ADVANCED TRAJECTORIES FOR SOLAR SAIL SPACECRAFT

by

Colin R. McInnes B.Sc.

Thesis

Submitted to the

University of Glasgow

for the degree

of Ph.D.

Department of Physics and Astronomy,

The University,

Glasgow G12 8QQ

August 1991

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

To Karen, for the final chapter

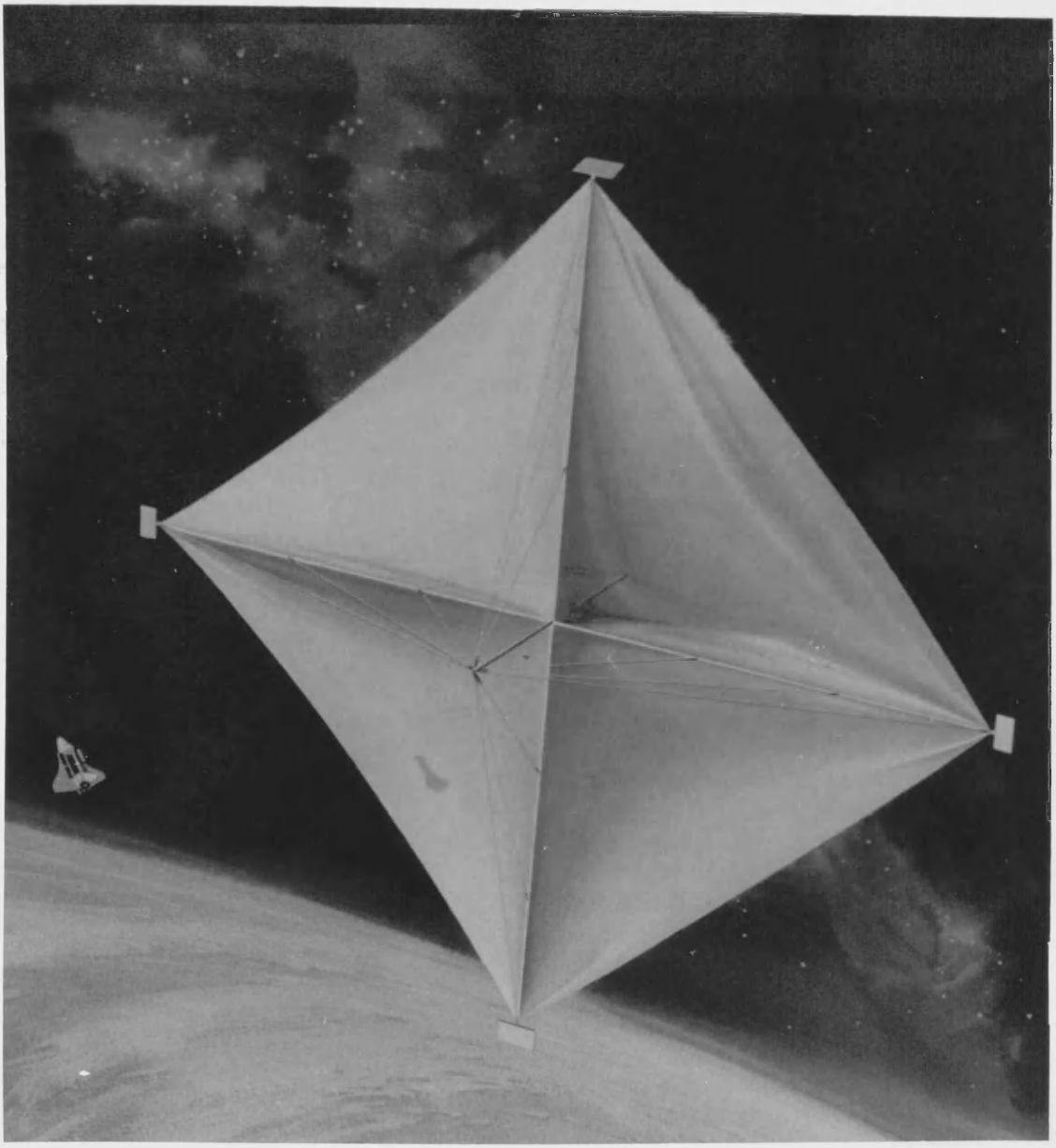
and my parents for the

opening pages

"Our traveller knew marvellously the laws of gravitation, and all the attractive and repulsive forces. He used them in such a timely way that, once with the help of a ray of sunshine, another time thanks to a cooperative comet, he went from globe to globe, he and his kin, as a bird flutters from branch to branch."

Voltaire - Micromegas, 1752

Frontispiece An artist's impression of an 800x800 metre solar sail designed in 1977 by NASA's Jet Propulsion Laboratory for a mission to rendezvous with comet Halley. The sail is shown fully deployed after being transported into Earth orbit by the shuttle orbiter. Reflective vanes at the corners of the sail are used to generate torques to control the attitude of the spacecraft, (NASA).



CONTENTS

PREFACE

The concept of solar sailing is by no means a recent innovation, having its origins in the writings of the Soviet pioneers of astronautics of the 1920s. The idea has been subject to detailed engineering scrutiny during design studies in the 1970s for a proposed rendezvous mission to comet Halley. Having been verified as a technologically viable means of spacecraft propulsion, solar sailing has matured to the extent that contemporary solar sail designs have the potential of remarkable performance.

With such high performance spacecraft, advanced new trajectories are attainable with unique applications for space science missions. The dynamics and applications of such trajectories are investigated in this thesis. For example, circular heliocentric orbits displaced high above the ecliptic plane are capable of continuous out-of-plane observations of the heliosphere. Similarly, circular geocentric orbits displaced along the Sun-Earth line offer new mission opportunities for the investigation of the geomagnetic tail. Since such trajectories and missions are unique to solar sail spacecraft they provide a strong motivation for the further development of solar sailing. Such development may then lead to the flight testing of solar sailing and to its eventual adoption for a major mission.

The original work of this thesis is contained within chapters 3-7. The contents of these chapters have appeared, or are to appear, as a series of papers in the ESA Journal, Journal of Spacecraft and Rockets, Celestial Mechanics and Acta Astronautica. Aspects of this research have also been presented at various conferences, meetings and seminars.

ACKNOWLEDGEMENTS

The research work for this thesis was carried out from 1988 to 1991 while the author was a research student in the Astronomy Group of the Department of Physics and Astronomy at the University of Glasgow. The funding was provided by a Robert Cormack Bequest Fellowship from the Royal Society of Edinburgh, to whom I am greatly indebted.

During the past three years many people have contributed to this thesis, either knowingly in scientific collaboration or quite unwittingly through moral support, providing brief diversions into the real world or nodding vigorously at strange ideas of sailing through space.

Amongst those who have contributed scientifically I must thank my two supervisors Prof. John Brown and Dr. John Simmons for ideas, enthusiasm and encouragement. Thanks for believing there was something in solar sailing John, and thanks for the relativistic interstellar trips, John. My scientific thanks goes to Keith Macpherson for his collaboration with the work of section 3.5 and to James Cluckie for his collaboration on the mission analysis codes. Outwith the university many solar sail enthusiasts have contributed, in particular; Bob Forward, formerly of the Hughes Research Laboratories, Malibu; Elena Poljakhova of the University of Leningrad and Alexandre Shvartsburg of the Central Design Bureau for Unique Instrumentation in Moscow. My thanks also goes to Rodney Buckland of Brunel University for his infectious enthusiasm for solar sailing and for the opportunity to indoctrinate future generations of space scientist with my own ideas.

Aside from the scientific input many people have contributed in

other no less important ways (in no particular order); Martin, for philosophical exchanges most Mondays along Maryhill road; Christos, for lessons on the Greek language and culture (especially for the Malachas warning); Alec, Lyndsay and Keith for some memorable ventures into the stratosphere (above 3000 ft, of course); Moray and Richard, for providing a bench mark on minimum working days; Alan, for unpaid (at least yet) computer consultancy; David, for the telepathic transfer of the skills of scientific paper writing; Daphne, for unrivalled secreterial capabilities; Geoff, for the bale-out on practical astronomy; Kenny, for a daily glimpse of the real world and Craig, for a semi-regular dosage of pubs, movies and philosophical excursions into the meaning of life, (but not thankfully the universe and everything).

Lastly, yet by no means least, my thanks go to my pal, companion, shadow and wife-to-be Karen, for love and friendship. In this respect it was in fact Karen who developed most of the theory of halo orbits - the fact that my name appears on this thesis is a mere editorial oversight. I only hope that I can contribute as much to the understanding of the molecular biology of cis-Diamminedichloroplatimum (II), (known colloquially as plain 'cisplatin'), in the coming years as you have contributed to astrodynamics in the past three.

Colin M^{c}Innes

University of Glasgow

August 1991

SUMMARY

In this thesis the dynamics and applications of new advanced trajectories for solar sail spacecraft are investigated. By utilising the continuously available solar radiation pressure force exotic non-Keplerian trajectories which are unique to solar sail spacecraft are developed. Although large families of these new trajectories are found to be unstable, simple feedback control schemes are designed to ensure asymptotic stability. The unique nature of these trajectories opens up new space science mission opportunities, impossible for conventional spacecraft. These missions are shown to offer the possibility of interesting new observations of many aspects of solar system physics.

In chapter 1 the concept of solar sailing is introduced and the history of its development reviewed. The fundamental design parameters for solar sail spacecraft are investigated and the fabrication of potential sail materials discussed. Square and heliogyro type solar sails are then described and their relative merits discussed. Modern state-of-the-art sail design and performance is then considered. Previous studies of solar sail mission applications are reviewed and recent concepts for future advanced applications discussed.

The fundamental physics of solar radiation pressure is considered in chapter 2 along with the modelling of the solar radiation pressure force exerted on a real solar sail. Heliocentric solar sail trajectories are investigated and in particular the logarithmic spiral trajectory is derived. The limitations of such trajectories and the necessity of time optimal trajectories satisfying the two point boundary conditions of an

interplanetary transfer trajectory are then indicated. The dynamics of geocentric escape trajectories are investigated and the use of optimised trajectories is again discussed.

The solar radiation pressure model used in chapter 2 is expanded in chapter 3 to include astrophysical effects which have implications for solar sail dynamics. By the use of the radiation pressure tensor the effect of the finite angular size of the solar disk on the functional form of the solar radiation pressure force is obtained. The resulting deviation of the solar radiation pressure force from an inverse square variation with heliocentric distance is shown to have a de-stabilising effect on solar sails in stationary and circular orbital configurations. The effect of small time variations in the solar luminosity is also considered.

In chapter 4 the first family of advanced solar sail trajectories is investigated. By obtaining stationary solutions to the heliocentric dynamical equations in a co-rotating reference frame the conditions for heliocentric halo orbits are obtained. These orbits are circular heliocentric orbits displaced out of the ecliptic plane by a component of the solar radiation pressure force exerted on the sail. Using a linear perturbation analysis the stability characteristics of the system are investigated and unstable families of trajectories found. Simple feedback control schemes are then obtained to ensure asymptotic stability. Lastly, by patching individual halo orbits and Keplerian orbits together, elaborate new patched trajectories are shown to exist.

A similar analysis is used in chapter 5 to investigate geocentric halo orbits, which are near polar circular orbits displaced along the Sun-line by the solar radiation pressure force. Unstable families are again found with feedback control schemes developed to ensure

asymptotic stability. The fundamental family of geocentric halo orbits is enlarged by the patching of individual halo orbits and Keplerian orbits.

In chapter 6 artificial stationary solutions to the circular restricted three-body problem are considered. The addition of the solar radiation pressure force leads to extensions of the five classical stationary points to a family of extended stationary surfaces. For the Earth-Sun system these new solutions are truly time independent. However, for the Earth-Moon system small trims in the sail area are required to compensate for the rotation of the Sun-line during the synodic lunar month. Again, the stability and control of the system is investigated and the instability of the solutions demonstrated. An out-of-plane trajectory at the lunar L_2 point is also developed.

The application of these new trajectories for potential space science missions is discussed in chapter 7. The scientific benefit of out-of-plane observations for solar system physics is explored and potential missions investigated. A simple, twin solar sail mission utilising heliocentric halo orbits is described. Similarly, the utilisation of geocentric halo orbits for geomagnetic tail observations is also investigated. Several applications of the new three-body stationary solutions are considered, such as the use of payload transfer from Lagrange point stationary surfaces to eliminate lengthy planetary spiral trajectories.

Lastly, the conclusions of chapter 8 outline further possible developments in the areas of dynamics, control and mission analysis. Recommendations for the progress of advanced solar sail missions are given.

1. INTRODUCTION TO SOLAR SAILING

1.1 Solar Sailing

For all of its short thirty year history practical spacecraft propulsion has been dominated by the unaltering principles of Newton's third law. All forms of propulsion from the solid rocket motor to the solar-electric ion drive rely on a reaction mass which is accelerated by some exothermal or electromagnetic means into a high velocity jet. A unique and elegant means of propulsion which transcends this reliance on reaction mass is the solar sail. Of course solar sail spacecraft must also obey Newton's third law. However, solar sails gain momentum from an ambient source, namely photons, the quantum packets of energy of which sunlight is composed.

The momentum carried by an individual photon is vanishingly small. Therefore, in order to intercept large numbers of photons solar sails must be vast structures. Furthermore, to generate as high an acceleration as possible from the momentum of the intercepted photons solar sails must be extraordinarily light. For a practical solar sail the mass per unit area of the whole spacecraft must be an order of magnitude less than the paper this text is written upon. Not only must solar sails be vast in area and small in mass, they must also be near perfect reflectors. Then the momentum transferred to the sail can be almost double the momentum of the incident photons and the solar radiation pressure force is directed normal to the sail surface. Therefore, by controlling the orientation of the sail relative to the incoming radiation the sail can then gain or loose momentum so that its orbit can be controlled.

The picture then is clear. A solar sail is a vast sheet of

reflective film, typically the size of a football field, stretched taught over a lightweight structure. Using the momentum gained from ambient photons the spacecraft is slowly, but continuously accelerated to begin any number of possible missions. Solar sailing, with its analogies with terrestrial sailing may be a romantic notion. However, as will be shown in this thesis, the romanticism is overshadowed by the immense practicability and quiet efficiency with which solar sails can be put to use for scientific investigation.

1.2 Historical Perspective

Although solar sailing has been considered as a practical means of spacecraft propulsion only relatively recently, the fundamental ideas are by no means new. The actual concept of solar sailing has a long and rich history dating back to the Soviet pioneers of astronautics. As early as the 1920s the Soviet father of astronautics Konstantin Tsiolkovsky and his co-worker Fridrickh Tsander both wrote of 'using tremendous mirrors of very thin sheets' and 'using the pressure of sunlight to attain cosmic velocities', Tsiolkovsky (1921) and Tsander (1924). The origins of the work of Tsiolkovsky and Tsander can however be traced back as far as the seventeenth century.

1.2.1 Theories of Solar Radiation Pressure

In 1619 Johannes Kepler proposed qualitatively that the material of comet tails was pushed outward from the Sun due to some pressure from sunlight. At this time the corpuscular theory of light was the favoured view of optics and the outward pressure due to sunlight was a natural consequence of this theory, Lebedew (1902) (and references contained within). Newton, a strong proponent of the corpuscular

theory accepted Kepler's view as a possible explanation but attempted in 1687 to explain the phenomenon within his theory of universal gravitation. He made the hypothesis that there was an ambient ether denser than the material of comet tails. Therefore, the repulsion was due to buoyancy forces and the Sun had only an attractive gravitational force. Sometime later in 1744 Euler returned to Kepler's view. However, Euler adopted the longitudinal wave theory of light due to Huygens. With this theory Euler was able to show that a longitudinal wave would exert a repulsive force on a body in its path. Later still, as Coulomb's experiments of 1785 with electrostatics became known, Olbers in 1812 rejected all previous explanations of comet tail repulsion and proposed that the Sun had a nett electrical charge. Particles leaving the comet nucleus had then to be charged with the same sign as the solar charge. The fact that electrostatic forces have an inverse square variation (as does solar radiation pressure) supported Olber's theory which then became prevalent.

The theoretical basis for the existence of radiation pressure came independently of the astronomical theories. Maxwell predicted the existence of radiation pressure in 1873 as a consequence of his theory of electromagnetic radiation. Apparently independently, Bartoli demonstrated the existence of radiation pressure as a consequence of the second law of thermodynamics. The experimental verification of the existence of radiation pressure and the verification of Maxwell's quantitative results then came in 1901 when Lebedew finally suceeded in isolating the radiation pressure force using a series of torsion balance experiments, Lebedew (1902). A detailed account of Lebedew's, and others, experimental work is given by Nichols and Hull (1903).

1.2.2 The Development of the Solar Sail

After the initial writings of Tsiolkovsky and Tsander in the 1920s the concept of solar sailing appears to have remained dormant, bar a few science fiction stories, for over thirty years. It was not until 1958 that Richard Garwin, of the IBM Watson laboratory at Columbia University, re-examined solar sailing in the light of modern materials technology, Garwin (1958). Garwin recognised the unique and elegant features of solar sailing. Namely that solar sails require no propellant and are continuously accelerated, therefore allowing large velocity changes over extended periods of time. Such was Garwin's enthusiasm and optimism for solar sailing that his analysis showed a round trip to Venus was possible in less than one year using commercially available thin film.

Following the discussion of solar sailing by Garwin more detailed studies of the dynamics of solar sails were undertaken by Tsu. By approximating the heliocentric equations of motion Tsu was able to show that, for a fixed sail attitude, solar sail trajectories were of the form of spirals with a fixed opening angle, Tsu (1959). A simple comparison with chemical and ion propulsion systems showed that solar sails could match, and in many cases out perform, these systems. The approximations used by Tsu were removed by London, who solved the heliocentric equations of motion exactly and obtained a true logarithmic spiral solution, London (1960). It was recognised that since the logarithmic spiral trajectory did not satisfy the boundary conditions of an interplanetary transfer, large velocity impulses were required at the departure and terminal points. This problem was later addressed in studies of time optimal trajectories. London also explored the case of a sail oriented normal to the Sun-line such that the solar radiation

pressure force reduced the effective solar gravitational force. The resulting family of modified conic section solutions were found, in some cases, to yield shorter transfer times than the logarithmic spiral, as discussed later by Kiefer (1965). These early studies stimulated interest in solar sailing and led to more detailed analyses, for example Villers (1960), Gordon (1961), Cotter (1973). However, these studies were at a low level and lacked a specific solar sail mission.

In the early 1970s the development of the space shuttle promised the prospect of being able to transport and deploy large volume payloads in orbit. Also, the technological development of deployable space structures and thin films suggested that solar sailing could be considered for a specific mission. By 1973 NASA was funding studies of solar sailing at the Battelle laboratories, Wright (1974), which gave positive recommendations for further investigation. During the continuation of this work Jerome Wright discovered a trajectory which would allow a solar sail rendezvous with comet Halley at its perihelion in the mid 1980s. The flight time of only four years would allow for a late 1981, early 1982 launch. Until then a difficult rendezvous mission was thought to be impossible in such a short time using the technology of the day. A seven to eight year mission had been envisaged using solar-electric ion propulsion. These positive results prompted NASA Jet Propulsion Laboratory director Bruce Murray to initiate an engineering assessment study of the potential readiness of solar sailing. Following the assessment a formal proposal was put to NASA in September 1976. The design of a comet Halley rendezvous mission using solar sailing was initiated in November of the same year, Friedman et. al (1978).

In the initial design study an 800x800 m three-axis stabilised

square sail design was considered, but was dropped in May 1977 due to the associated high risk deployment. The design work then focused on a spin stabilised heliogyro type sail. The heliogyro, which was to use twelve 7.5 km long blades of film rather than a single sheet, had been developed by Richard MacNeal under NASA contract ten years earlier, MacNeal (1967). The heliogyro could be easily deployed by simply unrolling the individual blades. As a result of the design study the structural dynamics and control of the heliogyro were understood and potential sail materials were manufactured and characterised. Also important for NASA institutional considerations, the solar sail work had sparked public interest and excitement in the comet Halley rendezvous mission.

As a result of the interest in solar sailing the solar-electric propulsion group re-evaluated their performance estimates and had in the end been directly competing with the solar sail group for funding. A detailed account of this internal competition is given by Logsdon (1989). As a result of an evaluation of these two advanced propulsion concepts NASA chose the solar electric system in September 1977, on its merits of being a less, but still considerable risk for a comet Halley rendezvous. A short time later the solar electric system was also dropped, as eventually was a dedicated NASA comet Halley mission.

Although dropped by NASA for near term mission applications, the design studies of the mid-1970s stimulated worldwide interest in solar sail technology. Low level European studies were taken up by CNES in Toulouse to assess the potential of the Ariane launcher for deep space missions, Riviere et. al (1977). Perhaps more importantly for the long term prospects of solar sailing was the formation of the World Space Foundation in California and the Union Pour la Promotion de

Propulsion Photonique in Toulouse. The WSF, formed principally of JPL engineers shortly after the termination of the JPL solar sail work, attempted to raise private funds to continue solar sail development and to undertake a small scale demonstration flight. Shortly afterwards the U3P group was formed and in 1981 proposed an ambitious moon race to promote solar sail technology. Both of these groups were joined by the Solar Sail Union of Japan in 1985 and have proposed that the moon race be adopted as a project for the international space year in 1992, Perret et. al (1989). More recently the Columbus Quincentenial Jubilee commission, formed to organise celebrations of the quincentenary of Columbus discovering the new world, have been attempting to stimulate interest in a solar sail race to Mars. The proposal has generated international interest with some of the most technically advanced and innovative sail designs to date, for example Johnson et. al (1989), Fox et. al (1989) and von Flotow et. al (1989). At the time of writing most, if not all, of the international solar sail groups are having extreme difficulty in raising the necessary funds. However, the many new solar sail designs and organisations involved have brought about a renaissance in solar sailing not seen since the comet Halley studies of almost twenty years ago. Where this leads to for the future of solar sailing remains to be seen.

1.3 Solar Sail Design

The main considerations in the engineering design of a solar sail spacecraft are driven primarily by the mission application and launch vehicle. The volume of the payload bay of the launcher dictates the maximum sail area and the mass allocation dictates the maximum overall

spacecraft mass. In terms of mission analysis, the spacecraft acceleration determines the transfer time to a particular target, or even whether a particular mission is possible. Secondary considerations such as attitude and thermal control are driven by the trajectory design. That is, rapid attitude manoeuvres may be required for planetocentric trajectories or the heliocentric phase may take the sail close to the Sun.

