the orbit of", Monthly Note of the Royal Astronomical Society, Vol. 68, pp. 576, 1908.

[173] Cowell, P.H., Crommelin, A.C.D., "Halley's, perturbations of, in the past", Monthly Note of the Royal Astronomical Society, Vol. 68, pp. 665-670, 1908.

[174] Bate, R.R., Mueller, D.D., White, J.E., "Fundamentals of Astrodynamics", Dover Publications, Toronto, 1971.

[175] Kyner, W.T., Bennet, M.M., "Modified Encke Special Perturbation", The Astronomical Journal, Vol. 71, No. 7, pp. 579-582, 1966.

[176] Stumpff, S., Weiss, E.H., "A Fast Method of Orbit Computation", NASA Technical Note D-4470, April 1967.

[177] Kaula, W.M., "Theory of Satellite Geodesy", Blaisdell, Massachusetts, 1966.

[178] Broucke, R.A., Cefola, P.J., "On the Equinoctial Orbit Elements", Celestial Mechanics, Vol. 5, pp. 303-310, 1972.

[179] Allan, R.R., Proceedings of the Royal Society of London, A228, Vol. 60, 1965.

[180] Cefola, P.J., "Equinoctial Orbit Elements – Application to Artificial Satellite Orbits", AIAA 72-937, AIAA/AAS Astrodynamics Conference. Palo Alto, California, 1972.

[181] Walker, M.J.H., Ireland, B., Owens, J., "A Set of Modified Equinoctial Elements", Celestial Mechanics, Vol. 36, pp. 409-419, 1985.

[182] Betts, J. T., "Optimal Interplanetary Orbit Transfers by Direct Transcription", Journal of Astronautical Sciences, Vol. 42, No. 3, pp 247-268, July-Sept. 1994.

[183] Dormand, J.R., Price, P.J., "A Family of Embedded Runge-Kutta Formulae", Journal of Computing and Applied Mathematics, Vol. 6, pp 19 – 26, 1980.

[184] Science Software, "Celestial Computing with MATLAB", Littleton, Colorado, 1999.

[185] van der Ha, J.C., Modi, V.J., "Analytical Evaluation of Solar Radiation Induced Orbital Perturbations of Space Structures", The Journal of the Astronautical Sciences, Vol. 25, No. 4, pp. 283-306, 1977.

[186] Cook, G.E., "Luni-Solar Perturbations of the Orbit of an Earth Satellite", Geophysics Journal of the Royal Astronomical Society, Vol. 6, No. 3, pp 271-291, 1962.

[187] Pisacane, V.L., Moore, R.C., (Eds.), "Fundamentals of Space Systems", Oxford University Press, Inc., New York, 1994.

[188] Nacozy, P.E., Dallas, S.S., "The Geopotential in Nonsingular Orbital Elements", Celestial Mechanics, Vol. 15, pp. 453-466, 1977.

[189] Macdonald M., McInnes C. R., "Seasonal Efficiencies of Solar Sailing Planetary Orbit", IAC-02-S.6.01, Proceedings of 53rd International Astronautical Congress, Houston, October 2002.

[190] European Space Agency, "Science & Technology: INTEGRAL", INTEGRAL spacecraft, URL: http://sci.esa.int/science-e/www/area/index.cfm?fareaid=21 [cited 09 March 2004].

[191] McInnes, C.R., "Final Report: Solar Sail Option for the PROBA II Mission", European Space Agency (ESTEC) Contract Report, Contract No. 15003/00/NL/PB, September 2001.

[192] Delta IV Users Manual, available at URL: www.boeing.com [cited 01 April 2005].

[193] Isakowitz, S.J., Hopkins Jr., J.P., Hopkins, J.B., "International reference guide to space launch systems" 3rd Edition, American Institute of Aeronautics and Astronautics, Reston, 1999.

[194] Yuzhnoye Design Office, URL: www.yuzhnoye.com [cited 10 March 2004].

[195] Arianespace, "Vega User's Manual", Issue 0, Revision 0, available at URL: http://www.arianespace.com/site/images/VEGAUsersManual.pdf [cited 09 March 2004].

[196] Burt, E.G.C., "On Space Manoeuvres With Continuous Thrust", Planet Space Science, Vol. 15, pp 103-122, 1967.

[197] Burt, E.G.C., "The Dynamics of Low-Thrust Spacecraft Manoeuvres", Proceedings of the Eleventh Lanchester Memorial Lecture, 9 May 1968.

[198] Sharma, D. N., Scheeres, D.J., "Solar-System Escape Trajectories Using Solar Sails", Journal of Spacecraft and Rockets, Vol. 41, No. 4, pp. 684 – 687, 2004.

[199] Dachwald, B., "Interplanetary Mission Analysis for Non-perfectly Reflecting Solar Sailcraft Using Evolutionary Neurocontrol", AAS 03-579, Advances in the Astronautical Sciences, Astrodynamics, Vol. 116, No. 2, pp. 1247-1262, 2003.

[200] Dachwald, B., "Minimum Transfer Times for Nonperfectly Reflecting Solar Sailcraft", Journal of Spacecraft and Rockets, Vol. 41, No. 4, pp 693-695, 2004.

[201] Colasurdo, G., Casalino, L., "Optimal Control Law for Interplanetary Trajectories with Nonideal Solar Sail", Journal of Spacecraft and Rockets, Vol. 40, No. 2, pp 260-265, 2003.

[202] Mengali, G., Quarta, A. A. "Optimal Three-Dimensional Interplanetary Rendezvous Using Non-Ideal Solar Sail", Journal of Guidance, Control, and Dynamics, Vol. 28, No. 1, pp 173-177, 2005.

[203] Koblik V. V., Polyakhova E. N., Sokolov L. L., Shmyrov A. S., "Controlled Solar Sail Transfers to Near-Sun Orbits with Restrictions on Sail Temperature", Journal of the Astronautical Sciences, Vol. 47, No. 3-4, pp 189-206, 1999.

[204] Koblik V. V., Polyakhova E. N., Sokolov L. L., Shmyrov A. S., "Controlled Solar Sail Transfers into the Near-Sun Orbits", Advances in Space Research, Vol. 29, No. 12, pp. 2083-2088, 2002.

[205] Koblik V. V., "Solar Sail Motion in Near-Sun Regions", Doctoral Thesis, Tuorla Observatory, Univeristy of Turku, Finland, May 2003.

[206] Balthasar, H., Stark, D., Wöhl, "The Solar Rotation Elements i and Ω Derived from Recurrent Single Sunspots," Astronomy and Astrophysics, Vol. 174, pp. 359-360, 1987.