1.3.1 Fundamental Design Parameters

The crucial design parameter for a solar sail spacecraft is the characteristic acceleration. This is defined as the acceleration of a sail oriented normal to the Sun-line at a heliocentric distance of one astronomical unit. At this distance the magnitude of the solar radiation pressure P is 4.57×10^{-6} Nm^{-2} (this value will be derived in section 2.1). Therefore, allowing for a finite sail efficiency η, the characteristic acceleration is given by

$$a_0 = \frac{2\eta P}{\sigma} \quad , \quad \sigma = \frac{m}{A} \tag{1.1}$$

where σ is the solar sail mass per unit area, with m the total spacecraft mass and A the total area. The sail efficiency is a function of the optical properties of the sail material and the sail shape. The total mass may be divided into two components, m_s due to the sail material and structure and m_p the payload mass. If the sail mass is further divided into the mass of the sail material m_{sl} and the structural mass then $m_s = m_{sl}(1+k)$ where k is the structural mass as a fraction of the sail mass, typically of order 0.5 or less. Therefore, the characteristic acceleration becomes

$$a_o = \frac{2\eta P}{\sigma_{sl}(1+k) + (m_p/A)} \qquad (1.2)$$

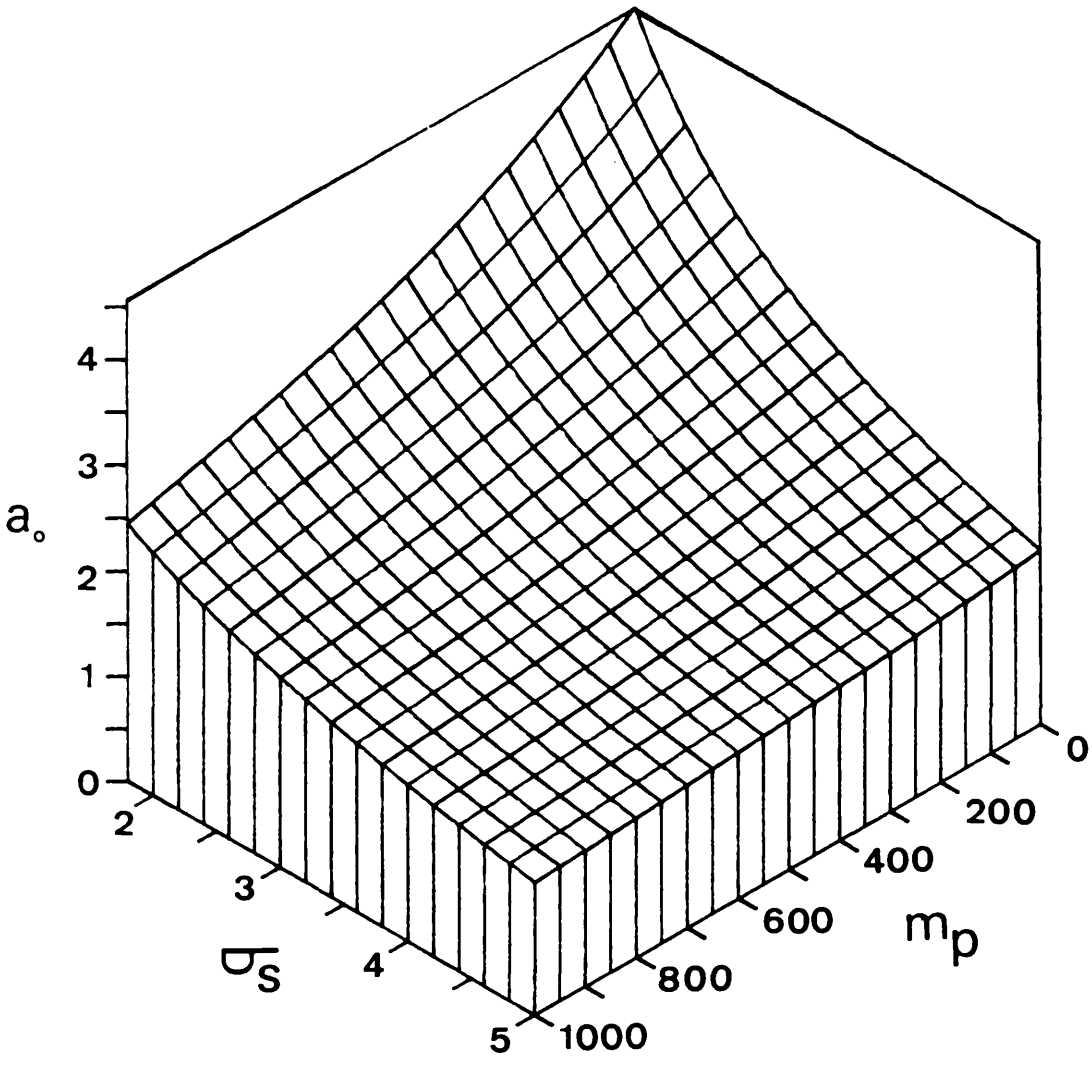
where σ_{sl} is the mass per unit area of the sail material. Using equation (1.2) the performance of a practical solar sail may now be analysed.

For a fixed sail area and efficiency equation (1.2) becomes a function of the total (sail material and structure) mass per unit area σ_s and the payload mass m_p. Therefore, surfaces of characteristic acceleration may be generated, Figure 1.1. For an 800x800 m square sail with η=0.9 a characteristic acceleration of order 1 mms^{-2} is possible with a 10^3 kg payload and $\sigma_s \approx 5$ gm^{-2}. Such an acceleration level, as designed for the JPL comet Halley sail using 1977 technology parameters, is suitable for most interplanetary missions. It can be seen that for a large value of σ_s the sail characteristic acceleration is relatively insensitive to variations in payload mass. This is due to the sail mass exceeding the payload mass so that the term $\sigma_{sl}(1+k)$ dominates in equation (1.2). Similarly, for a large payload mass the characteristic acceleration is relatively insensitive to variations in the technology level of the sail, (ie. σ_{sl} and k). Therefore, to obtain a high characteristic acceleration of 5-6 mms^{-2}, as required for some advanced trajectories, it is clear from Figure 1.1 that not only must the sail material and structure be light but it is crucial that the mass of the payload and onboard hardware must be kept to a minimum. Inverting equation (1.2) an expression for the payload mass is obtained, viz

$$m_p = A\left\{\frac{2\eta P}{a_o} - \sigma_{sl}(1+k)\right\} \qquad (1.3)$$

Therefore, for a fixed characteristic acceleration and sail mass per

Figure 1.1



Surface of sail characteristic acceleration (mms^{-2}) as a function of payload mass (kg) and sail mass per unit area (gm^{-2}).

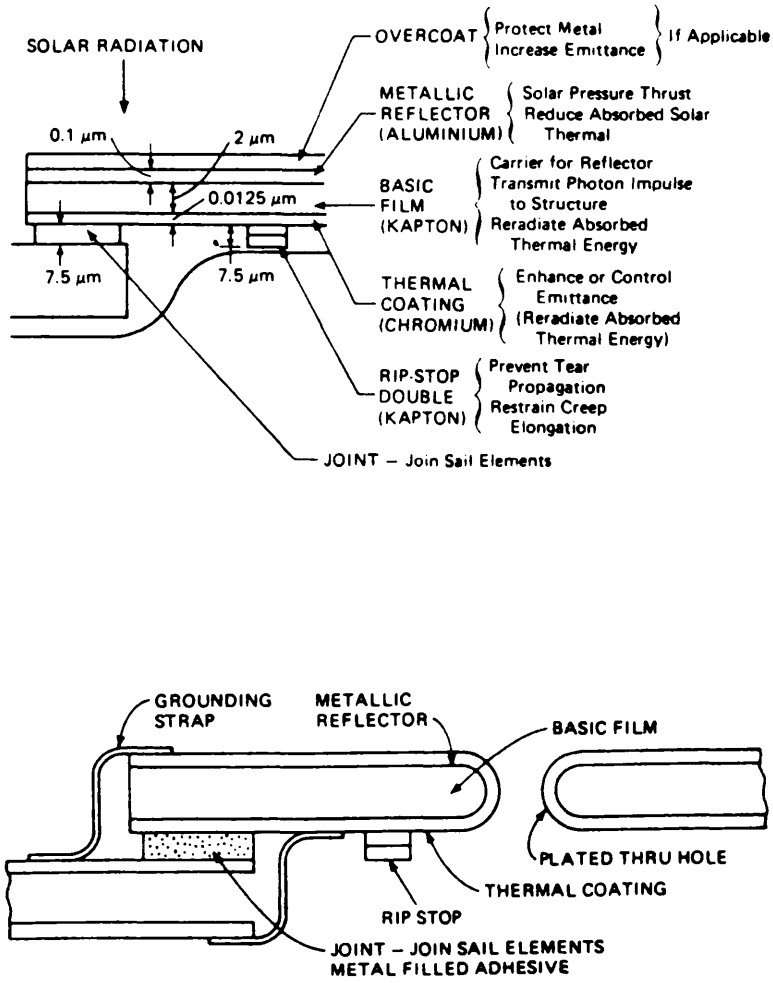
unit area any increase in payload mass must be offset with a corresponding proportional increase in sail area.

1.3.2 Sail Material Design

Aside from the fundamental solar sail design parameters there are many secondary considerations which must be addressed. Firstly, a suitable material must be chosen for the actual sail. The material must have a large enough tensile strength so that when fully deployed and under tension the material does not fail and create tears which may propagate through the entire sail. Furthermore, since the reflective surface of the sail will not be perfect a small fraction of the incident solar radiation will be absorbed. This absorbed energy must be dissipated through a thermally emitting rear surface. Since the absorbed energy will increase the material temperature there must be dimensional stability so that the sail has low thermal expansion and shrinkage. Once deployed in orbit the sail material must be free of wrinkles which may cause multiple reflections and intense hot spots. The sail shape must be simulated using a finite element method in combination with the sail structural dynamics to calculate the actual shape due to billowing. Multiple reflections can then be traced and eliminated.

A cross section of a potential sail material developed during the JPL comet Halley studies is shown in Figure 1.2. The sail has a 2 μm Kapton plastic substrate upon which a thin 0.1 μm Aluminium layer is deposited. This substrate allows the sail to be folded and packed into a small volume for launch and to be safely unfurled in orbit. The rear surface of the sail has a 0.0125 μm Chromium coating for thermal control of the sail. The Kapton film is directly manufactured at 2 μm

Figure 1.2



Section of 2 μm aluminised Kapton sail film, as developed for the JPL comet Halley rendezvous mission, (Friedman 1988).

thickness giving the sail material a mass per unit area of 2.9 gm^{-2}. Individual sheets of sail material are bonded using a high speed heat sealing technique and a suitable adhesive. More recent studies and laboratory tests have investigated the use of plasma etched Kapton giving a mass per unit area of only $1gm^{-2}$, Fox et. al (1989). Unetched bands can be formed to increase the tensile strength of the material. Potentially lighter sail materials may also be manufactured using Lexan film, von Flotow et. al (1989), and Polyimide, Johnson et. al (1989).

If in-orbit manufacturing of sail material were to become possible vast improvements in sail performance would be achieved. Since the sail material would not be folded and packed a plastic substrate would not be required. The construction of such sails has been considered by Drexler (1979). Using thin film techniques samples of 10-100 nm sail film were produced. Small triangular panels of sail material would be produced and bonded to construct the entire sail.

In order that any tears due to failure of the sail material, such as micrometeorite impacts, do not propagate the sail must be provided with ripstops at regular intervals. These are formed from either Kapton tapes joined to the rear sail surface or from double folds in the sail forming seams, Figure 1.2. Furthermore, since the sail is exposed to the solar wind the sail material will acquire a differential electrical charge between the front and rear surfaces. This is due to the incident proton flux from the solar wind and the photoelectric effect, Hilliard and Whipple (1985). To prevent electrical discharges from the front to rear of the sail, which is a potential source of failure and tearing, both surfaces of the sail must be in electrical contact, Figure 1.2. Finally, although sail materials have been

designed in detail and samples fabricated the material lifetime in the space enviroment is uncertain. The material is exposed to hard solar radiation and cosmic rays which will almost certainly degrade the sail surface reflectivity and reduce tensile strength. These questions will however only be addressed with detailed in-orbit testing of samples. Preliminary testing has recently been carried out onboard the Soviet Almaz satelltie, Shvartzburg (1991).

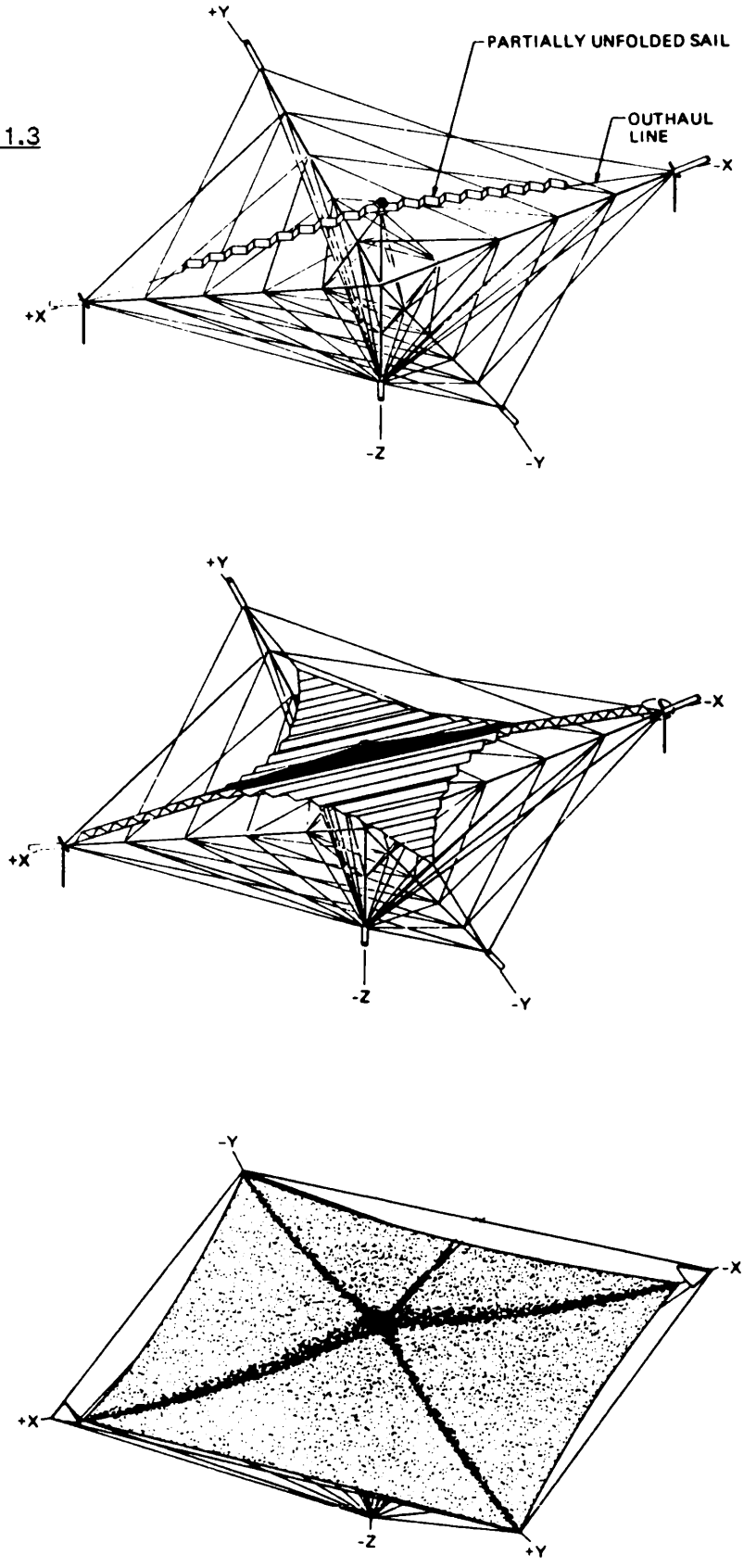
1.4 Archetypal Solar Sail Configurations

During the JPL comet Halley studies two individual solar sail configurations were investigated in detail. Firstly, a three-axis stabilised square sail design and laterly a spin stabilised heliogyro. Although fundamentally different in their design, construction and control it was found that their overall mass per unit area differed by a few per cent only. The square sail and heliogyro differ greatly however in their method of deployment. It was due to the simpler deployment of the heliogyro that it was selected for the comet Halley rendezvous.

1.4.1 Three-Axis Stabilised Square Sail

The square sail configuration uses a single sheet of sail material which is kept in tension using diagonal spars extending from a central boom normal to the sail surface on the reflective side. The spars are connected to the boom with stays to reduce structural loads, Figure 1.3. For the 800x800 m JPL comet Halley sail design the spars were 1.2m diameter open lattice structures, constructed of Titanium to prevent thermal expansion. Due to the lattice nature of the structure the spars could be coiled for storage. The stays were flat tapes, to

Figure 1.3



Two stage deployment of a single square sail film sheet over a fully deployed boom and stay structure, (Friedman 1988).

prevent snarling during deployment.

For automated in-orbit deployment the central boom and diagonal spars are extended from packaging canisters. During this extension the stays are pulled along the spars from storage reels. Once the entire structure is erect the sail itself is deployed. The sail is firstly pulled into a ribbon along a diagonal spar and then pulled flat along the perpendicular diagonal, Figure 1.3. If the sail were constructed from four individual triangular sheets the same two stage procedure would take place for each element. Although this deployment scheme has been shown to be feasible the large number of serial operations leads to a large number of potential failures. Since the scheme could not be tested on the ground prior to implementation, even partially, the deployment had a high element of risk.

The three-axis attitude control of the square sail may be provided by several techniques. One method is to use a shift in the centre of mass relative to the centre of pressure. This can be achieved by small displacements in the payload position, using deflections in the central boom. This technique has recently been studied in detail by Angrilli and Bortolami (1990). However, the use of small reflective vanes at the sail corners provides the greatest control. Using combinations of vanes arbitrary roll, pitch and yaw torques may be generated, even when the main sail is in a null attitude edge on to the Sun-line. It can be shown that suitable control can be achieved with vanes which rotate about axes along the spars to which they are attached. Lastly, the entire sail may be displaced across the structure to shift the centre of pressure, while the centre of mass remains fixed. The sail is displaced by reeling in outboard support lines while reeling out the opposite set. This is

perhaps the simplest technique and least costly in terms of hardware mass.

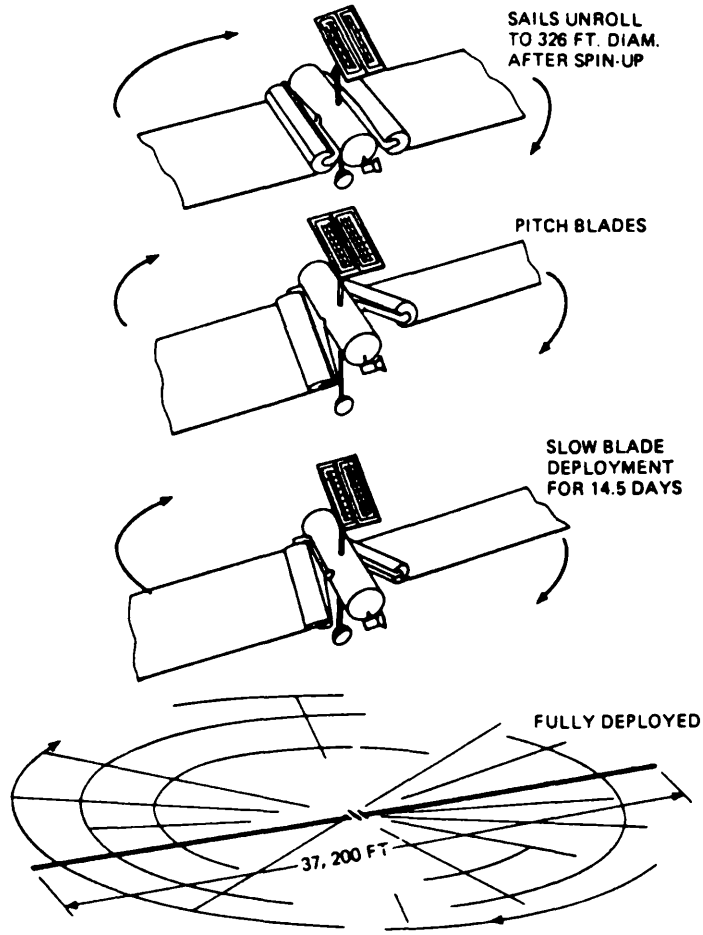
1.4.2 Spin-Stabilised Heliogyro

Whereas the square sail relies on the three-axis stabilisation of a large single sail the heliogyro has several long blades of sail material, rotating to provide tension and spin stabilisation. The JPL comet Halley heliogyro consisted of twelve blades each 8 m wide and 7.5 km in length. The blades are joined to a central hub containing the payload and are pitched to provide attitude control. Centrifugal loads are carried in tension members at the edges of the blades. These tension members are constructerd from several tapes of high tension 2 mm wide graphite Polyimide fibres. The tapes have periodic crossovers where they are bonded to provide multiple load paths. This gives each tension member the ability to withstand several failures. If a blade were to become detached there is a high probability of impact with the remaining blades leading to a catastrophic failure. The blades have transverse battens in compression, spaced along their length to provide torsional stiffness.

The deployment of the heliogyro is a simpler procedure than for the square sail. An initial spin-up is provided by small thrusters. Then a set of rollers allow the blades to unwind, with the central hub directed along the Sun-line and the blades held at the same collective pitch angle. The radiation pressure torque generated by the blades adds angular momentum to the system allowing the blades to fully extend, Figure 1.4.

The heliogyro attitude may be controlled using the blade pitch in a cyclic or collective manner to alter the spin axis of the vehicle. A

Figure 1.4



Deployment of a single blade heliogyro. The initial spin is provided by thrusters, with the solar radiation pressure force exerted on the blades adding momentum for full deployment, (Friedman 1988).

collective pitch of the blades is used to spin up or spin down the vehicle whereas the cyclic pitch is used to generate asymmetric forces to provide attitude control torques. The pitch drive is provided by motors on the central hub. Although the attitude turning rates are smaller than those of the square sail the magnitude of the radiation pressure force may be modulated by up to 20 per cent. A recent innovative approach by von Flotow et. al (1989) has considered the possibility of using piezo-electric actuators with shorter lightweight blades.

The most serious problem with the heliogyro is the possibility of dynamic instability. Various instabilities due to thermo-elastic and purely mechanical effects can occur, although blade flutter appears potentially the most serious, MacNeal (1971). This is due to a photo-dynamic coupling between the blade pitch and blade bending. Flutter may however be avoided by separating the frequencies of the first few torsional and bending modes of the blades. Therefore, although more dynamically complex the heliogyro is controllable and is a more reliable spacecraft to deploy.

1.4.3 State-of-the-Art Solar Sail Design

The JPL square and heliogyro sails are the only two designs to have undergone detailed engineering studies. However, a number of recent sail designs prompted by the proposed solar sail races have the promise of greater performance than either of the JPL designs. In particular a 276 m disc sail developed by Cambridge Consultants Ltd is perhaps the most advanced sail yet designed, Johnson et. al (1989). The CCL sail is in the form of a circular disc of sail material, supported on a structure of 36 radial carbon fibre reinforced plastic

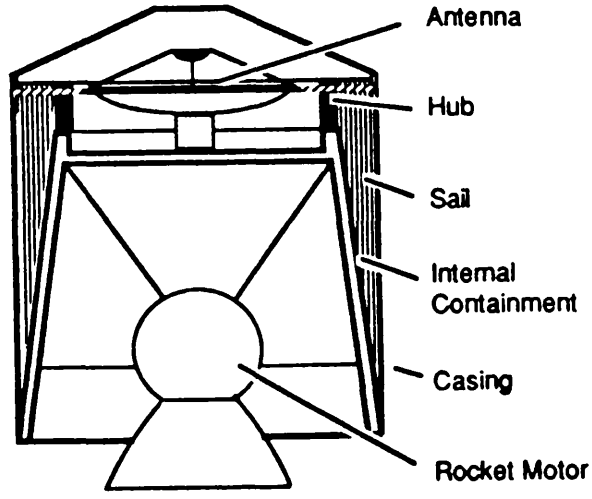
spars which are cantilevered from a central load bearing hub. A unique spiral fold allows both the sail and spars to be packaged into a compact form for launch.

When constructed the sail and spars are wound around the central hub in a unique packing arrangement. Therefore, the sail is easily deployed by allowing the 36 spars to elastically unwind from the central hub, progressively stretching the sail, Figure 1.5. The rate of deployment is controlled by the viscous forces of initially unstretched polymer threads linking the spars at regular intervals. These threads stretch visco-elastically as the sail deploys limiting the deployment rate and so the angular velocity of the hub. Otherwise the hub would gain a large amount of angular momentum which would create a mechanical shock as it was dissipated through the vehicle once the deployment was complete.

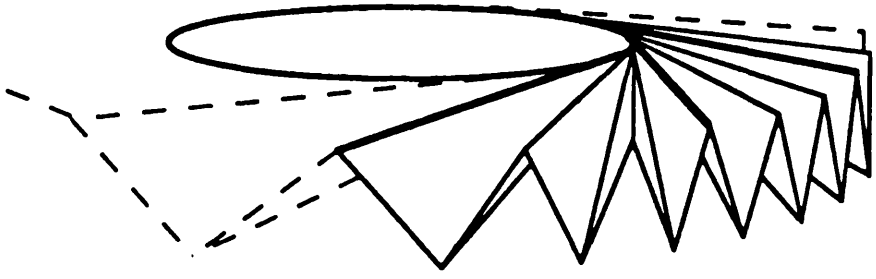
The sail attitude may be controlled by inducing small distortions in the entire sail shape by actuating the spars. The spar actuation is achieved by applying bending moments at the hub using thermo-elastic bracing wires connected to the spars. By applying differential electrical heating to these wires the sail geometry can be altered from a flat disk to a cone for passive stability, or a saddle for active attitude manoeuvering. With spar bending of less than one degree only, extremely rapid attitude control may be achieved. The sail spin rate may also be controlled by bending alternate spars to form a turbine geometry.

Using the baseline design accelerations of order 2 mms^{-2} may be achieved with a 60 kg payload. However, since the sail structural design leads to low stresses on the sail much thinner materials than the 2 μm Polyimide film proposed may be used. By plasma etching the

Figure 1.5



Secondary Space Segment Assembly



Partially deployed sail

Packing and deployment of the CCL disc sail. The structural spars are wound around the central hub during launch. Once released the spars unwind elastically deploying the sail, (Johnson et. al 1989).

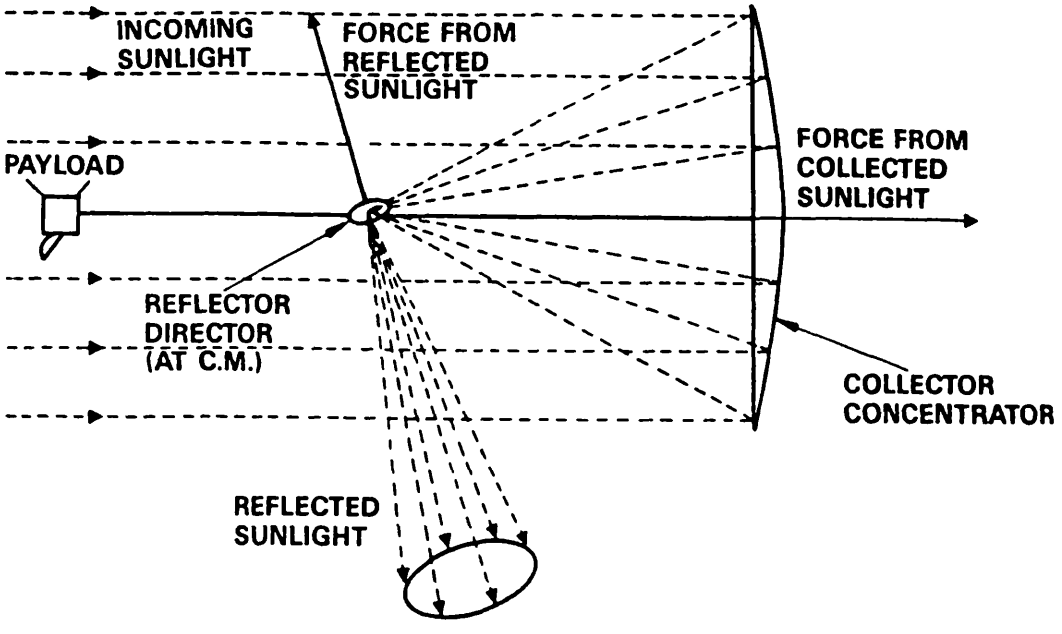
rear surface large improvements in sail performance and characteristic acceleration would be possible, Oswald (1990). Such a high performance vehicle would be suitable for advanced solar sail trajectories.

Finally, the solar photon thruster, another variant on the planar type solar sail has recently been discussed, Forward (1989a). The concept was originally proposed in the Soviet literature by Malanin and Repyakh (1974). The solar photon thruster is constructed from a large Sun pointing reflector which directs the solar radiation to a small movable secondary reflector, Figure 1.6. By separating the functions of collecting and directing the solar radiation, the solar photon thruster has a significant performance improvement over other sail designs at large pitch angles. However, as yet no detailed engineering design studies have been undertaken for this type of sail.

1.5 Solar Sail Mission Applications

Now that the design, deployment and control of various solar sail types have been discussed, possible mission applications will now be investigated. Since solar sail spacecraft require no propellant they offer considerable advantages for transporting large payloads and for round trip sample return missions. Furthermore, since the solar sail has a virtually unlimited ΔV capability multiple mission objectives such as asteroid surveys are possible. In fact both of these advantages are inherent in the solar sail interplanetary shuttle concept, Wright and Warmke (1976). By using a square sail inherited from the JPL comet Halley vehicle with autonomous on board systems it was proposed that multiple payloads could be transported to bodies within the solar system. After the payload delivery the sail would return

Figure 1.6



A solar photon thruster sail. The main reflector directs radiation to a small steerable secondary reflector at the centre of mass of the system, (Forward 1989a).

to Earth orbit for refurbishment and a new set of payloads.

In operational terms a solar sail may be used to simply augment the performance of a launch vehicle upper stage by taking the payload to escape from an initial high Earth orbit, or the sail may be used for all the propulsion beyond low Earth orbit. Escape spirals must however begin from altitudes above 700 to 1000 km due to atmospheric drag on the sail. From these altitudes escape times may be of the order of months so that it is desirable to place the sail on a parabolic escape trajectory by conventional means. Although Earth orbit is not the optimal operating environment for solar sails an orbital transfer vehicle using solar sail technology has been investigated, Teeter (1977).

Unlike ballistic transfers the solar sail has an essentially open launch window. The trajectory may be continuously modified for any launch date. For planetary rendezvous short spiralling capture times are possible in the inner solar system, however at the outer planets these times are prohibitively long due to the greatly diminished solar radiation pressure. Payloads would be inserted into planetary orbit using burns from space storable propellants, with the sail using a gravity assist to return to the inner solar system in a shorter time than it took for the outward trip.

1.5.1 Inner Solar System Missions

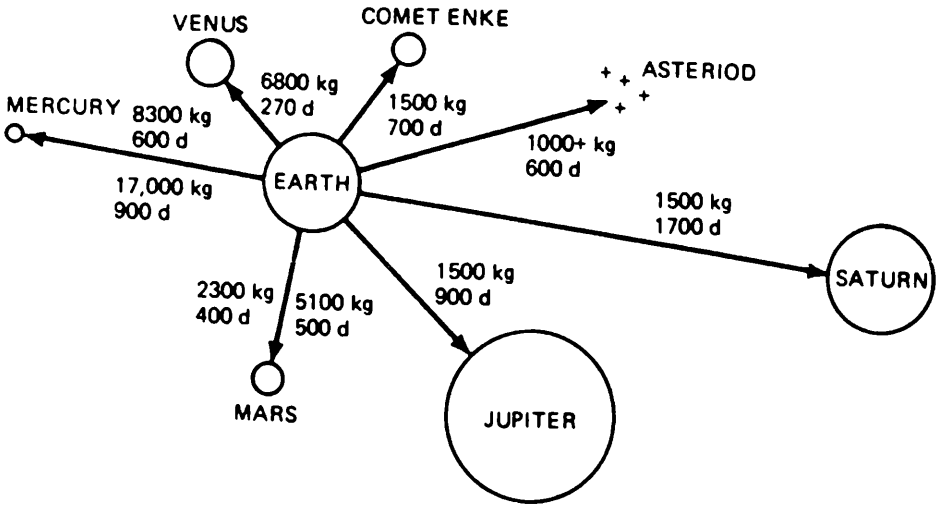
Due to the increased solar radiation pressure in the inner solar system solar sail spacecraft can deliver large payloads to close, high inclination heliocentric orbits. The payload would be delivered by spiralling inwards while directing a component of the solar radiation pressure force out of the ecliptic plane to increase the orbital

inclination. Before the desired inclination was reached however, the sail would have spiralled to its closest heliocentric distance, determined by the thermal tolerance of the sail material (typically 0.3 AU). The sail would then use a cranking orbit, alternately directing the solar radiation pressure force above and below the ecliptic plane, to achieve the final desired inclination. Moderate performance sails with characteristic accelerations of 1 mms^{-2} can deliver 10^3 kg payloads into a 0.3 AU polar orbit in 750 days (including escape from Earth orbit), Wright and Warmke (1976).

Other inner solar system missions, such as a round trip to Mercury, have spectacular possible payload masses. For example, a $1.7x10^4$ kg payload may be delivered into orbit around Mercury in 900 days, Figure 1.7. Such a mission could include both surface landers for a sample return and permanent orbiters, Wright and French (1987). A sample return from Mercury using conventional spacecraft technology is at best extremely difficult and costly in terms of propellant mass. The ballistic two-impulse transfer to Mercury requires a ΔV of 17.4 kms^{-1}, although this can be reduced somewhat using gravity assists.

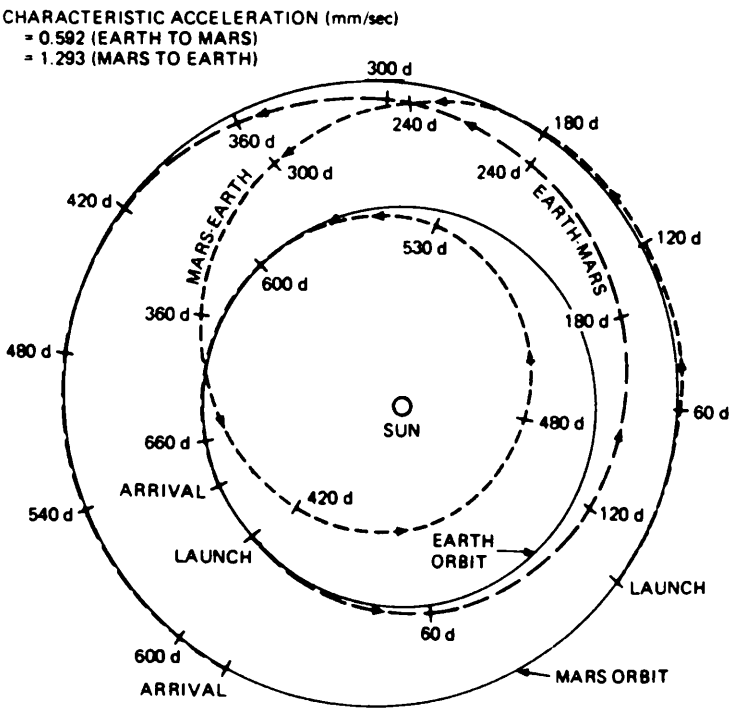
For missions to Mars the outward transfer times tend to be somewhat longer than for ballistic transfers. However, the solar sail transfer is not constrained by the 1.2 year waiting period of the ballistic transfer. Therefore, for a waiting time at Mars of a few months the total round trip time is of the same order as that for ballistic transfers, Figure 1.8. On the outward trip over $5x10^3$ kg may be delivered in 500 days. The payload delivered is however different than that from a ballistic transfer. For a solar sail the entire payload can be used for obtaining a sample from the Martian surface,

Figure 1.7



Payload capabilites and transfer times for an 800x800 m square sail with a 2.5 μm Kapton sail, (Friedman 1988).

Figure 1.8



Time optimal round trip Earth-Mars trajectory for 1988. Departure is on 1/5/88 with arrival at Mars on 10/1/90. The return trajectory begins on 11/6/90 with return to Earth on 13/4/92, (Friedman 1988).

whereas for the ballistic mission the delivered payload must include propellant for the return trajectory. It has been shown that a 200 kg sample could be returned using a solar sail vehicle in contrast to less than a 1 kg sample using a ballistic mission, Wright and Warmke (1976). Other studies have shown that large sails may greatly reduce the total mass required in low Earth orbit for a manned mission to Mars, Staehle (1982). In particular it has been shown that advanced sails may be constructed and deployed from tethered facilities at the international space station, Garvey (1987).

1.5.2 Outer Solar System Missions

Due to the diminished solar radiation pressure in the outer solar system insertion of payloads into planetary orbit would be achieved using a conventional chemical propulsion system, or aerobraking. Payloads of 1.5×10^3 kg may be delivered to Jupiter and Saturn, with transfer times of 900 and 1700 days respectively, allowing useful masses to be placed in orbit. For solar system escape a 1.5×10^3 kg payload may be taken as far as 30 AU in 3000 days, allowing long baseline astrometric measurments to be made, Wright and Warmke (1976). Time optimal trajectories to the outer planets usually include an inward spiral close to the Sun to gain momentum before the outward trajectory. The use of gravity assists at the outer planets to reduce transfer times would be of limited use due to the high approach velocity. Transfer times may however be reduced by increasing the hyperbolic arrival velocity at the target.

Again, due to the virtually unlimited ΔV capability of solar sail spacecraft multiple asteroid rendezvous missions are possible, as are asteroid sample returns. A total round trip time of 1400 days to Eros

is possible with a $1.5x10^3$ kg payload, Wright and Warmke (1976). Similarly for high eccentricity, high inclination comets rendezvous and sample returns are possible at the expense of extended mission durations.

1.5.3 Advanced Solar Sail Missions

Due to the continually available solar radiation pressure force solar sail spacecraft are capable of many advanced missions impossible for any other spacecraft type. Although many of these missions require advanced, high performance solar sail design others are possible using the technology level of the JPL comet Halley sail.

The delivery of payloads to close heliocentric orbit was discussed in section 1.5.1. Although the comet Halley sail was constrained to operate at greater than 0.25 AU the studies of sail materials technology indicated that sustained operation within 0.2 AU would be possible. A sail with a specialised thermal coating may survive the thermal environment at only 0.06 AU (15 solar radii), Wright (1990). At these extremely close heliocentric distances it has been proposed that the solar radiation pressure force may be used to suitably modify the spacecraft orbital period. In particular a 25 day period would allow active regions near the solar equator to be tracked across the solar disk, Forward (1986). Furthermore, for an advanced sail it would be possible to choose the total mass per unit area so that the solar radiation pressure force exactly balanced the solar gravitational force (σ=1.53 gm^{-2}). This would allow the spacecraft to remain stationary above the solar poles, making continuous observations, Drexler (1979), Forward (1986).

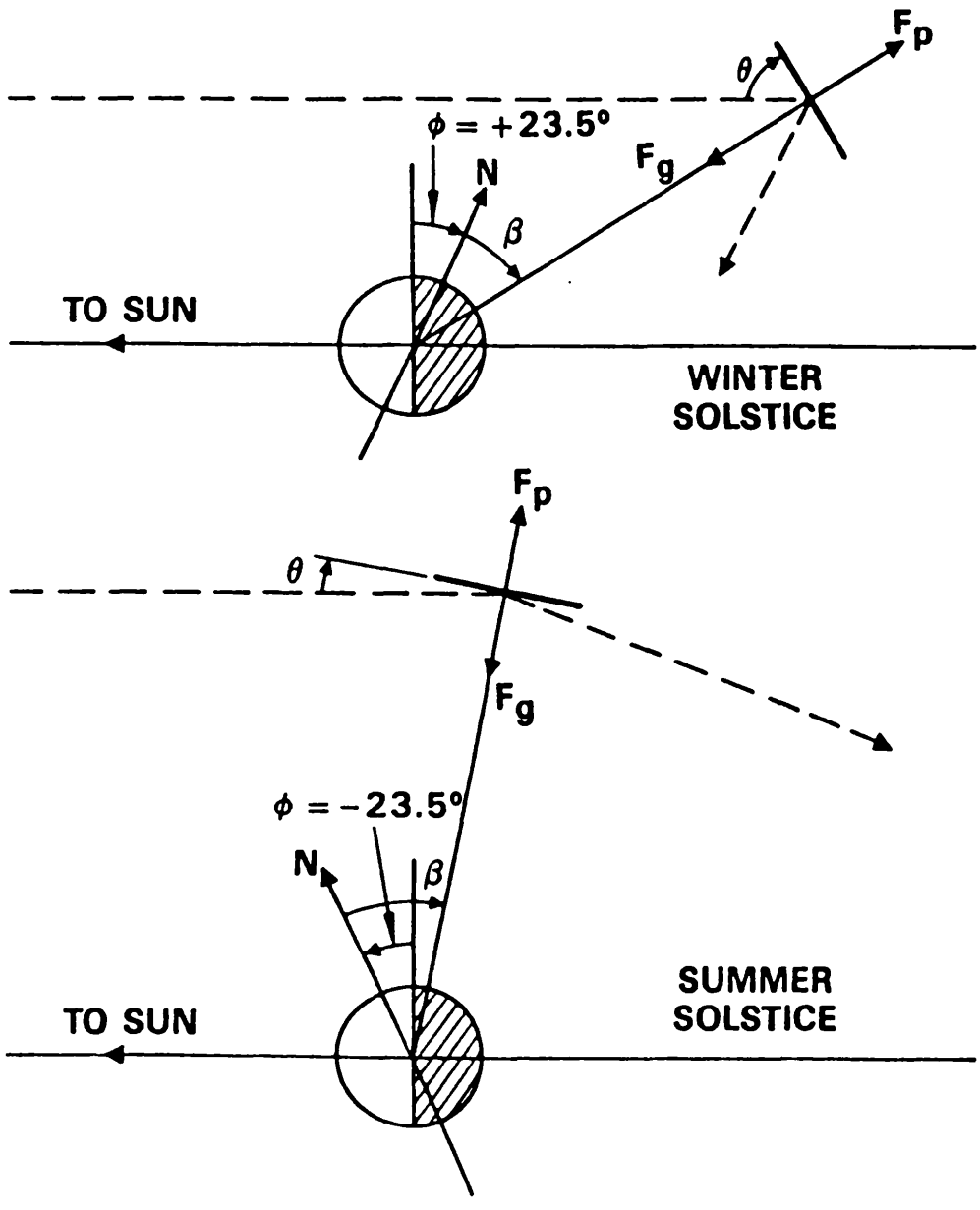
For geocentric applications it has been demonstrated that solar

sails may be used to displace communication satellites above and below geostationary orbit, Forward (1981,1984). This would allow satellites to be stacked above and below the equatorial plane, greatly increasing the number of available locations. Angular separations in the sky of 0.3^{o} are possible using 1 μm Kapton sails with a payload mass of one third of the sail mass. Furthermore, communications payloads may be placed in static equilibrium high above the night side of the Earth allowing line of sight viewing of the spacecraft from high lattitude regions, Forward (1989b). For conventional geostationary satellites the spacecraft appear low on, or below, the horizon in high lattitude and polar regions. This static equilibrium 'statite' concept requires large geocentric distances of order 50 Earth radii or greater so that the local gravitational acceleration is low, Figure 1.9. However, since the spacecraft would appear to rotate about the pole star once per day, clock drives would be required for tracking.

1.6 Overview

Although solar sailing is by no means a new concept it has undergone many detailed engineering design studies by various groups. As a result of these studies it has been demonstrated that solar sailing is a viable means of spacecraft propulsion and has significant capabilites for near term missions. By operating without propellant a solar sail may deliver significantly larger payloads than could be achieved by other means. The vehicle may then still return to Earth orbit, perhaps with planetary soil samples, and be refurbished for subsequent missions. Solar sailing may also provide the means to accomplish extremely difficult missions, such as rendezvous with high eccentricity, high inclination comets.

Figure 1.9



Static equilibrium 'statite' concept for communication services to high lattitude regions. For a fixed polar angle β the altitude at the summer solstice must be greater than the winter solstice due to the oblique incidence of the solar radiation, (Forward 1989b).

For current advanced solar sail designs new unique and advanced applications become possible. Several of these applications, such as the displacement of geostationary satellites, have a useful near term application. Others, such as static equilibiria above the solar poles, require sail performance levels between two and three times that of current designs. These designs can however be improved upon with further development. In particular the use of plasma etching on sail material offers the potential for vast improvements. However, as discussed in section 1.3.1, with a low mass per unit area sail high performance can only be obtained with a minimised payload mass. Therefore, the application of other areas of advanced space technology such as lightweight solar cells and high strength, low mass composite structural materials must be considered.

Finally, in analysing solar sail missions and applications comparison must be made with other viable propulsion schemes, such as advanced solar-electric propulsion. While solar-electric propulsion was ultimately chosen for the comet Halley rendezvous due to its greater technological maturity it is not suitable for advanced missions requiring a continuously available thrust. That is, a solar-electric propulsion system may generate the required thrust magnitude but the mission lifetime will be finite due to its dependence on propellant. For this reason the possibilites for the large scale development of solar sailing relies on the investigation of advanced missions and applications which are unique to solar sail spacecraft.

2. SOLAR SAIL ORBITAL DYNAMICS

2.1 Solar Radiation Pressure

The source of motive force for solar sail spacecraft is the momentum transferred to the sail by radiative energy from the Sun. Using the electromagnetic description of light the momentum is transported to the sail by electromagnetic radiation. Physically, the electric field component **E** of the incident electromagnetic wave generates a current **J** in the sail surface. This current then couples to the magnetic field component of the wave **B** to generate a Lorentz force **JxB** in the direction of propagation of the wave. The induced current then generates another electromagnetic wave which is observed as the reflected component of the incident wave.

Alternatively, solar radiation pressure can be envisaged as being due to the momentum transferred to the sail by photons. Using the photon representation the magnitude of the solar radiation pressure can be calculated from the specific intensity of the radiation field, $I_\nu(\mathbf{r},\mathbf{n};t)$, (see appendix A). The radiation pressure tensor $\mathbf{P}(\mathbf{r};t)$ is then defined as the second angular moment of the specific intensity of the radiation field integrated over the entire frequency spectrum, viz

$$\mathbf{P}(\mathbf{r};t) = \frac{1}{c}\int_0^\infty \oint_{4\pi} I_\nu(\mathbf{r},\mathbf{n};t)\ \mathbf{nn}\ d\Omega d\nu \qquad (2.1)$$

If the sail heliocentric distance $r \gg R_\odot$, the solar radius, the photons arrive at the sail surface along approximately parallel rays in direction **k**. The specific intensity of the radiation field can then be defined as $I_\nu(\mathbf{r},\mathbf{n}) = I_\nu{}^{\odot}\Theta_\epsilon(\mathbf{n}-\mathbf{k})$, assuming the Sun to be a time independent source. The function $\Theta_\epsilon(\mathbf{n}-\mathbf{k})$ is a unit step function with a small, finite width

ϵ. Therefore, within a solid angle ϵ of direction **k** all photons arrive along parallel rays. Substituting for the specific intensity in equation (2.1) the radiation pressure tensor then becomes

$$\mathbf{P}(r) = \frac{1}{c}\,\mathbf{kk}\int_0^\infty I_\nu^0\, d\nu \oint_{4\pi} \Theta_\epsilon(\mathbf{n}-\mathbf{k})\, d\Omega \tag{2.2}$$

For a surface element $d\Sigma$ of the solar disk the solid angle element may be written as $d\Omega=d\Sigma/r^2$. Therefore, integrating over the solar disk it is found that

$$\mathbf{P}(r) = \frac{I_0}{c}\frac{\pi R_0^2}{r^2}\,\mathbf{kk} \tag{2.3}$$

where I_0 is the frequency integrated specific intensity. For a planar solar sail the sail area may be represented as a directed surface **A=An**, Figure 2.1. Therefore, the force exerted on the sail due to the incident photons is given by

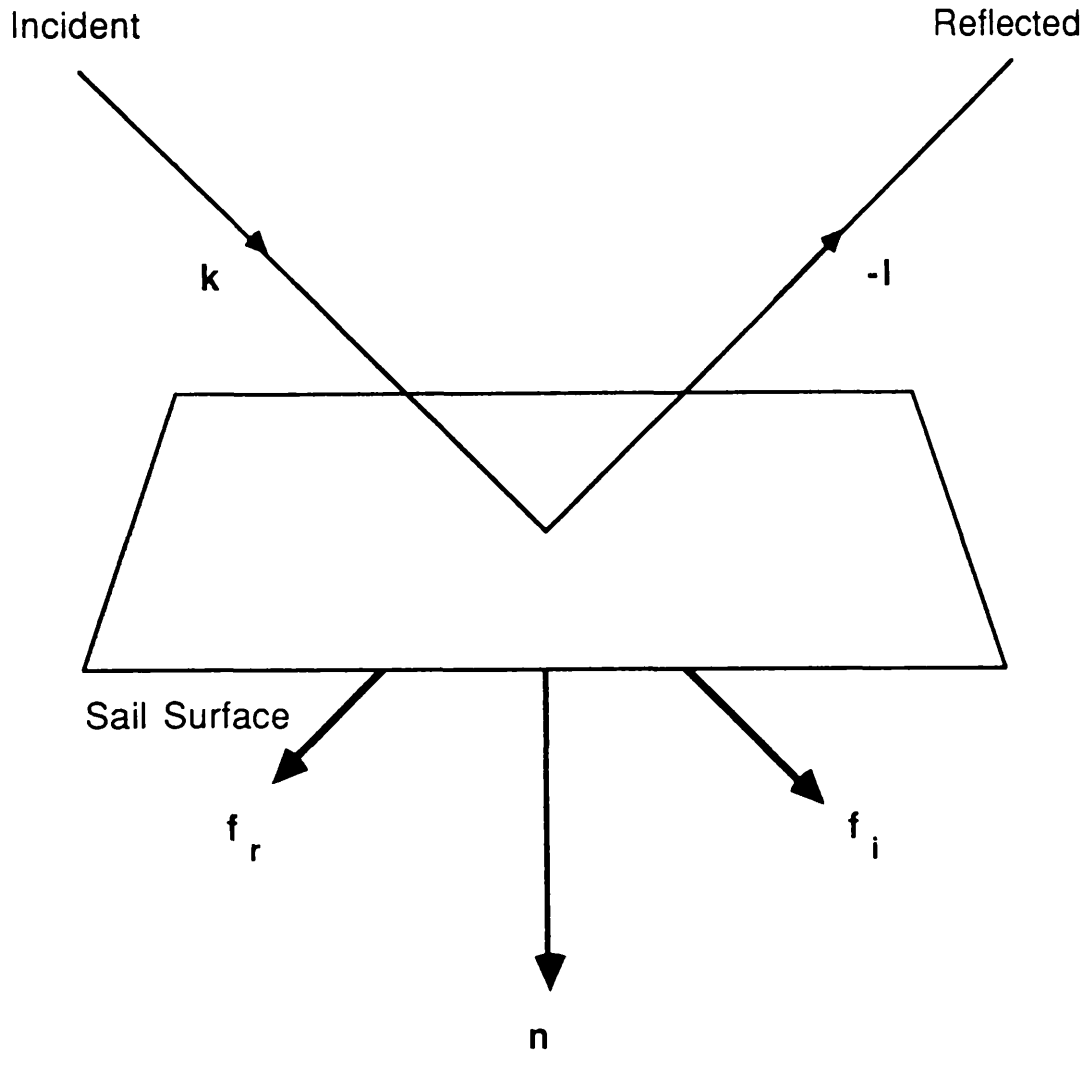
$$\mathbf{f}_i = \frac{I_0}{c}\frac{\pi R_0^2}{r^2}\,A(\mathbf{k}.\mathbf{n})\,\mathbf{k} \tag{2.4a}$$

If the sail is now considered to be a perfect reflector the component of the photon momentum along the sail normal is reversed so that the reaction force due to the reflected photons is given by

$$\mathbf{f}_r = \frac{I_0}{c}\frac{\pi R_0^2}{r^2}\,A(\mathbf{k}.\mathbf{n})\,\mathbf{l} \tag{2.4b}$$

The total radiation pressure force exerted on the sail is then given by the sum of the incident and reflected components. Using the vector relation $(\mathbf{k}+\mathbf{l})=2(\mathbf{n}.\mathbf{k})\mathbf{n}$ the total force may be written as

Figure 2.1



Solar radiation incident on a solar sail from direction **k** and reflected in direction **-l**. The sum of the force due to incident radiation $\mathbf{f}_i$ and the force due to reflected radiation $\mathbf{f}_r$ generates a total radiation pressure force normal to the sail surface in direction **n**.

$$\mathbf{f} = \frac{2I_o}{c} \frac{\pi R_o^2}{r^2} A(\mathbf{k}.\mathbf{n})^2 \mathbf{n} \qquad (2.5)$$

so that the solar radiation pressure force has an inverse square variation with heliocentric distance. In fact the solar radiation pressure force deviates from an inverse square form, as will be discussed in chapter 3.

The frequency integrated specific intensity I_o may be obtained by calculating the energy flux at the sail surface. The flux is defined as the first angular moment of the specific intensity, (see appendix A)

$$F(\mathbf{r};t) = \int_0^\infty \oint_{4\pi} I_\nu(\mathbf{r},\mathbf{n};t)\, \mathbf{n}\, d\Omega d\nu \qquad (2.6)$$

However, the flux F at the sail surface is given in terms of the solar luminosity L_o simply by $L_o/4\pi r^2$. Therefore, equation (2.6) gives

$$\frac{L_o}{4\pi r^2} = \int_0^\infty I_\nu{}^o \, d\nu \oint_{4\pi} \frac{d\Sigma}{r^2} \qquad (2.7)$$

Integrating over the solar disk the frequency integrated specific intensity is found to be $I_o = L_o/4\pi^2 R_o^2$. Therefore, substituting for I_o in equation (2.5) and dividing by the total spacecraft mass m the sail acceleration is found to be

$$\mathbf{a} = \frac{2}{c} \left\{\frac{L_o}{4\pi r^2}\right\} \frac{1}{\sigma} (\mathbf{k}.\mathbf{n})^2 \mathbf{n} \qquad (2.8)$$

where σ is the total spacecraft mass per unit area. The magnitude of the solar radiation pressure is therefore F/c, which is 4.57×10^{-6} Nm^{-2} at a heliocentric distance 1 AU.

For a solar sail in heliocentric orbit the direction of incidence of

the radiation **k** is given by the unit radial vector $\mathbf{r}/|\mathbf{r}|$. Furthermore, the sail acceleration may be made dimensionless with respect to the solar gravitational acceleration $\mu/|\mathbf{r}|^2$, where μ is the product of the solar mass M and gravitational constant G. The radiation pressure acceleration may then be conveniently written as

$$\mathbf{a} = \beta \frac{\mu}{|\mathbf{r}|^4}(\mathbf{r}.\mathbf{n})^2 \mathbf{n} \quad , \quad \mathbf{r}.\mathbf{n}>0 \tag{2.9}$$

where the dimensionless sail loading parameter β is defined to be the ratio of the solar radiation pressure acceleration to the solar gravitational acceleration. Since both the solar radiation pressure acceleration and the solar gravitational acceleration have an inverse square variation the sail loading parameter is independent of the sail heliocentric distance. Using equation (2.8) the sail loading parameter may be written as

$$\beta = \frac{\sigma_*}{\sigma} \quad , \quad \sigma_* = \frac{L_0}{2\pi GMc} \tag{2.10}$$

where the critical mass per unit area $\sigma_*=1.53\ gm^{-2}$. With this mass per unit area $\beta=1$ so that the sail radiation pressure acceleration is equal to the local solar gravitational acceleration.

If the assumption of perfect reflectivity is relaxed a more exact model of the solar radiation pressure acceleration may be constructed. Taking into account the reflectance, transmittance and emittance of the sail material the radiation pressure acceleration may be written as

$$\mathbf{a} = \beta \frac{\mu}{|\mathbf{r}|^4}(\mathbf{r}.\mathbf{n}) \left\{\Lambda_1 \mathbf{r} + \left\{\Lambda_2|\mathbf{r}| + \Lambda_3(\mathbf{r}.\mathbf{n})\right\}\mathbf{n}\right\} \quad , \quad \mathbf{r}.\mathbf{n}>0 \tag{2.11}$$

where the sail material parameters Λ_j (j=1,3) are defined as

$$\Lambda_1 = 2(1 - \Lambda_3 - \tau) \tag{2.12a}$$

$$\Lambda_2 = \left\{\lambda_1(1 - \lambda_2) + \frac{\kappa}{3}(1 - \lambda_1 - \tau)\right\} \operatorname{sgn}\left\{\mathbf{n}.\frac{\mathbf{r}}{|\mathbf{r}|}\right\} \tag{2.12b}$$

$$\Lambda_3 = \lambda_1\lambda_2 \tag{2.12c}$$

with the auxiliary parameter κ is defined as a function of the sail temperature by

$$\kappa = \frac{e_1T_1^4 - e_2T_2^4}{e_1T_1^4 + e_2T_2^4} \tag{2.12d}$$

The parameter λ_1 represents the total fraction of incident radiation reflected from the sail while λ_2 represents the fraction of that radiation which is specularly reflected. The front and rear sail emissivities and temperatures are given by e_1, T_1 and e_2, T_2 with τ representing the fraction of incident radiation transmitted through the sail, Van der Ha and Modi (1977a). For state of the art sail materials these parameters have yet to be determined experimentally. However, for a perfectly reflective sail the parameters are given simply by Λ_3=1 and Λ_1=Λ_2=0.

During the JPL solar sail studies for the comet Halley mission even more precise models were developed using finite element simulations of the sail shape combined with experimental data on potential sail materials. These models may be conveniently represented as trigonometric series in the sail attitude angles and have been used in trajectory analysis software, for example Sackett (1977). Although the solar radiation pressure force may be empirically parameterised by the above means, in-flight calibration is required in practice for accurate guidance and control, Jacobson and Thornton (1978).

Although solar radiation pressure generates the largest force on

solar sails, other secondary forces are also present. For example, the solar wind will exert a small pressure on the sail due to the momentum transported by solar wind protons. During periods of high solar wind speed the mean proton density ρ at 1 AU is approximately 4×10^6 m^{-3} with a wind speed v of 700 kms^{-1}, Schwartz (1985). The pressure exerted on the sail can then be estimated from the momentum transport as

$$P_w \approx m_p \rho v^2 \qquad (2.13)$$

where m_p is the proton mass. Using the above parameters a solar wind pressure of 3×10^{-9} Nm^{-2} is obtained. That is, a pressure of nearly 10^{-4} less than the direct solar radiation pressure. The main effect of the solar wind is then to electrically charge the sail, as discussed in section 1.3.2. First order relativistic effects are proportional to the ratio of the sail speed to the speed of light, typically of order 10^{-4}. Similarly, for solar sails in Earth orbit the secondary pressure due to radiation scattered from the Earth is also several orders of magnitude less than that due to the direct solar radiation pressure, Green (1977).

2.2 Heliocentric Solar Sail Trajectories

The orbital dynamics of solar sail spacecraft are similar in many respects to the orbital dynamics of other spacecraft utilising low thrust propulsion systems. That is, a small continuous thrust is used to modify the spacecraft orbit over an extended period of time. However, a solar-electric propulsion system may orient its thrust vector in any direction, whereas solar sails are constrained to thrust vector orientations within 90^o of the Sun-line. For some mission

applications this leads to significant differences in the final trajectory. For example, to transfer from a prograde to a retrograde orbit a solar-electric system will direct its thrust vector perpendicular to the Sun-line to lose prograde angular momentum and then to gain retrograde angular momentum. However, for solar sails the transfer is made by increasing the spacecraft ecliptic inclination to greater than 90^{o} by alternately orienting the sail above and below the ecliptic plane. This scheme was to be implemented for the comet Halley rendezvous mission, Sauer (1977). These heliocentric 'cranking orbit' manoeuvres have been investigated by Van der Ha and Modi (1979).

The dynamical equation for a heliocentric orbiting solar sail may be obtained by considering the Sun-sail system in an inertial frame I, Figure 2.2. In this frame the centre of mass of the system C is at a position **R**, viz

$$\mathbf{R} = \frac{M\mathbf{r}_1 + m\mathbf{r}_2}{M + m} \tag{2.14}$$

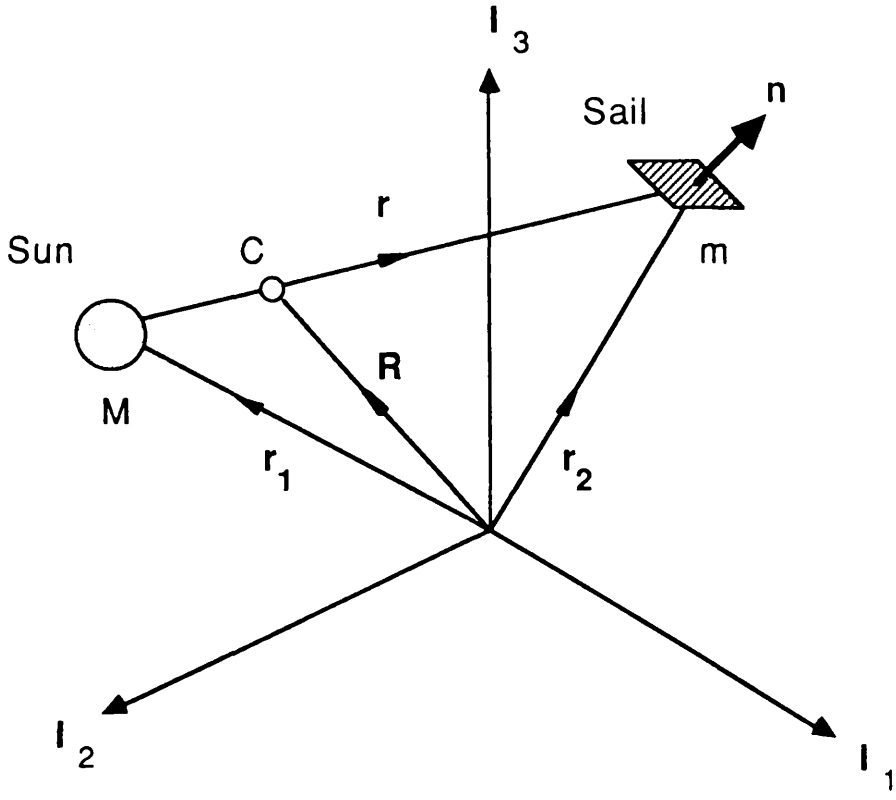
where $(d^2\mathbf{R}/dt^2)=0$ for a closed system, (see appendix B). Using equation (2.9) the forces acting on the Sun and the spacecraft in the inertial frame I are given by

$$M \frac{d^2\mathbf{r}_1}{dt^2} = \frac{GMm}{|\mathbf{r}|^3}\,\mathbf{r} \tag{2.15a}$$

$$m \frac{d^2\mathbf{r}_2}{dt^2} = -\frac{GMm}{|\mathbf{r}|^3}\,\mathbf{r} + \beta \frac{GMm}{|\mathbf{r}|^4}(\mathbf{r}.\mathbf{n})^2\,\mathbf{n} \tag{2.15b}$$

Adding equations (2.15a) and (2.15b) the acceleration of the centre of mass of the system C is obtained as

Figure 2.2



The Sun-sail system in an inertial reference frame I. The Sun (M) is located at $\mathbf{r}_1$ with the sail (m) at $\mathbf{r}_2$. The centre of mass of the system C is located on the Sun-sail vector $\mathbf{r}$ at position $\mathbf{R}$.

$$\frac{d^2\mathbf{R}}{dt^2} = \beta \frac{GMm}{M + m} \frac{(\mathbf{r}.\mathbf{n})^2}{|\mathbf{r}|^4} \mathbf{n} \qquad (2.16)$$

Therefore, since $(d^2\mathbf{R}/dt^2) \neq 0$ the centre of mass of the Sun-sail system accelerates. This is due to the gravitational acceleration of the Sun towards the sail as the sail is accelerated by the solar radiation pressure. However, since $m \ll M$ the term $Mm/(M+m) \simeq m$ so that the acceleration of the centre of mass is, of course, negligable.

If a Galilean transformation is now used to transform to a new inertial frame I', with an origin at the centre of mass C, then

$$M\mathbf{r}_1 + m\mathbf{r}_2 = 0 \quad , \quad \mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 \qquad (2.17)$$

so that the relative acceleration of the Sun and the spacecraft is given by

$$\frac{d^2\mathbf{r}}{dt^2} = \left\{1 + \frac{m}{M}\right\} \frac{d^2\mathbf{r}_2}{dt^2} \qquad (2.18)$$

Substituting from equation (2.15b) the general dynamical equation for a perfectly reflecting solar sail in heliocentric orbit is then given by

$$\frac{d^2\mathbf{r}}{dt^2} + \mu \frac{\mathbf{r}}{|\mathbf{r}|^3} = \beta \frac{\mu}{|\mathbf{r}|^4} (\mathbf{r}.\mathbf{n})^2 \mathbf{n} \quad , \quad \mathbf{r}.\mathbf{n} > 0 \quad , \quad \mu = G(M + m) \qquad (2.19)$$

where $\mu \simeq GM$ since $m \ll M$.

The fundamental aspects of heliocentric solar sail orbital dynamics can be appreciated by calculating the spacecraft orbital angular momentum and energy. Therefore, taking vector and scalar products of equation (2.19) with the sail position and velocity vectors the rate of change of orbital angular momentum $\mathbf{h}$ and energy E are obtained, viz

$$\frac{d\mathbf{h}}{dt} = \beta \frac{\mu}{|\mathbf{r}|^4}(\mathbf{r}.\mathbf{n})^2 \, \mathbf{r} \times \mathbf{n} \quad , \quad \mathbf{h} = \mathbf{r} \times \left\{\frac{d\mathbf{r}}{dt}\right\} \tag{2.20a}$$

$$\frac{dE}{dt} = \beta \frac{\mu}{|\mathbf{r}|^4}(\mathbf{r}.\mathbf{n})^2 \, \mathbf{n}.\left\{\frac{d\mathbf{r}}{dt}\right\} \quad , \quad E = \frac{1}{2}\left|\frac{d\mathbf{r}}{dt}\right|^2 - \frac{\mu}{|\mathbf{r}|} \tag{2.20b}$$

It can be seen from equations (2.20) that to maximise the instantaneous rate of change of angular momentum the sail normal, and so the solar radiation pressure force, must be directed perpendicular to the Sun-line. Similarly, to maximise the rate of change of energy the sail normal must be parallel to the velocity vector. However, if the sail normal is oriented perpendicular to the Sun-line the magnitude of the radiation pressure force is of course zero.

For an initially circular orbit an increase in energy corresponds to an increase in semi-major axis, whereas the eccentricity is a function of the orbital angular momentum. Therefore, depending on the manner in which the spacecraft orbit is to be modified the sail attitude can be chosen such that the azimuthal component of the solar radiation pressure force or the component in the direction of the velocity vector is maximised. One particular trajectory which can be obtained in closed form is the logarithmic spiral.

2.2.1 Logarithmic Spiral Trajectories

The first quantitative investigation of solar sail orbital dynamics approximated the heliocentric dynamical equations to obtain approximate solutions and transfer times, Tsu (1959). However, the exact set of equations were solved by London (1960) using the logarithmic spiral solution of Bacon (1959). This solution requires that the spacecraft thrust has an inverse square variation with heliocentric distance, which is the case for solar sail spacecraft, and that the velocity vector maintains a fixed angle with respect to the

instantaneous radius vector. Writing the general heliocentric dynamical equation, equation (2.19), in polar coordinates it is found that

$$\frac{d^2r}{dt^2} - r\left\{\frac{d\theta}{dt}\right\}^2 = -\frac{\mu}{r^2} + \beta\frac{\mu}{r^2}\cos^3\alpha \qquad (2.21a)$$

$$\frac{1}{r}\frac{d}{dt}\left\{r^2\frac{d\theta}{dt}\right\} = \beta\frac{\mu}{r^2}\cos^2\alpha\sin\alpha \qquad (2.21b)$$

where $\alpha=\cos^{-1}(\mathbf{r}.\mathbf{n}/|\mathbf{r}|)$ is the Sun-sail pitch angle. For a fixed pitch angle α a solution to equations (2.21) is given by

$$r = r_0 e^{\theta\tan\psi} \qquad (2.22)$$

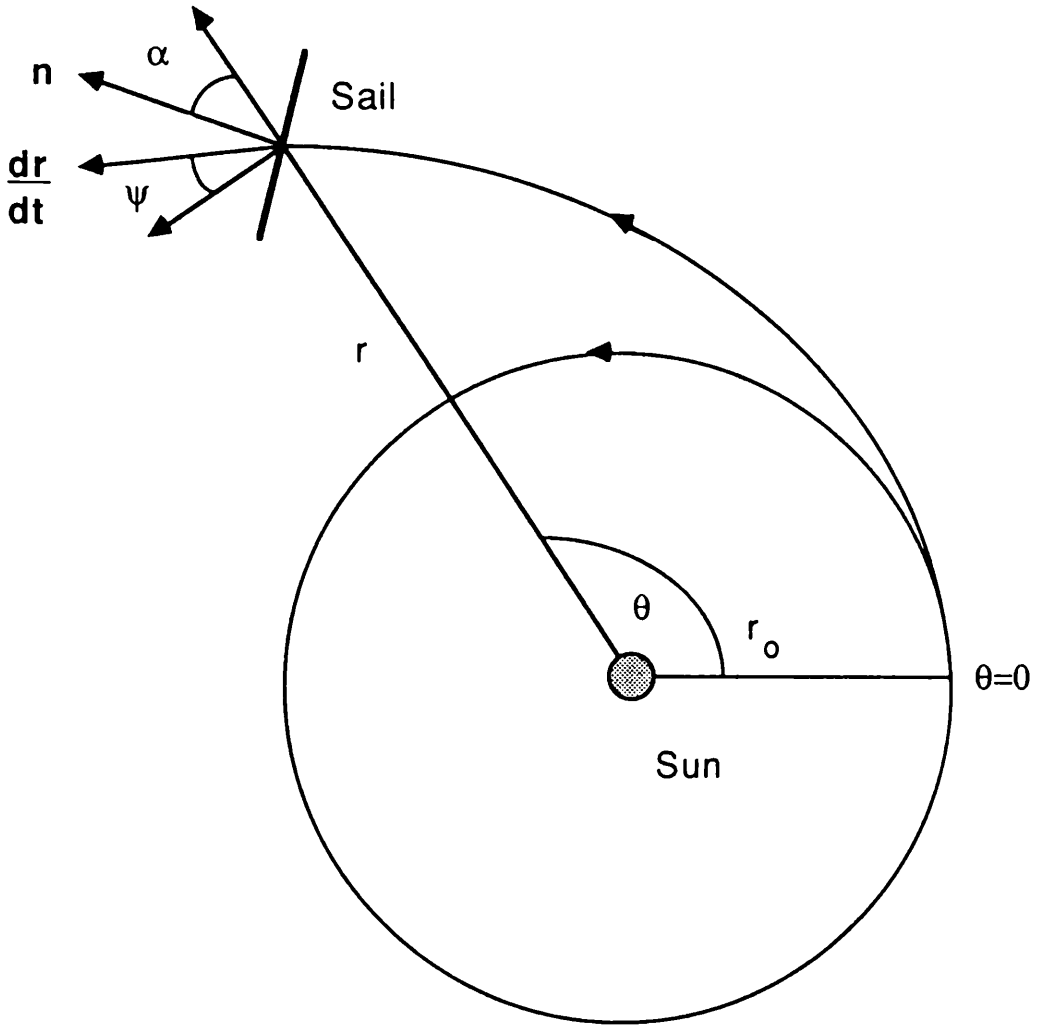
where the spiral angle ψ is the angle between the sail velocity vector and the normal to the Sun-line, Figure 2.3. Substituting equation (2.22) into equations (2.21) it may be shown that there is an implicit relationship between the Sun-sail pitch angle, the sail loading parameter and the spiral angle, viz

$$\frac{\sin\psi\cos\psi}{2-\sin^2\psi} = \frac{\sin\alpha\cos\alpha}{\beta^{-1}-\cos^3\alpha} \qquad (2.23)$$

For a given sail loading parameter and spiral angle equation (2.23) may be solved for the required Sun-sail pitch angle. It should be noted that since the spiral angle is constant an initial impulse is required to rotate the sail velocity vector from an initial circular orbit to the angle ψ required for the logarithmic spiral. In fact, unless the sail loading parameter and the spiral angle are small the impulses required at the initial and final points on the spiral are greater than the total impulse required for a minimum energy ballistic transfer.

By eliminating the azimuthal coordinate θ the transfer time

Figure 2.3



A solar sail logarithmic spiral trajectory with a fixed Sun-sail pitch angle α. The velocity vector is at a fixed angle Ψ with respect to the normal to the Sun-line.

between any two heliocentric distances may be obtained as

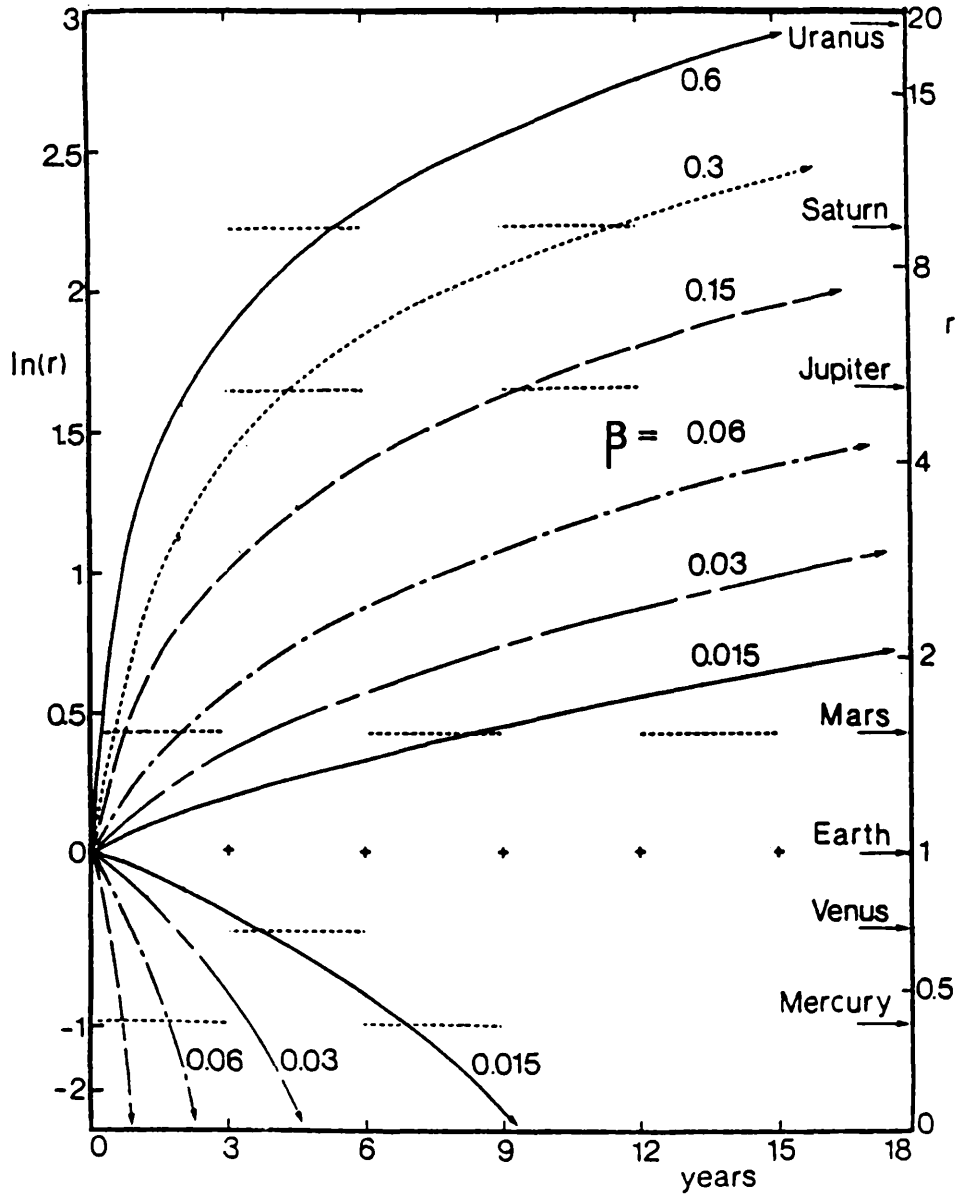
$$t - t_o = \frac{\sqrt{2}}{3} (\beta\mu)^{-1/2} \frac{r^{3/2} - r_o^{3/2}}{r_o} \left\{ \frac{\cot\psi}{\sin\alpha \cos^2\alpha} \right\}^{1/2} \qquad (2.24)$$

The Sun-sail pitch angle required for a minimum time transfer may then be found by setting $(dt/d\alpha)=0$ and solving simultaneously with equation (2.23). The optimised transfer time is shown as a function of heliocentric distance for various loading parameters in Figure 2.4. It is found that for sail characteristic accelerations of less than 1 mms^{-2} the Sun-sail pitch angle required for a time optimal spiral is approximately 35^o. It is interesting to note that this angle maximises the azimuthal component of the solar radiation pressure force.

A second family of closed solutions may be obtained with the sail oriented along the Sun-line. Setting $\alpha=0$ in equations (2.21) the standard two-body dynamical equations are obtained with a reduced gravitational parameter $\mu'=\mu(1-\beta)$. The resulting family of modified conic sections can yield shorter transfer times than the logarithmic spiral in certain cases. A detailed comparison of spiral, modified conic and ballistic trajectories has been made by Kiefer (1965).

The logarithmic spiral solutions have been extended to three-dimensions and explicit asymptotic series solutions for the optimal Sun-sail pitch angle developed, Van der Ha and Modi (1979). Furthermore, the initial conditions of the logarithmic spiral may be relaxed and arbitrary initial conditions allowed. Using the two variable expansion technique, Nayfeh (1973), asymptotic solutions with a fixed sail attitude have been obtained, Van der Ha and Modi (1979), Van der Ha (1980).

Figure 2.4



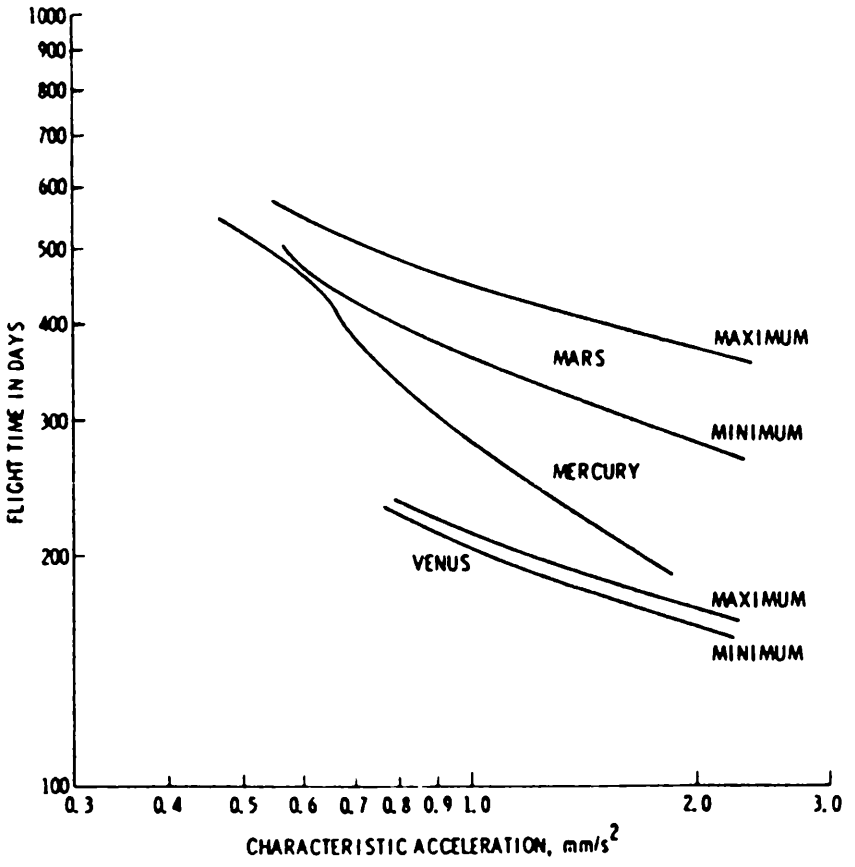
Transfer times for optimal logarithmic spiral trajectories as a function of heliocentric distance and dimensionless sail loading parameter, Van der Ha (1980).

2.2.2 Time Optimal Heliocentric Trajectories

For practical mission analysis purposes the logarithmic spiral solutions and asymptotic series solutions with a fixed sail attitude can be used only as a guide in finding true time optimal solutions. Such solutions require the use of optimal control techniques to obtain the required sail attitude as a function of time. Since the sail attitude is time varying the boundary conditions may be met without the use of initial and final impulses.

Time optimal rendezvous trajectories between circular, coplanar heliocentric orbits have been investigated by Zhukov and Lebedev (1964) who applied the Pontryagin maximum principle of the calculus of variations to obtain the required time varying sail attitude. The two point boundary conditions were satisfied using a numerical iteration scheme. An alternative approach using a numerical gradient method has also been investigated, Kelley (1960). Although both of these approaches provide time optimal trajectories the assumptions of circular, coplanar initial and final orbits are too restrictive for the analysis of practical interplanetary trajectories. These restrictions were removed in detailed studies of three-dimensional time optimal trajectories by Sauer (1976) which took account of the eccentricity and inclination of the initial and final orbits. Using a large data base maximum and minimum transfer times were obtained as a function of the sail characteristic acceleration, Figure 2.5. It is found in general that for outward transfers with a low sail characteristic acceleration the Sun-sail pitch angle is near that of the time optimal logarithmic spiral. For larger accelerations however, a 180^{o} rotation of the sail attitude takes place. That is, the spacecraft is accelerated along the trajectory and then decelerated to rendezvous with the target planet.

Figure 2.5



Transfer times for time optimal heliocentric trajectories. The transfers are between non-coplanar, non-circular orbits with the minimum and maximum transfer times shown. Only the minimum transfer time is shown for Mercury, Sauer (1976).

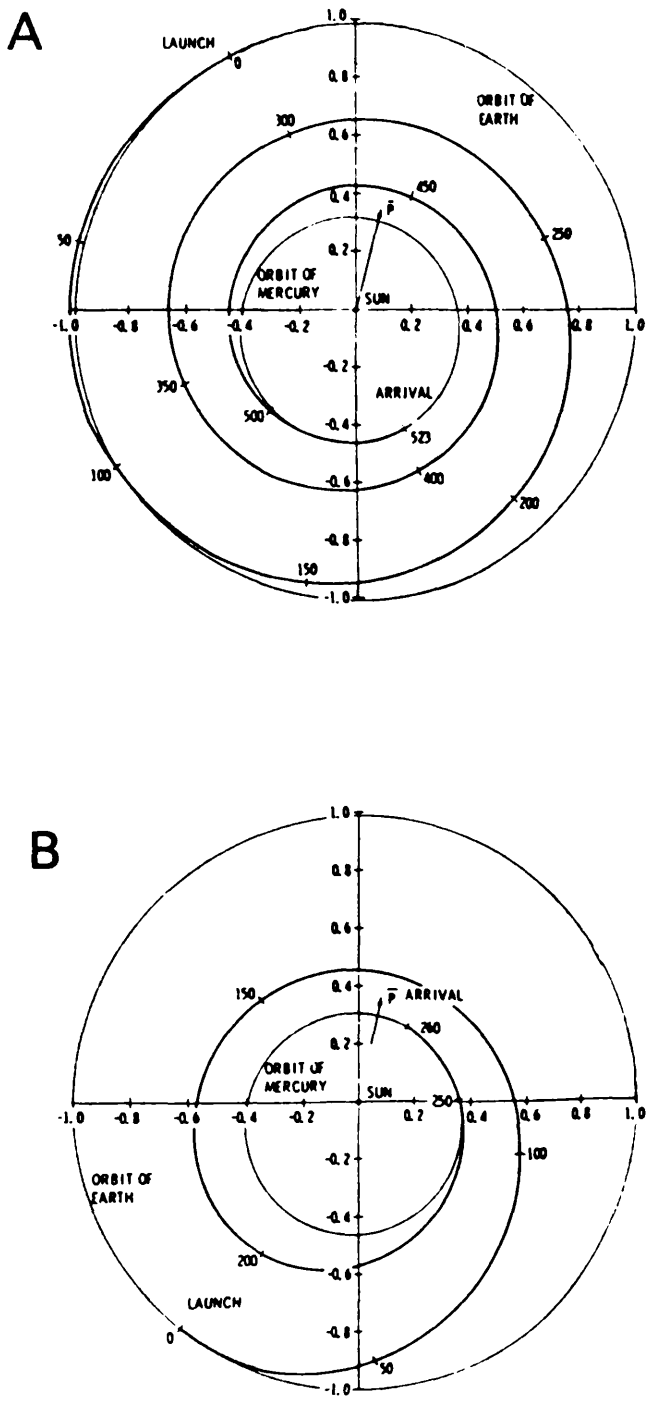
In general the convergence to a true optimal solution is difficult owing to the insensitivity of the transfer time to small variations in the time varying sail attitude.

As part of the investigation by Sauer hyperbolic excess velocities were included in the boundary conditions to investigate their effect on total transfer times. Figure 2.6a shows an ecliptic projection of a time optimal transfer to Mercury with zero relative departure and arrival velocities. A total flight time of 523 days is required with the spacecraft performing 2.5 revolutions of the Sun. With an initial excess velocity of 5 kms^{-1} at Earth escape the transfer angle is reduced to 1.5 revolutions and the transfer time is halved, Figure 2.6b. Therefore, since transfer times can be sensitive to initial conditions the possibility of a hyperbolic excess velocity on escaping from Earth orbit must be considered in the overall mission analysis. The use of a single lunar gravity assist to achieve a hyperbolic excess has recently been considered by Fox et. al (1989).

2.3 Geocentric Solar Sail Trajectories

Now that heliocentric solar sail trajectories have been discussed the question of geocentric escape trajectories must be addressed. Since the solar radiation pressure force cannot be directed sunward, solar sail escape trajectories are distinctly different from escape trajectories for other low thrust propulsion systems. In particular almost no energy can be gained for half of the trajectory while the sail is moving sunward and while the sail is in eclipse. These problems can be partially alleviated by using a polar escape trajectory normal to the Sun-line. However, there is the added cost of initially injecting the spacecraft into a high inclination polar orbit.

Figure 2.6



Time optimal solar sail trajectories from Earth orbit to Mercury with (a) zero hyperbolic excess and a 523 day transfer, (b) 5 kms^{-1} hyperbolic excess and a 260 day transfer, Sauer (1976).

The fundamental aspects of geocentric solar sail orbital dynamics may be understood by again calculating the spacecraft orbital angular momentum and energy. The general dynamical equation for a perfectly reflecting solar sail in geocentric orbit is given by, (cf. equation (2.19))

$$\frac{d^2\mathbf{r}}{dt^2} + \mu \frac{\mathbf{r}}{|\mathbf{r}|^3} = \beta(\mathbf{S}.\mathbf{n})^2\,\mathbf{n} \quad , \quad \mathbf{S}.\mathbf{n} > 0 \tag{2.25}$$

where the parameter β is now the sail characteristic acceleration. It is therefore assumed that the magnitude of the solar radiation pressure force is constant over the scale of the orbit. Furthermore, the Sun-line direction **S** will have a slow annual rotation due to the heliocentric motion of the Earth, but will be assumed to be fixed. It should be noted that if the sail attitude is fixed equation (2.25) has a closed solution in terms of elliptic functions, as will be shown in section 5.9.

Taking vector and scalar products of equation (2.25) the rate of change of orbital angular momentum and energy are obtained, viz

$$\frac{d\mathbf{h}}{dt} = \beta\,(\mathbf{S}.\mathbf{n})^2\,\mathbf{r} \times \mathbf{n} \tag{2.26a}$$

$$\frac{dE}{dt} = \beta\,(\mathbf{S}.\mathbf{n})^2\,\mathbf{n}.\left\{\frac{d\mathbf{r}}{dt}\right\} \tag{2.26b}$$

If the sail orbit is initially near circular, is in the ecliptic plane and the sail normal is fixed along the Sun-line such that **n=S**, then the total change in orbital angular momentum and energy over one orbit may be written as

$$|\Delta \mathbf{h}| = \frac{1}{\Omega}\int_0^{2\pi} \left|\frac{d\mathbf{h}}{dt}\right| d\theta \quad , \quad \left|\frac{d\mathbf{h}}{dt}\right| = \beta |\mathbf{r}| \sin\theta \tag{2.27a}$$

$$\Delta E = \frac{1}{\Omega}\int_0^{2\pi} \left|\frac{dE}{dt}\right| d\theta \quad , \quad \frac{dE}{dt} = \beta \left|\frac{d\mathbf{r}}{dt}\right| \sin\theta \tag{2.27b}$$

where θ is the azimuthal position angle of the sail measured from the Sun-line and Ω is the orbital angular velocity, Figure 2.7. Therefore, it can be seen that if the sail geocentric distance and velocity are slowly varying functions the total change in orbital angular momentum and energy over one orbit is zero. Physically, the energy and momentum gained during the half orbit when the sail is accelerating away from the Sun is lost during the half orbit when the sail is moving towards the Sun.

A simple escape scheme would then be to fix the sail attitude along the Sun-line when the sail is moving away from the Sun, $0<\theta\leqslant\pi$, and to fix the sail attitude normal to the Sun-line when moving towards the Sun, $\pi<\theta\leqslant 2\pi$, Figure 2.7. The total change in orbital angular momentum and energy are then found to be

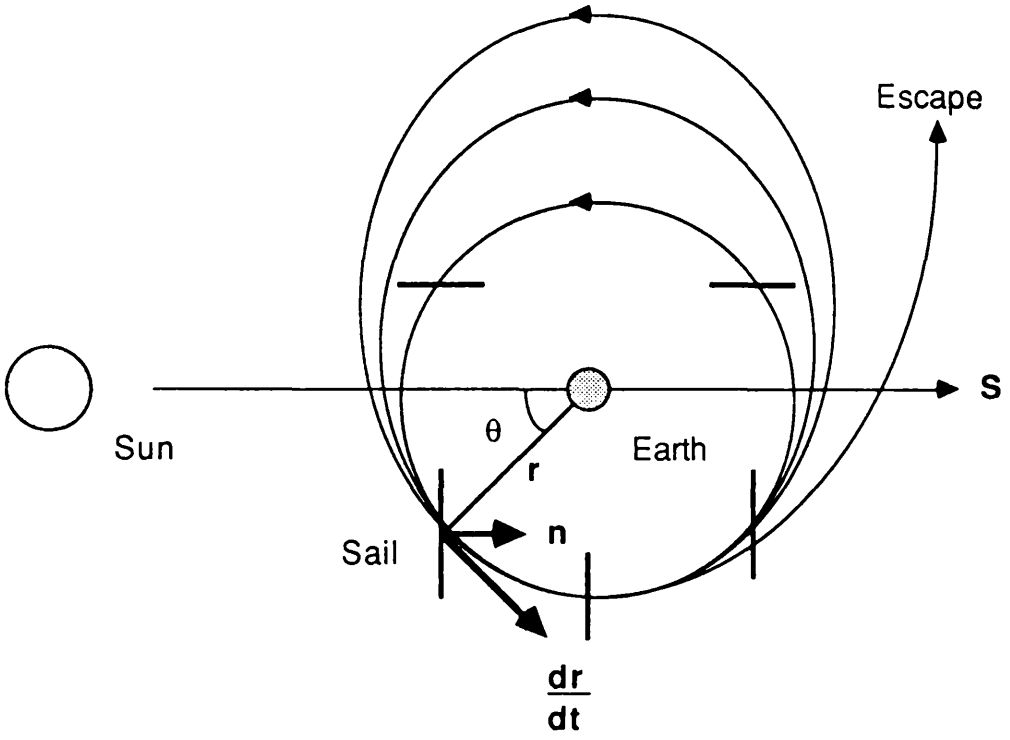
$$|\Delta \mathbf{h}| \approx \frac{2\beta}{\Omega} |\mathbf{r}|_0 \tag{2.28a}$$

$$\Delta E \approx \frac{2\beta}{\Omega} \left|\frac{d\mathbf{r}}{dt}\right|_0 \tag{2.28b}$$

where the spacecraft geocentric distance and velocity are evaluated at the initial circular orbit.

The change in orbital angular momentum and energy may be related to the change in semi-major axis a and eccentricity e of the orbit using the relation $|\mathbf{h}|^2=\mu a(1-e^2)$, (see for example Roy (1982)). For small values of e an expansion gives

Figure 2.7



A solar sail geocentric escape trajectory with the sail attitude **n** fixed along the Sun-line **S** for half of each orbit. The sail reaches escape velocity near the perigee point.

$$\frac{d|\mathbf{h}|}{dt} = \frac{1}{2}\frac{da}{dt}\left\{\frac{\mu}{a}\right\}^{1/2}\left\{1 - \frac{1}{2}e^2\right\} - \frac{de}{dt}e(\mu a)^{1/2} + O(e^4) \qquad (2.29)$$

The semi-major axis may be related to the energy of the system as $E=(-\mu/2a)$. Therefore, eliminating da/dt from equation (2.29) it is found that

$$\frac{de}{dt} = \frac{-1}{e(\mu a)^{1/2}}\left\{\frac{d|\mathbf{h}|}{dt} - a\left\{\frac{a}{\mu}\right\}^{1/2}\frac{dE}{dt}\left\{1 - \frac{1}{2}e^2\right\}\right\} + O(e^4) \qquad (2.30)$$

Substituting for the rates of change of orbital angular momentum and energy the change in eccentricity over one orbit is obtained from equation (2.30) using

$$\Delta e = \frac{1}{\Omega}\int_0^{\pi}\frac{de}{dt}\,d\theta \qquad (2.31)$$

Performing the integration the change in eccentricity over one orbit is then found to be

$$\frac{\Delta e}{e} = -\frac{\beta}{\Omega}\left|\frac{d\mathbf{r}}{dt}\right|_o^{-1} \qquad (2.32a)$$

Similarly, using the relation $E=(-\mu/2a)$ the change in semi-major axis is found to be

$$\frac{\Delta a}{a} = \frac{4\beta}{\Omega}\left|\frac{d\mathbf{r}}{dt}\right|_o^{-1} \qquad (2.32b)$$

It can be seen then that an initially near circular orbit will slowly become elliptical, as shown schematically in Figure 2.7. Furthermore, with a suitable choice of sail attitude control, energy is added to the system so that the semi-major axis of the orbit increases and the spacecraft may be taken to escape. However, to minimise the

escape time more complex schemes are required.

2.3.1 Semi-Optimal Geocentric Trajectories

The first investigations of geocentric escape trajectories were undertaken by Sands (1961) using a simple scheme whereby the sail is rotated at half the orbital angular velocity. Although the sail is taken to escape the scheme is by no means optimal. A more effective orbit raising scheme has been developed by Fimple (1962) by maximising the component of the solar radiation pressure force along the instantaneous velocity vector. It can be seen from equation (2.26b) that this ensures that the instantaneous rate of energy increase is maximised. An orbit initially normal to the Sun-line is used so that there is a continuous gain of energy.

The sail unit normal vector **n** may be written as the sum of the direction cosines of three attitude angles α_j (j=1,3), viz

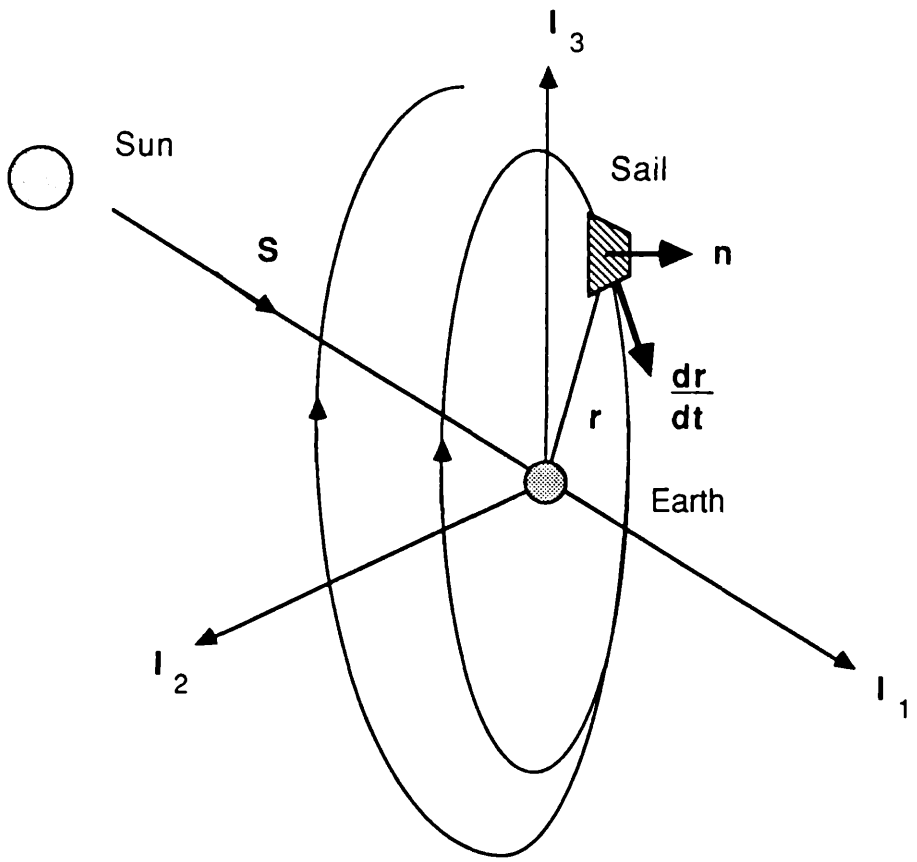
$$\mathbf{n} = \sum_{j=1}^{3} \cos\alpha_j \mathbf{u}_j \quad , \quad \mathbf{r} = \sum_{j=1}^{3} x_j \mathbf{I}_j \tag{2.33}$$

where $\mathbf{I}_j$ (j=1,3) are the unit vectors of an inertial geocentric cartesian coordinate system with $\mathbf{I}_1$ directed along the Sun-line, Figure 2.8. To maximise the instantaneous rate of increase of energy the function $J=(\mathbf{S}.\mathbf{n})^2\mathbf{n}.(d\mathbf{r}/dt)$ must be maximised subject to the normalisation constraint K=0 where

$$J = \cos^2\alpha_1 \sum_{j=1}^{3} \frac{dx_j}{dt}\cos\alpha_j \quad , \quad K = \sum_{j=1}^{3} (\cos^2\alpha_j - 1) \tag{2.34}$$

By using the techniques of constrained maxima with a set of arbitrary Lagrange multipliers λ_j (j=1,3) an optimal set of sail attitude angles

Figure 2.8



Semi-optimal Earth escape trajectory with the solar radiation incident along the Sun-line **S**. The sail attitude **n** is chosen to maximise the component of the solar radiation pressure force along the velocity vector.

may be obtained from

$$\frac{\partial J}{\partial \alpha_j} + \lambda_j \frac{\partial K}{\partial \alpha_j} = 0 \quad , \quad (j=1,3) \tag{2.35}$$

Solving equations (2.35) and eliminating the set of multipliers the required attitude angles are obtained as a function of the sail velocity as

$$\cos\alpha_1 = \Gamma^{-1} \left(3\cos\gamma_1 + (8 + \cos^2\gamma_1)^{1/2}\right) \tag{2.36a}$$

$$\cos\alpha_2 = \Gamma^{-1}\, 2\cos\gamma_2 \tag{2.36b}$$

$$\cos\alpha_3 = \Gamma^{-1}\, 2\cos\gamma_3 \tag{2.36c}$$

where the auxiliary coefficients are given by

$$\Gamma = (6\cos^2\gamma_1 + 6\cos\gamma_1(8 + \cos^2\gamma_1) + 12)^{1/2} \tag{2.37a}$$

$$\cos\gamma_j = \left|\frac{dr}{dt}\right|^{-1} \left\{\frac{dr}{dt}\right\}.\mathbf{I}_j \quad , \quad (j=1,3) \tag{2.37b}$$

With this choice of attitude control the sail is accelerated in the orbital plane while being simultaneously accelerated along the Sun-line. Since the ratio of the local gravitational acceleration to the solar radiation pressure acceleration will initially be large there will be little displacement along the Sun-line. However, as the sail geocentric distance increases the motion along the Sun-line dominates. Integrating equation (2.25) with equations (2.36) it is found that a suitable family of escape trajectories may be obtained, Fimple (1962).

General, long term geocentric solar sail trajectories have been investigated by Van der Ha and Modi (1977a, 1977b) using the two variable expansion technique. Orbit raising schemes have been constructed by switching the sail attitude at various points along the

orbit between a null configuration normal to the Sun-line and a Sun facing attitude along the Sun-line, Van der Ha and Modi (1977c). By calculating analytically the change in orbital elements over a particular arc of the orbit with the sail attitude fixed, various schemes may be quickly evaluated. The analytic solutions for the changes in the spacecraft orbital elements are used to up-date the true orbit in a numerical rectification procedure. In particular it is found that switching the sail attitude at the points where the sail velocity vector is normal to the Sun-line leads to a rapid increase in energy.

2.3.2 Time Optimal Geocentric Trajectories

Time optimal escape trajectories have been investigated by Sackett and Edelbaum (1977) using the Pontryagin maximum principle of the calculus of variations. By averaging the dynamical equations over one orbital period short period terms may be eliminated, leading to a reduction in numerical computation time. Since the method of averaging is only valid for a small ratio of solar radiation pressure acceleration to local gravitational acceleration the solutions are valid to geocentric distances of 10^5 km only. For an initial high Earth orbit with a geocentric distance of 2.1×10^4 km a sub-escape point of 10^5 km was attained in 116 days with a spacecraft characteristic acceleration of 0.6 mms^{-2}. This transfer time falls to 70 days with a characteristic acceleration of 1 mms^{-2}. For many trajectories a rapid increase in eccentricity is found with a lowering of the sail perigee distance. To avoid perigee distances of less than the radius of the Earth and to avoid the rapid attitude manoeuvres associated with high eccentricity orbits a penalty function was added to bias against a low perigee.

The analysis of Sackett and Edelbaum was extended by Green

(1977) to large geocentric distances to investigate the final escape manoeuvre. From a high Earth orbit with a geocentric distance of 1.5×10^5 km orbit a solar sail with a characteristic acceleration of 1 mms^{-2} reached escape velocity in 20 days. It was found that a velocity dependent attitude control similar to section 2.3.1 produced solutions close to optimum with an increase in escape time of only 1.2×10^{-2}. Furthermore, it was found that polar escape trajectories required slower attitude manoeuvres but were rather longer than low inclination escape trajectories. More recently the escape problem has been investigated by Borja (1984) who considered the combined rigid body dynamics and orbital dynamics of a freely precessing, spinning disc sail. By maximising the energy gain per orbit near optimal escape trajectories were generated.

2.4 Conclusions

It has been shown that solar sail spacecraft can be manoeuvred in geocentric and heliocentric orbit by several methods. For the heliocentric case the logarithmic spiral solution provides a simple analytic solution to the problem. However, the large initial and final impulses required renders the trajectory impractical. For practical mission analysis purposes time optimal trajectories which satisfy the boundary conditions for transfer between non-coplanar, non-circular orbits are required.

For geocentric escape time optimal trajectories may again be generated. While providing the best solution in terms of the orbital dynamics of the problem time optimal escape may require large attitude turning rates which can be impractical for large sails. Excessive turning rates can however be avoided by choosing the sail normal to

be fixed perpendicular to the sail radius vector. The sail may then be switched into a null attitude when moving sunward by rotation about the radial axis, while the structure remains gravity gradient stabilised.

Finally, nearly all of the studies of solar sail trajectories to date have investigated the use of solar radiation pressure as a means of modifying the solar sail trajectory for interplanetary transfer. However, other advanced applications, such as the static equilibrium 'statite' discussed in section 1.5.3, require the solar radiation pressure force only to modify the local gravitational acceleration. This means of utilising the solar radiation pressure force for advanced applications will be investigated in the remainder of this thesis.

3. AN EXACT SOLAR RADIATION PRESSURE MODEL

3.1 Introduction

In previous solar sail trajectory studies, as discussed in chapter 2, it has been assumed that the solar radiation pressure force has an inverse square variation with heliocentric distance. This allows in some cases a closed analytic solution to the dynamical equations, such as the logarithmic spiral trajectory. In this chapter it will be shown that for a planar solar sail the assumption of an inverse square variation is in fact not valid when account is taken of the finite angular size and limb darkening of the solar disk. This new astrophysical modelling of the source of solar radiation pressure is distinct from the modelling of the solar radiation pressure force, which is dependent on the optical properties of the sail material.

Starting from the fundamental definition of radiation pressure through the radiation pressure tensor, to take account of the varying direction of incidence of solar radiation from different parts of the solar disk, it will be shown that the solar radiation pressure is modified from an inverse square form by a function of the sail heliocentric distance and the solar radius. A more precise calculation is also carried out using a limb darkened solar disk which gives a closed, but complex functional form for the limb darkened solar radiation pressure.

It should of course be noted that at large heliocentric distances these effects are small and can become comparable to other secondary forces, as discussed in section 2.1. However, as will be shown using the simpler non-limb darkened solar radiation pressure, this modification results in the further de-stabilisation of the previously

assumed marginal instability of solar sails stationary above the solar poles, as discussed in section 1.5.3. Furthermore, the sail loading parameter required for these stationary solutions (β=1) will no longer be independent of the sail heliocentric distance, as is the case for an inverse square variation of the solar radiation pressure. The stability analysis is also extended to solar sail circular orbital motion. It is found that for a given sail orbital period there is an inner heliocentric distance where the sail motion becomes unstable.

Finally, the effects of small, time variations in the solar radiation pressure are investigated to attempt to model the effect of short period fluctuations in the solar luminosity. It is found that for one-dimensional sail motion a reduced form of Mathieu's equation is obtained and that Floquet stability analysis shows that the motion always remains bound. For the circular orbital case Mathieu's equation is obtained in full. It is again found that for short period fluctuations the sail motion is bound with non-periodic, bound solutions to Mathieu's equation. However, for long period variations (ie. the eleven year solar cycle) the circular orbital motion can become unstable for certain ranges of parameters.

3.2 Solar Radiation Pressure with an Extended Source

The functional form of the solar radiation pressure exerted on a planar, perfectly reflecting solar sail will now be obtained using the radiation pressure tensor. The frequency integrated radiation pressure tensor $\mathbf{P}(\mathbf{r};t)$ is defined as the second angular moment of the specific intensity of the radiation field integrated over the entire frequency spectrum, as used in section 2.1

$$\mathbf{P}(\mathbf{r};t) = \frac{1}{c}\int_0^{\infty}\oint_{4\pi} I_\nu(\mathbf{r},\mathbf{n};t)\ \mathbf{nn}\ d\Omega d\nu \qquad (3.1)$$

However, in section 2.1 the solar radiation incident on the sail surface was assumed to be incident along parallel rays. The varying direction of incidence of radiation from different parts of the solar disk will now be included in the integration of the radiation pressure tensor. Therefore, equation (3.1) will be used to obtain the exact radiation pressure from a uniformly bright and limb darkened, finite sized solar disk.

3.2.1 Uniformly Bright Solar Disk

In the case of the non-limb darkened solar radiation field where the specific intensity is time independent and isotropic across the solar disk, the radiation pressure on a radially oriented, perfectly reflecting sail at a heliocentric distance r may be written as

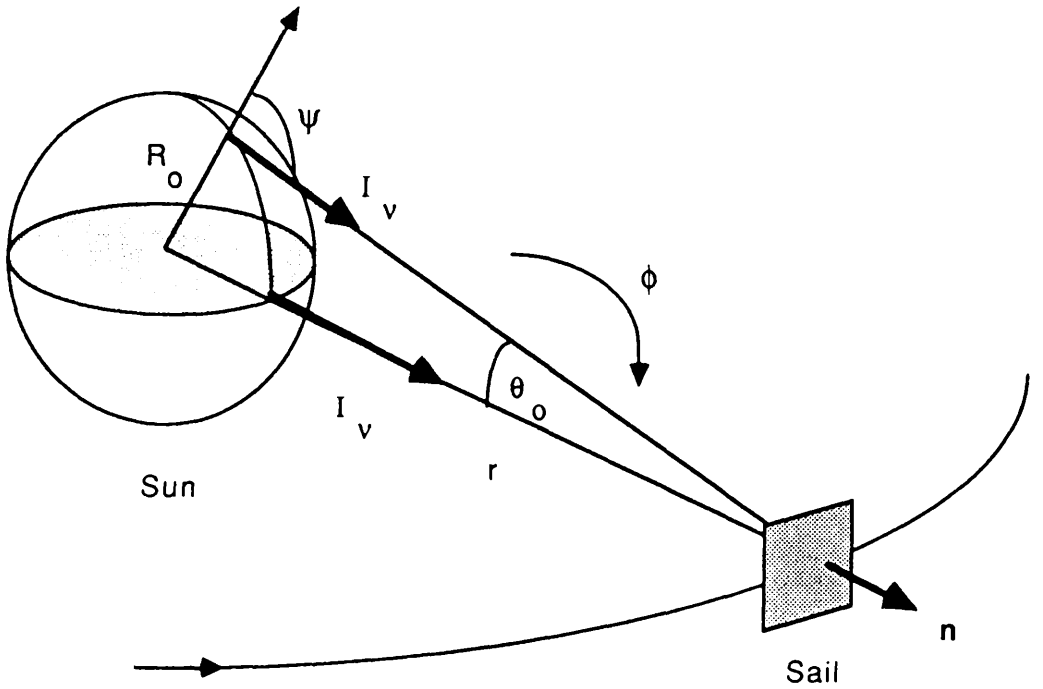
$$P(r) = \frac{2}{c}\int_0^{\infty}\int_0^{2\pi}\int_0^{\theta_o} I_\nu \cos^2\theta\ d\Omega d\nu \quad , \quad d\Omega = \sin\theta\ d\theta d\phi \qquad (3.2)$$

where the geometry of the system is specified in Figure 3.1. Making use of the conservation of specific intensity along rays, so that I_ν is independent of r, and noting the azimuthal symmetry of the geometry, equation (3.2) reduces to the integral

$$P(r) = \frac{4\pi}{c} I_o \int_{\eta_o}^{1} \eta^2\, d\eta \quad , \quad \eta = \cos\theta \quad , \quad \eta_o = \cos\theta_o \qquad (3.3)$$

where I_o is the frequency integrated specific intensity. Performing this integration and substituting for η_o it is found that

Figure 3.1



Sun-sail geometry with a finite angular sized solar disk. The solar disk has an apparent angular diameter of e_o from a heliocentric distance r and the azimuthal angle Φ is defined about the radial direction. The viewing angle Ψ is the angle between the normal to the solar surface and the line of sight.

$$P(r) = \frac{4\pi}{3c} I_o \left\{1 - \left\{1 - \left\{ \frac{R_o}{r} \right\}^2 \right\}^{3/2}\right\} \tag{3.4}$$

Equation (3.4) may be expanded in $(R_o/r)^2$ and for $r \gg R_o$ may be written to first order as

$$P(r) = \frac{2\pi}{c} I_o \left\{\frac{R_o}{r}\right\}^2 + O\left((R_o/r)^4\right) \tag{3.5}$$

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

$$P^*(r) = \frac{2}{c} \left\{\frac{L_o}{4\pi r^2}\right\} \tag{3.6}$$

where L_o is the solar luminosity. Hence, by comparing equations (3.5) and (3.6) the frequency integrated specific intensity I_o is identified as

$$I_o = \frac{L_o}{4\pi^2 R_o^2} \tag{3.7}$$

as derived in section 2.1. Substituting for I_o in equation (3.4) an expression for the solar radiation pressure exerted on a radially oriented, perfectly reflecting solar sail from a uniformly bright, finite sized solar disk is obtained as

$$P(r) = \frac{L_o}{3\pi c R_o^2} \left\{1 - \left\{1 - \left\{ \frac{R_o}{r} \right\}^2 \right\}^{3/2} \right\} \tag{3.8}$$

A more useful way of representing equation (3.8) is to express it in terms of the point source, inverse square variation $P^*(r)$ defined by equation (3.6), viz

$$P(r) = P^*(r)F(r) \ , \ F(r) = \frac{2}{3}\left\{\frac{r}{R_o}\right\}^2 \left\{1 - \left\{1-\left\{ \frac{R_o}{r} \right\}^2\right\}^{3/2} \right\} \quad (3.9)$$

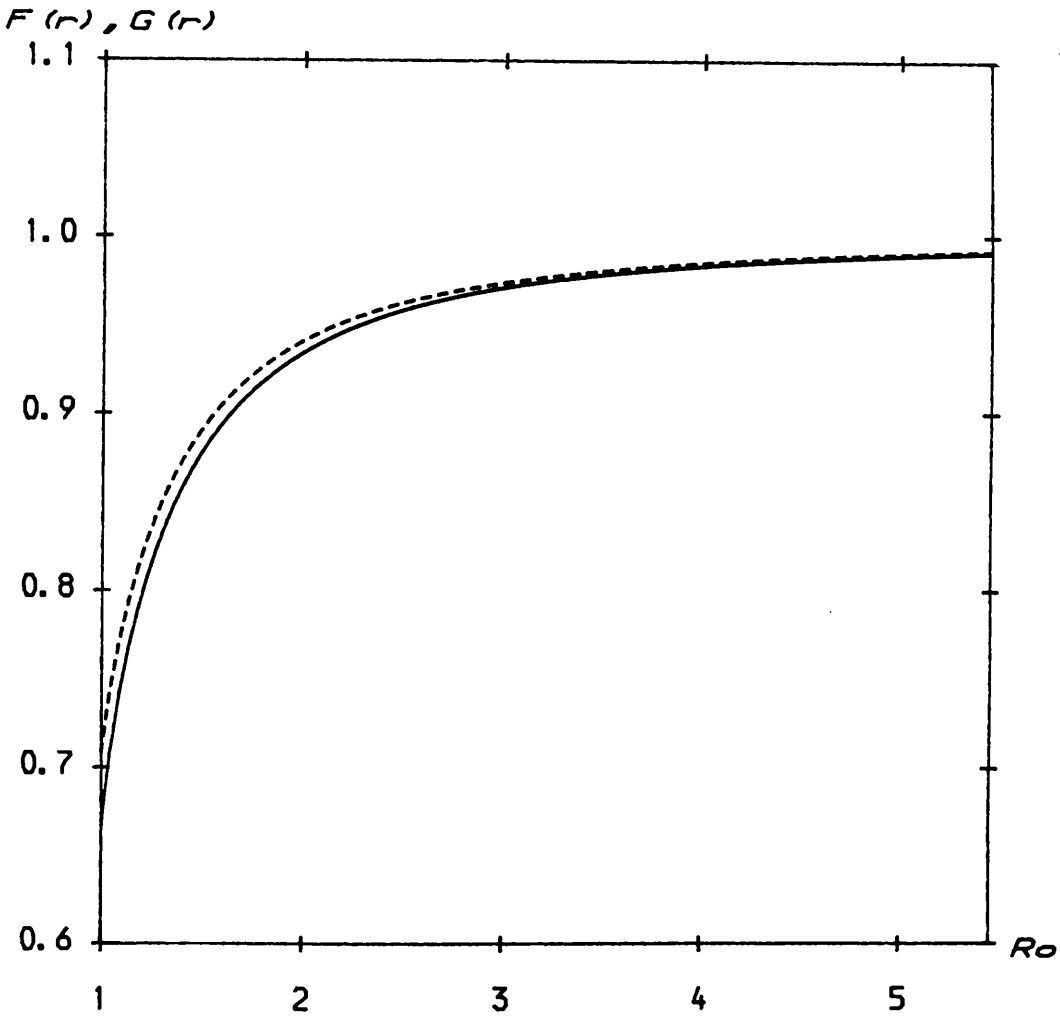
It is seen that the function F(r) attains its minimum value at $r=R_o$, where $F(R_o)=2/3$, giving the greatest deviation of the solar radiation pressure from an inverse square variation. Furthermore, as $r\rightarrow\infty$, $F(r)\rightarrow 1$ since the solar disk becomes more point-like, as shown in Figure 3.2. It can be seen that F(r) approaches unity over a scale of a few solar radii so that the magnitude of the deviation from an inverse square form is extremely small at large heliocentric distances.

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

3.2.2 Limb Darkened Solar Disk

A more accurate model of the solar radiation pressure may be obtained by the inclusion of solar limb darkening in the functional form of the specific intensity. Limb darkening is an effect due to the specific intensity of the solar radiation field having a directional dependence. That is, as the radiation from a point on the solar surface is viewed from an oblique angle the associated specific intensity falls so that the limb of the solar disk appears darker than the disk centre. Empirically, solar limb darkening has a complex functional form, Allen (1955). However, using an approximate model of

Figure 3.2



Deviation of the solar radiation pressure from an inverse square form. F(r) (——) gives the deviation for a uniformly bright disk and G(r) (----) for a limb darkened disk.

the solar atmosphere an analytic expression for the limb darkening can be obtained. The Eddington grey solar atmosphere model, Mihalas and Mihalas (1984), which assumes that the solar atmosphere is convection free and is in both radiative and local thermodynamic equilibrium, allows such an analytic solution to the radiative transfer equations. The specific intensity of the solar radiation field may then be written as

$$I = \frac{I_o}{4} (2 + 3\cos\Psi) \tag{3.10}$$

where I_o is the frequency integrated specific intensity defined in equation (3.7) and the viewing angle Ψ is shown in Figure 3.1. It is seen from equation (3.10) that the solar limb will appear darker than the centre of the solar disk by a factor of 0.4 using this grey atmosphere approximation. It would appear then that solar limb darkening will have an important effect when the radiation pressure tensor is integrated.

The angle Ψ may be related to the integration variable θ through the equation

$$\cos\Psi = \left\{1 - \left\{\frac{r}{R_o}\right\}^2 \sin^2\theta\right\}^{1/2} \tag{3.11}$$

so that the required integral now becomes

$$P(r) = \frac{2}{c} \int_o^{2\pi} \int_o^{\theta_o} \frac{1}{4} I_o \left\{2 + 3\left\{1-\left\{\frac{r}{R_o}\right\}^2 \sin^2\theta\right\}^{1/2}\right\} \cos^2\theta \, d\Omega \tag{3.12}$$

After some lengthy integration an expression for the solar radiation pressure from a finite angular sized, limb darkened solar disk is

obtained as

$$P(r) = \frac{2\pi}{3} I_o \left\{1 - \left\{1-\left\{\frac{R_o}{r}\right\}^2\right\}^{3/2}\right\} + \frac{3\pi}{2} I_o \left\{\frac{R_o}{r}\right\}^3 \left\{\frac{1}{4}\left\{\frac{r}{R_o}\right\}\left\{1 + \left\{\frac{r}{R_o}\right\}^2\right\}\right.$$

$$\left. + \frac{1}{8}\left\{1 - \left\{\frac{r}{R_o}\right\}^2\right\}^2 \ln\left\{\frac{(r/R_o) - 1}{(r/R_o) + 1}\right\}\right\} \qquad (3.13)$$

This expression can again be written in terms of the point source, inverse square radiation pressure $P^*(r)$, viz

$$P(r) = P^*(r)G(r) \qquad (3.14)$$

where the function G(r) is defined by

$$G(r) = \frac{1}{3}\left\{\frac{r}{R_o}\right\}^2\left\{1 - \left\{1 - \left\{\frac{R_o}{r}\right\}^2\right\}^{3/2}\right\} + \frac{3}{4}\left\{\frac{R_o}{r}\right\}\left\{\frac{1}{4}\left\{\frac{r}{R_o}\right\}\left\{1 + \left\{\frac{r}{R_o}\right\}^2\right\}\right.$$

$$\left. + \frac{1}{8}\left\{1 - \left\{\frac{r}{R_o}\right\}^2\right\}^2 \ln\left\{\frac{(r/R_o) - 1}{(r/R_o) + 1}\right\}\right\} \qquad (3.15)$$

The functional form of equation (3.15) is shown in Figure 3.2. At the solar surface $G(R_o)=0.708$ so that the limb darkened solar radiation pressure deviates less from an inverse square form than the non-limb darkened pressure. This is due to the reduced momentum transfer from the solar limb. However, at large heliocentric distances $G(r)\rightarrow 1$ as expected. It can be seen then that the functions F(r) and G(r) have the same overall functional behaviour but differ somewhat in precise numerical values, their fractional difference being of order 10^{-2} or less for all but the closest heliocentric distances. For this reason the uniformly bright solar disk approximation with its much simpler functional form will be used in the following analysis.

3.3 Consequences for Stability: 1-Dimensional Motion

Using the expression for the solar radiation pressure derived in section 3.2.1 the stability of a solar sail stationary above the solar poles will be investigated. By linearising the dynamical equations about a stationary point the form of the resulting sail motion will be examined to determine the dynamical stability of the system.

A radially oriented solar sail with $\mathbf{n}=\mathbf{r}/|\mathbf{r}|$ and with a total mass per unit area σ will now be considered under the influence of solar gravity and an inverse square solar radiation pressure given by $P^*(r)$. Equation (2.19) then gives the dynamical equation as

$$\frac{d^2\mathbf{r}}{dt^2} = -\mu(1-\beta)\frac{\mathbf{r}}{|\mathbf{r}|^3} \quad , \quad \beta = \frac{\sigma^*}{\sigma} \tag{3.16}$$

From equation (3.16) it can be seen that there is a unique value of β which gives a stationary solution independent of r, (ie. $\beta=1$). This independence is of course due to the solar gravitational and radiation pressure accelerations both having an inverse square variation. The critical sail mass per unit area for a stationary solution σ^* is then given by equation (2.10) as

$$\sigma^* = \frac{L_0}{2\pi GMc} \quad , \quad (\sigma^* = 1.53\ \mathrm{gm}^{-2}) \tag{3.17}$$

If the sail is initially at rest at some heliocentric position $\mathbf{r}_0$, with $\beta=1$ and a perturbation $\boldsymbol{\varepsilon}$ is applied, such that $\mathbf{r}_0 \rightarrow \mathbf{r}_0+\boldsymbol{\varepsilon}$, the resulting sail motion is obtained from equation (3.16) as

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} = 0 \Rightarrow \boldsymbol{\varepsilon}(t) = \boldsymbol{\varepsilon}_{01} + \boldsymbol{\varepsilon}_{02}\, t \tag{3.18}$$

where $\boldsymbol{\varepsilon}_{01}$, $\boldsymbol{\varepsilon}_{02}$ are constants of the motion. Equation (3.18) is in fact

valid for $|\boldsymbol{\varepsilon}|\gg 1$ since the system is force free with $\beta=1$. It can be seen then that for an inverse square variation of solar radiation pressure a stationary sail has marginal type instability, (see appendix C). That is, if perturbed the sail will move from its stationary point, but with a linear rather than exponential growth.

Consider now the one-dimensional dynamical equation with the modified solar radiation pressure of a uniformly bright solar disk given by equation (3.9). The dynamical equation now becomes

$$\frac{d^2\mathbf{r}}{dt^2} = -\mu(1-\beta F(r))\frac{\mathbf{r}}{|\mathbf{r}|^3} \tag{3.19}$$

where the sail loading parameter β is the ratio of the inverse square radiation pressure force to the solar gravitational force exerted on the sail. Therefore, the parameter β is not, in this context, the ratio of the actual forces acting on the sail but is still defined as the ratio σ^*/σ. Clearly then, there is no longer a unique value of β giving a stationary solution at all heliocentric distances. The required value of β will now be a function of the sail heliocentric distance given by

$$\beta_c(r) = F(r)^{-1} \tag{3.20}$$

The sail will now be considered to be stationary at a heliocentric position $\mathbf{r}_o$ with $\beta=\beta_c(r_o)$ and a perturbation of $\boldsymbol{\varepsilon}$ applied, such that $\mathbf{r}_o\rightarrow\mathbf{r}_o+\boldsymbol{\varepsilon}$. Then, expanding equation (3.19) in powers of $\boldsymbol{\varepsilon}$, the equation of the subsequent sail motion is obtained. Considering a radial perturbation only it is found that

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} = -\frac{\mu}{r_o^3}\left\{1-\beta_c\left\{F(r_o)+\left.\frac{\partial F}{\partial r}\right|_{r=r_o}\boldsymbol{\varepsilon}\right\}\right\}\left\{1-3\frac{\boldsymbol{\varepsilon}}{r_o}\right\}(\mathbf{r}_o+\boldsymbol{\varepsilon}) \tag{3.21}$$

Substituting for β_c a second order variational equation defining the sail motion in the neighbourhood of the stationary position r_o is obtained. Retaining linear terms only it is found that

$$\frac{d^2\varepsilon}{dt^2} - D(r_o)\varepsilon = 0 \quad , \quad D(r_o) = \frac{\mu}{r_o^3}\left\{\frac{r}{F(r)}\frac{dF}{dr}\bigg|_{r=r_o}\right\} \tag{3.22}$$

The stability characteristics of the system may now be investigated by calculating the eigenvalues of the variational equation. This may be carried out by substituting an exponential solution of the form

$$\varepsilon = \varepsilon_o e^{\omega t} \tag{3.23}$$

Therefore, substituting this solution in equation (3.22) it is found that

$$(\omega^2 - D(r_o))\,\varepsilon_o = 0 \tag{3.24}$$

so that the system eigenvalues $\omega_{1,2}$ are given by $\pm D(r_o)^{1/2}$. Evaluating the derivative of F(r) and making the substitution $\nu=(R_o/r_o)^2$ the function $D(r_o)$ reduces to

$$D(r_o) = \frac{2\mu}{r_o^3} M(\nu) \quad , \quad M(\nu) = 1 - \frac{3}{2}\nu\frac{(1-\nu)^{1/2}}{1-(1-\nu)^{3/2}} \tag{3.25}$$

where the sign of the function $M(\nu)$ determines whether the system eigenvalues are real or purely imaginary.

The asymptotic behaviour of the function $M(\nu)$ in the limits of $\nu\to1$, $(r_o\to R_o)$ and $\nu\to0$, $(r_o\to\infty)$ will now be examined. Firstly, as $\nu\to1$ it is seen that $M(\nu)\to1$ and as $\nu\to0$ the function $M(\nu)$ may be expanded as

$$M(\nu) = 1 - \frac{1 - \frac{1}{2}\nu + O(\nu^2)}{1 - \frac{1}{4}\nu + O(\nu^2)} \to 0 \quad , \quad \nu\to0 \tag{3.26}$$

Since it may also be shown that the function $M(\nu)$ has only one real root at $\nu=0$, and so does not change sign for $0<\nu<1$, it is concluded that $M(\nu)>0$, $0<\nu<1$. Therefore, there are two real eigenvalues of opposite sign and an exponential solution of the form

$$\varepsilon(t) = \sum_{j=1}^{2} \varepsilon_{0j}\, e^{\omega j t} \tag{3.27}$$

where ε_{01}, ε_{02} are constants and the system eigenvalues $\omega_{1,2}$ are given by

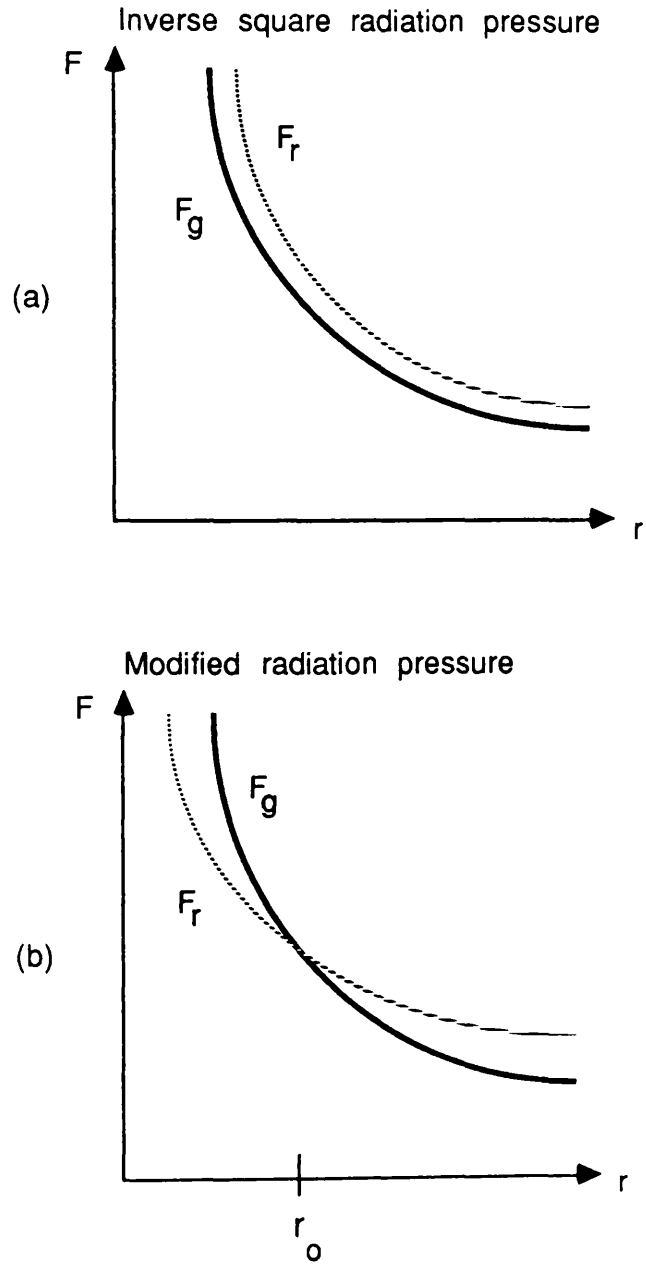
$$\omega_{1,2} = \pm\sqrt{2}\ \omega_0\ M(r/R_0)^{1/2} \quad , \quad \omega_0^2 = \frac{\mu}{r_0^3} \tag{3.28}$$

where ω_0 is the angular velocity of a Keplerian circular orbit at a heliocentric distance r_0.

It has been shown then that by considering the Sun as a uniformly bright, extended source of radiation that a solar sail of a given loading has only one possible stationary point and that this point is exponentially unstable, the instability being independent of the sail parameters. The timescale of the instability $\tau=2\pi/\omega_1$ is however dependent on the sail heliocentric distance. The instability timescale is itself large (eg. $\tau=0.96$ years for $r_0=0.1$ AU) but the existance of the instability adds to the need for active sail station-keeping so that the dynamics of the one-dimensional Sun-sail system are not as elementary as is at first thought.

The instability may be understood physically as shown schematically in Figures 3.3a and 3.3b. Figure 3.3a shows the variation of the solar gravitational force F_g and the inverse square solar radiation pressure force F_r. It can be seen that the solar

Figure 3.3



Schematic form of (a) the solar gravitational and inverse square radiation pressure forces and (b) the modified radiation pressure force with a stationary solution at $r=r_o$.

gravitational force may be balanced at all heliocentric distances by the solar radiation pressure force with a suitable choice of sail loading, therefore giving a stationary solution with marginal instability as discussed above. However, in Figure 3.3b with the deviation of the solar radiation pressure force from an inverse square form, it can be seen that for a given sail loading there is only one heliocentric distance r_o where F_g and F_r intersect and so there is only one stationary solution. For $r>r_o$, $F_r>F_g$ so that the sail is accelerated outward by the solar radiation pressure force and for $r<r_o$, $F_r<F_g$ so that the spacecraft falls sunwards accelerated by the solar gravitational force. Therefore the stationary solution at r_o is unstable.

3.4 Consequences for Stability: 2-Dimensional Motion

It has been shown that the modification to the inverse square form of the solar radiation pressure gives unstable stationary solutions for the one-dimensional Sun-sail system. The stability of a solar sail in a circular heliocentric orbit, such as a 25 day orbit following the solar equatorial rotation or a one year earth synchronous orbit, will now be investigated.

A solar sail in heliocentric orbit with a sail loading parameter β will now be considered. For a radial sail orientation with $\mathbf{n}=\mathbf{r}/|\mathbf{r}|$ the dynamical equations are given by equations (2.21) as

$$\frac{d^2r}{dt^2} - r\left\{\frac{d\theta}{dt}\right\}^2 = -\frac{\mu}{r^2} + \beta\frac{\mu}{r^2}F(r) \qquad (3.29a)$$

$$\frac{1}{r}\frac{d}{dt}\left\{r^2\frac{d\theta}{dt}\right\} = 0 \qquad (3.29b)$$

where β is again defined as the ratio σ^*/σ. From equation (3.29b) it is

seen that an angular momentum integral may be immediately obtained, viz

$$r^2 \left\{\frac{d\theta}{dt}\right\} = h \tag{3.30}$$

where h is the spacecraft angular momentum per unit mass. Using equation (3.30) to eliminate $(d\theta/dt)$ from equation (3.29a) a single radial equation is obtained as

$$\frac{d^2r}{dt^2} - \frac{h^2}{r^3} = -\frac{\mu}{r^2}\left(1 - \beta F(r)\right) \tag{3.31}$$

Firstly, the case of a purely inverse square variation of the solar radiation pressure will be considered with $F(r)=1$. A reduced gravitational constant $\mu^*=\mu(1-\beta)$ may therefore be defined. If the sail is on an initially circular orbit with $(d^2r/dt^2)=0$ at $r=r_o$ then $h^2=\mu^* r_o$. If a perturbation δ is applied, such that $r_o \rightarrow r_o+\delta$ (assuming without loss of generality that h remains constant) and equation (3.31) is linearised with respect to δ a variational equation is obtained defining the sail motion in the neighbourhood of the initial circular orbit, viz

$$\frac{d^2\delta}{dt^2} + \frac{\mu^*}{r_o^3}\,\delta = 0 \tag{3.32}$$

which has a solution of the form

$$\delta(t) = \sum_{j=1}^{2} \delta_{oj}\, e^{i\omega^* j t} \quad , \quad \omega^*_{1,2} = \pm\left\{\frac{\mu^*}{r_o^3}\right\}^{1/2} \tag{3.33}$$

where δ_{o1}, δ_{o2} are constants and ω^* is the angular velocity of a circular orbit at a heliocentric distance r_o with a reduced gravitational constant μ^*. Therefore, as expected, solar sail orbits with an inverse

square variation of solar radiation pressure are linearly stable.

Since the eigenvalues of the system are purely imaginary it would appear that the sail motion is Lyapunov stable, (see appendix C). However, the azimuthal drift in the sail motion due to the assumption that h remains constant has been ignored. This drift is due to the sail at perturbed position $r_o+\varsigma$ having a first order difference in angular velocity to the unperturbed orbit at r_o. If the applied perturbation is of the form $r_o \to r_o+\varsigma$ and $\theta \to \omega^* t+\Psi$ then equation (3.29b) gives the first order drift as

$$\frac{d\Psi}{dt} = -\frac{2\omega^*}{r_o}(\varsigma(t) - (\varsigma_{o1}+\varsigma_{o2})) \tag{3.34}$$

Therefore the sail motion is Lyapunov unstable (the sail diverges from its unperturbed position), although the heliocentric distance remains bound so that the motion is Poincaré stable (see appendix C).

The analysis is now repeated with the modified, uniformly bright form of the solar radiation pressure as defined by equation (3.9). The sail angular momentum per unit mass is now given by $h^2=\mu r_o(1-\beta F(r_o))$. A perturbation of ς is applied to equation (3.31) and expanding in powers of ς it is found that

$$\frac{d^2\varsigma}{dt^2} - \frac{h^2}{r_o^3}\left\{1 - \frac{3\varsigma}{r_o}\right\} = -\frac{\mu}{r_o^2}\left\{1 - \frac{2\varsigma}{r_o}\right\}\left\{1 - \beta F(r_o) - \beta\left.\frac{dF}{dr}\right|_{r=r_o} \varsigma\right\} \tag{3.35}$$

Substituting for h^2 and retaining linear terms only a modified variational equation is obtained, viz

$$\frac{d^2\varsigma}{dt^2} - \beta\frac{\mu}{r_o^3}\left\{F(r_o) + r_o\left.\frac{dF}{dr}\right|_{r=r_o} - 1\right\}\varsigma = 0 \tag{3.36}$$

which may be written more compactly as

$$\frac{d^2\varepsilon}{dt^2} - P(r_o)\varepsilon = 0 \tag{3.37}$$

Evaluating the derivative of F(r) and making the substitution $\nu=(R_o/r_o)^2$ the function $P(r_o)$ may be written as

$$P(r_o) = \frac{\mu}{r_o^3} Q(\nu) \ , \ Q(\nu) = 2\beta\left\{\frac{1}{\nu}\left(1-(1-\nu)^{3/2}\right)-(1-\nu)^{1/2}\right\}-1 \tag{3.38}$$

For stability it is required that the variational equation has purely imaginary eigenvalues so that $Q(\nu)<0$, $0<\nu<1$. This condition may then be written in terms of an inequality on β, viz

$$\beta(\nu) < \left\{ \frac{2}{\nu}\left(1-(1-\nu)^{3/2}\right) - 2(1-\nu)^{1/2} \right\}^{-1} \tag{3.39}$$

Equation (3.39) divides the parameter space (β,r) into two distinct regions of stability and instability, as shown in Figure 3.4. The regions of stability and instability are partitioned by β_*, defined by an equality in equation (3.39). There is a further hyperbolic region defined by $\beta>\beta_c(r)$. The solution to the variational equation in the regions of stability and instability is of therefore the form

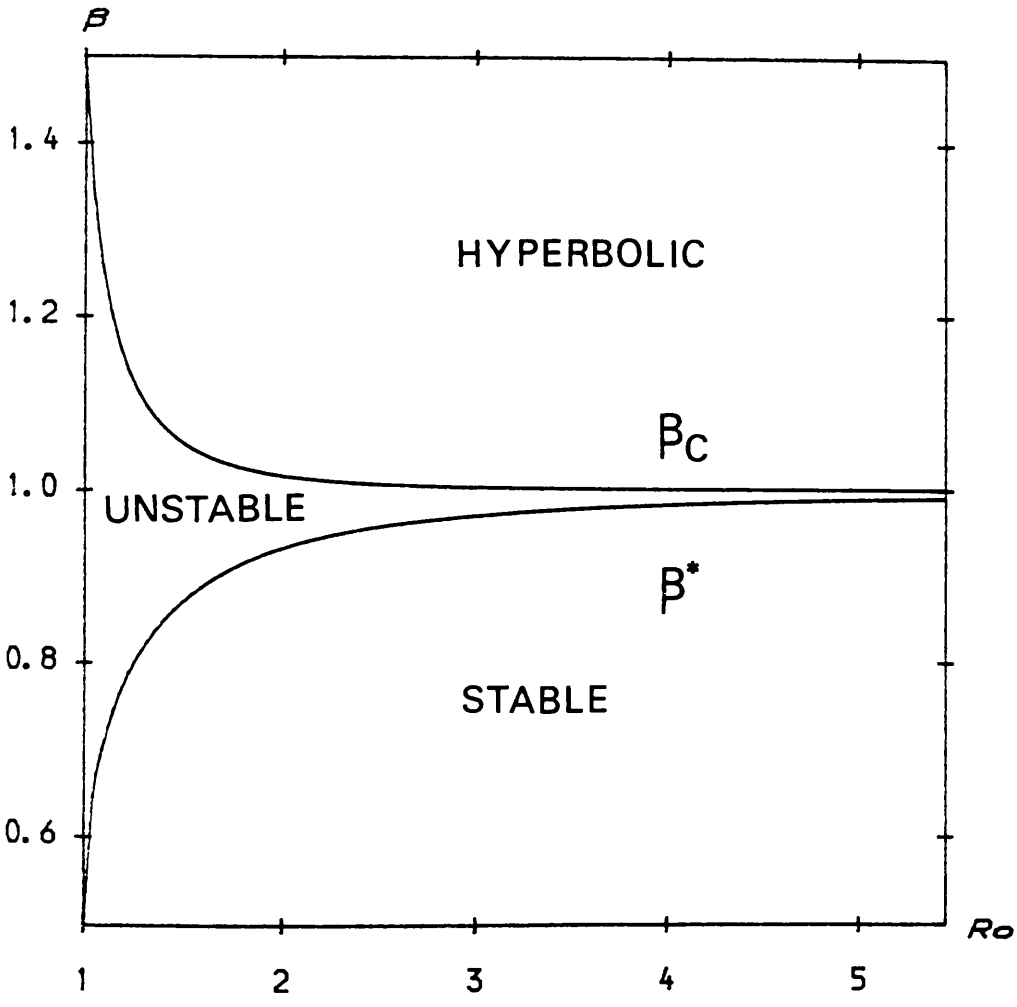
$$\varepsilon(t) = \sum_{j=1}^{2} \varepsilon_{oj}\, e^{\omega_j t} \tag{3.40}$$

where the eigenvalues of the system $\omega_{1,2}$ are defined by

$$\omega_{1,2} = \pm\omega_o\, Q(r/R_o)^{1/2} \tag{3.41}$$

It can be seen from Figure 3.4 that for all but the closest heliocentric

Figure 3.4



The three regions of the sail parameter space (β,r) with stable, unstable and hyperbolic regions partitioned by β^* and β_c respectively.

distances the condition for instability is confined to a very narrow region of the sail parameter space close to β=1, (ie. a narrow range of long orbital periods close to a stationary solution). In general therefore, a very specific set of parameters are required for an unstable orbit.

Substituting for h in equation (3.30) it is found that

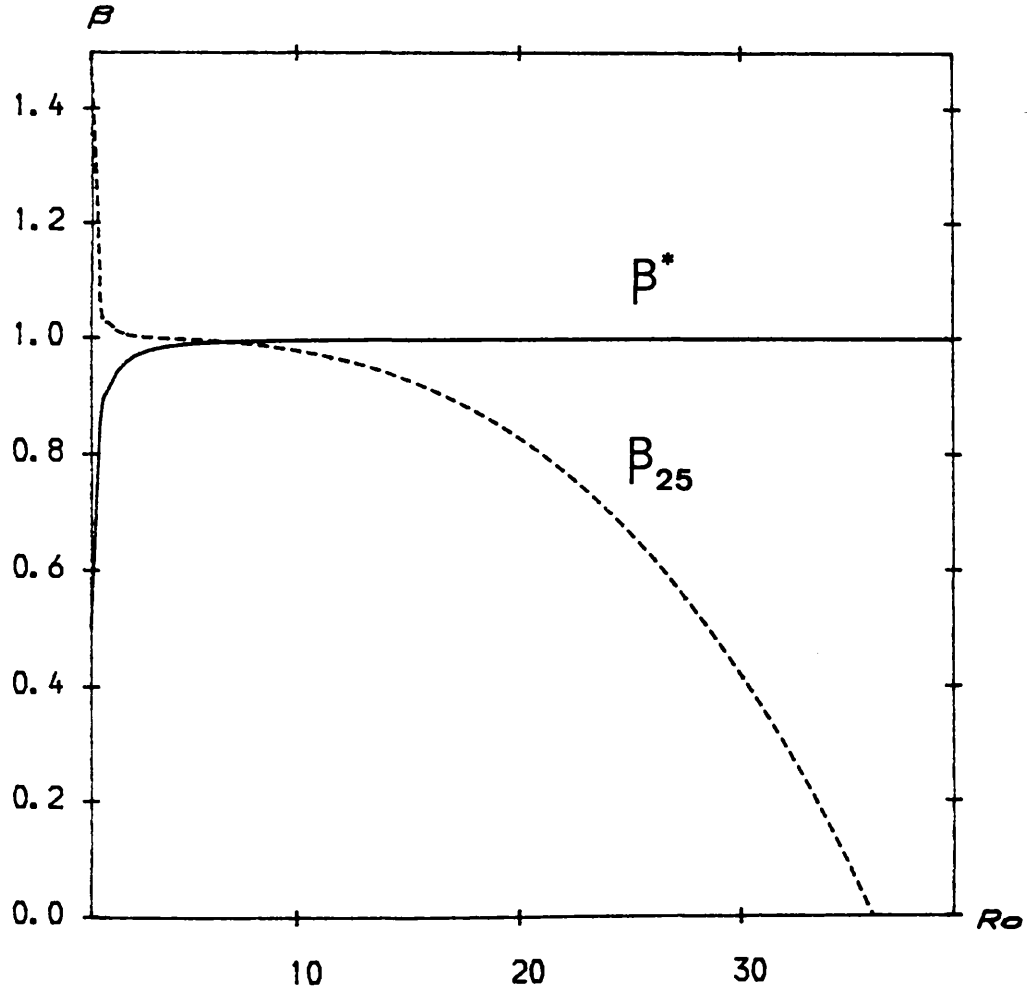
$$\omega = \omega_o \, (1 - \beta F(r))^{1/2} \qquad (3.42)$$

where ω is the angular velocity of a circular solar sail orbit with loading parameter β at a heliocentric distance r. For $r \gg R_o$ the function $F(r) \triangleq 1$ and so for $\beta > \beta^*$ the orbital angular velocity $\omega \ll 1$. If the sail then orbits with a higher angular velocity then a smaller value of β, within the stable region of the parameter space, will be required to maintain a heliocentric distance r. The value of the spacecraft orbital angular ω_c velocity required to obtain stability can be obtained by substituting for β^* in equation (3.42), viz

$$\omega_c = \omega_o \, (1 - \beta^*(r) F(r))^{1/2} \qquad (3.43)$$

However, a sail orbital angular velocity of less than that required for stability may be desired and so a value of β within the unstable region of the parameter space may be required. For the case of a solar sail in a heliostationary orbit following the 25 day solar equatorial rotation the value of β required to obtain this orbit is shown as a function of the sail heliocentric distance in Figure 3.5. It can be seen that $\beta \to 0$ as $r \to 35.76 \; R_o$, corresponding to a Keplerian orbit with a period of 25 days at 0.167 AU under the action of solar gravity only. However, the required value of β crosses into the unstable

Figure 3.5



Intersection of β_{25}, the sail loading parameter required for a 25 day orbit synchronous with the solar rotation, with β^* the sail loading parameter defining the region of instability. The intersection occurs at a heliocentric distance of 6.7 R_O.

region of the parameter space at a heliocentric distance of 6.7 R_o (0.03 AU) so that any 25 day solar sail orbit at a closer heliocentric distance will be necessarily unstable. Similarly for a sail in an Earth synchronous one year orbit the region of instability is bounded by a heliocentric distance of 19.1 R_o (0.09 AU). Therefore, it is concluded that if a solar sail orbits with even a small angular velocity it will require a value of $\beta<\beta^*$ to maintain the required heliocentric distance and so will be dynamically stable. However, when some particular orbital period is required, such as the heliostationary or Earth synchronous orbits there will always be a region of instability.

It has been shown that a dynamically unstable solar sail may be stabilised, depending on the required orbital period and heliocentric distance, if it has a small orbital angular velocity and that any perturbations will then result merely in periodic oscillations about its nominal circular orbit with a frequency given by equation (3.41). For solar sails stationary in the ecliptic plane this may be acceptable or even necessary for many purposes, such as a solar observation mission utilising the 25 day heliostationary orbit. However, for sails above the solar poles where the mission objective would be to have continuous observations of the poles, it would not be desirable to have the sail orbiting and so the sail would have to be actively controlled to remain at a dynamically unstable stationary ponit.

3.5 Time Varying Solar Radiation Pressure

Now that the dynamical effects of the finite angular size of the solar disk have been investigated the effects of variations in the solar luminosity, which relate directly to the solar radiation pressure, will be examined.

The solar luminosity has a well known and well defined eleven year period of variation due to internal physical processes. However, over short timescales the solar luminosity varies in an extremely non-periodic manner due to flares, sunspots etc. To attempt to obtain some understanding of how the time varying luminosity L(t) affects solar sail dynamics a simple sinusoidal variation will be superimposed upon the mean solar luminosity L_0, viz

$$L(t) = L_0(1 + \epsilon\cos(\Omega t)) \tag{3.44}$$

where $\epsilon \ll 1$ is the fractional variation in the solar luminosity with a frequency Ω. This functional form of the solar luminosity has in fact some empirical justification in that there appears to be variations with $\epsilon \approx 10^{-3}$ due to sunspot features. The sunspot features cover a small part of the solar disk therefore reducing the solar luminosity. These variations have an associated timescale of order 25 days due to the solar rotation carrying the sunspot features across and behind the solar disk, Wilson et. al (1981). It will be assumed that $r \gg R_0$ so that an inverse square solar radiation pressure model may be used.

3.5.1 Consequences for Stability: 1-Dimensional Motion

For a solar sail with $\beta=1$ at a stationary position $\mathbf{r}_0$ the one-dimensional dynamical equation with a time varying solar luminosity is given by (cf. equation (3.16))

$$\frac{d^2\mathbf{r}}{dt^2} = -\mu(1-\beta)\frac{\mathbf{r}}{|\mathbf{r}|^3} + \mu\frac{\mathbf{r}}{|\mathbf{r}|^3}\epsilon\beta\cos(\Omega t) \tag{3.45}$$

Substituting for β, applying a perturbation $\boldsymbol{\varsigma}$ such that $\mathbf{r}_0 \to \mathbf{r}_0 + \boldsymbol{\varsigma}$ and linearising about the stationary point $\mathbf{r}_0$ a variational equation with a

periodic coefficient is obtained. Considering a radial perturbation only it is found that

$$\frac{d^2\delta}{dt^2} = \epsilon \frac{\mu}{r_o^2}\left\{1 - \frac{2\delta}{r_o}\right\} \cos(\Omega t) \qquad (3.46)$$

which may be more conveniently written as

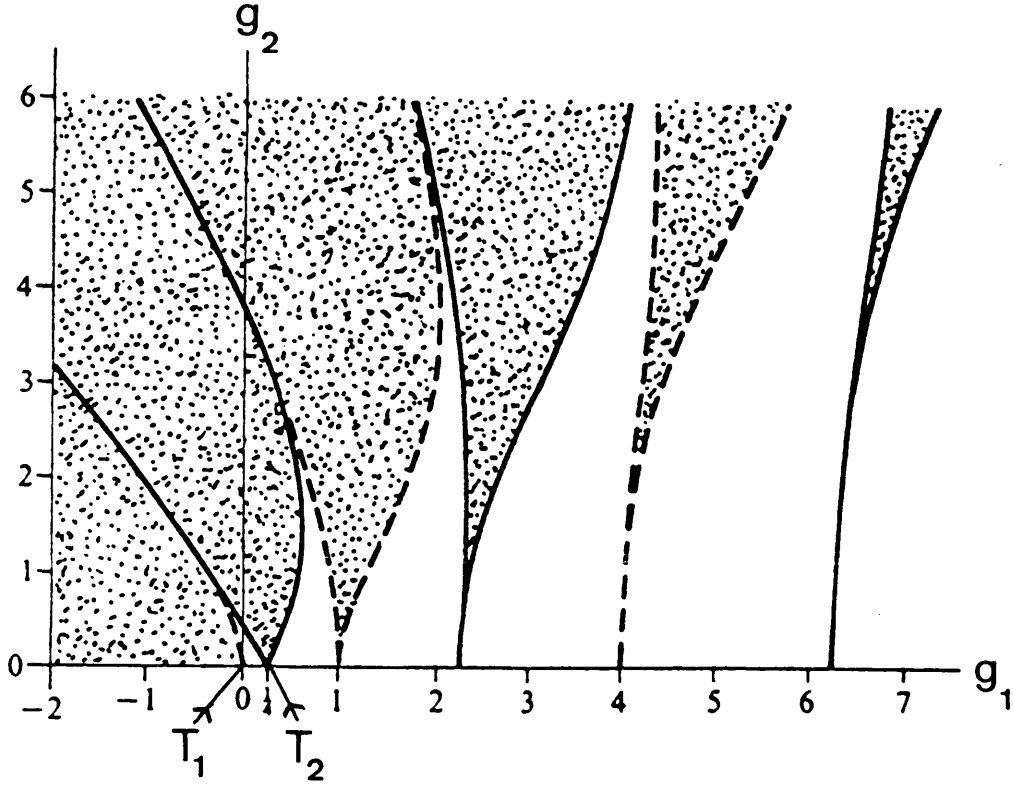
$$\frac{d^2\delta}{dt^2} + (\lambda_1\delta + \lambda_2)\cos(\Omega t) = 0 \ , \quad \lambda_1 = 2\epsilon\omega_o^2 \ , \quad \lambda_2 = -\epsilon r_o\omega_o^2 \qquad (3.47)$$

This equation may be reduced to a form of Mathieu's equation by using the variable transformation $\nu=\lambda_1\delta+\lambda_2$ and re-scaling the time variable such that $t'=\Omega t$

$$\frac{d^2\nu}{dt'^2} + g_2\cos(t')\nu = 0 \ , \quad g_1 = 0 \ , \quad g_2 = 2\epsilon\left\{\frac{\omega_o}{\Omega}\right\}^2 \qquad (3.48)$$

The stability characteristics of this type of equation may be investigated using Floquet theory, Jordan and Smith (1987). In this analysis the solution to equation (3.48) is written as the product of a purely periodic and an exponential solution, with the exponent determining the stability of the stationary solution. Whether the exponent is real or purely imaginary (unbound or bound solutions) depends on the values of the two coefficients $g_{1,2}$. It is found using Floquet theory that for stable, bound solutions to equation (3.48) it is required that $g_2<0.5$ (with $g_1=0$), as shown in Figure 3.6. That is, g_2 must lie between the two transition curves $T_{1,2}$ on the $g_1=0$ axis. This condition will always be satisfied since $\epsilon\approx10^{-3}$ at maximum and the variation has a timescale of order 25 days. Since this frequency is far greater than the Keplerian orbital angular velocity ω_o (consistent with the assumption that $r\gg R_o$) $0<g_2<0.5$ so that the

Figure 3.6



Stability map for Mathieu's equation, viz

$$\frac{d^2x}{dt^2} + (g_1 + g_2\cos(t))x = 0$$

The shaded region corresponds to unbound solutions and the unshaded region to bound solutions. The zero and first order transition curves $T_{1,2}$ are shown, (—— Solution of period 4π exists, ---- Solution of period 2π exists), (Adapted from Jordan and Smith (1987)).

stationary solution does not become unstable with exponential growth. It is therefore concluded that short period time variations in the solar luminosity have little effect on the stability of one-dimensional stationary solutions.

3.5.2 Consequences for Stability: 2-Dimensional Motion

For the circular orbital case a similar analysis can be carried out using the time varying solar luminosity. The dynamical equations now become (cf. equations (3.29))

$$\frac{d^2r}{dt^2} - r\left\{\frac{d\theta}{dt}\right\}^2 = -\frac{\mu}{r^2} + \beta\frac{\mu}{r^2}\left(1 + \epsilon\cos(\Omega t)\right) \qquad (3.49a)$$

$$\frac{1}{r}\frac{d}{dt}\left\{r^2\frac{d\theta}{dt}\right\} = 0 \qquad (3.49b)$$

These equations may be reduced to a single radial equation by again using the angular momentum integral $h=r^2(d\theta/dt)$, viz

$$\frac{d^2r}{dt^2} - \frac{h^2}{r^3} = -\frac{\mu}{r^2}\left((1-\beta) - \beta\epsilon\cos(\Omega t)\right) \qquad (3.50)$$

For an initial circular orbit at heliocentric distance r_o the spacecraft orbital angular momentum per unit mass is given by $h^2=\mu r_o((1-\beta)-\epsilon\cos(\Omega t))$ so that only the time averaged angular momentum is conserved. Applying a perturbation $r_o \rightarrow r_o+\xi$ and linearising about the initial circular orbit it is found that

$$\frac{d^2\xi}{dt^2} - \frac{h^2}{r_o^3}\left\{1 - \frac{3\xi}{r_o}\right\} = -\frac{\mu}{r_o^2}\left\{1 - \frac{2\xi}{r_o}\right\}\left\{(1-\beta) - \beta\epsilon\cos(\Omega t)\right\} \qquad (3.51)$$

Substituting for h^2 and retaining linear terms only a variational equation is obtained

$$\frac{d^2\delta}{dt^2} + \frac{\mu}{r_o^2}\left\{(1 - \beta) - \beta\epsilon\cos(\Omega t)\right\}\delta = 0 \qquad (3.52)$$

It has again been assumed that h remains constant with the applied perturbation so that the azimuthal drift of the spacecraft has been ignored. Equation (3.52) may be reduced directly to Mathieu's equation by transforming the time variable such that $t'=\Omega t$, viz

$$\frac{d^2\delta}{dt^2} + (g_1+g_2\cos(t'))\delta = 0 \ , \ g_1 = (1-\beta)\left\{\frac{\omega_o}{\Omega}\right\}^2, \ g_2 = -\beta\epsilon\left\{\frac{\omega_o}{\Omega}\right\}^2 \qquad (3.53)$$

where $g_{1,2} \ll 1$. For small coefficient values the transition curves $T_{1,2}$ of Figure 3.6 may be approximated, Jordan and Smith (1987). For the first transition curve T_1 the approximation is given by a parabolic relation $g_1=-(1/2)g_2^2$ and the stability condition is therefore

$$g_1 > -\frac{1}{2}g_2^2 + O(g_2^3) \qquad (3.54)$$

where $g_2^3 \approx 10^{-9}$ in this case. Since $g_1>0$ for $0<\beta<1$ this condition is always satisfied. The second transition curve T_2 has a linear approximation $g_1=(1/4)+(1/2)g_2$ and a resulting stability condition

$$g_1 < \frac{1}{4} + \frac{1}{2}g_2 + O(g_2^2) \qquad (3.55)$$

Substituting for $g_{1,2}$ the resulting condition on β is given by

$$\beta > \frac{(\omega_o/\Omega)^2 - (1/4)}{(\omega_o/\Omega)^2\,(1 - (\epsilon/2))} \qquad (3.56)$$

To first order in ϵ this condition may then be more conveniently written as

$$\beta > 1 - \frac{1}{4}\left\{\frac{\Omega}{\omega_o}\right\}^2 + \frac{\epsilon}{2}\left\{1 - \frac{1}{4}\left\{\frac{\Omega}{\omega_o}\right\}^2\right\} \qquad (3.57)$$

Therefore, for short period, sunspot related fluctuations in the solar luminosity $(\Omega/\omega_o)^2 \gg 1$, so that for $0<\beta<1$ the sail motion is always stable with purely imaginary characteristic exponents.

It has been shown then that short period fluctuations in the solar luminosity have little effect on solar sail orbital stability provided the timescale of fluctuation is far less than the sail orbital period. Even for the short period 25 day orbit following the solar rotation there will be little effect on orbital stability since the sail is co-rotating with the sunspot features which give rise to the luminosity variations. However, for the long period eleven year variations in the solar luminosity the sail motion will in fact become unstable, although the instability timescale will be extremely long. It can be seen from Figure 3.6 that for $\epsilon \ll 1$, and so $g_2 \ll 1$, the sail motion will become unstable for $g_1 < 0.25$. Therefore if, for example, the sail has an eleven year orbital period so that, $(\Omega/\omega_o)^2 \triangleq 1$ the motion will become unstable for $\beta > 0.75$.

3.6 Conclusions

It has been shown in this chapter that due to the finite angular size and limb darkening of the solar disk the solar radiation pressure exerted on a solar sail spacecraft does not have an inverse square variation with heliocentric distance, as assumed in previous studies. In general the effect of this deviation will be to introduce small errors into numerical calculations of solar sail trajectories. However, by identifying that this deviation of the functional form of the solar radiation pressure exists more accurate nominal solar sail trajectories

may be obtained. This is particularly important for trajectories in the inner solar system and for any future solar sail missions with advanced sail materials making extremely close passes (<0.1 AU) to the Sun. Such close passes are found for time optimal trajectories to the outer planets, as discussed in section 1.5.2.

Although the actual magnitude of the deviation of the solar radiation pressure from an inverse square form is small except at close heliocentric distances, it has been shown that the one-dimensional dynamics of a stationary solar sail is not as simple a system as is at first thought. A solar sail of a given loading now has only one possible stationary solution that is exponentially unstable. Furthermore, with the sail in a circular orbit it has been shown that for a sufficiently large orbital angular velocity the sail orbit becomes linearly stable against perturbations. However, with specific requirements on the sail orbital period there exists a region within which the sail orbit necessarily becomes unstable.

Finally, it has been shown that short period time variations in the solar luminosity appear to have little effect on solar sail stability in stationary and circular orbital configurations. However, for long period variations with the timescale of variation of the same order as the sail orbital period the motion can become unstable with a long instability timescale.

4. SOLAR SAIL HELIOCENTRIC HALO ORBITS

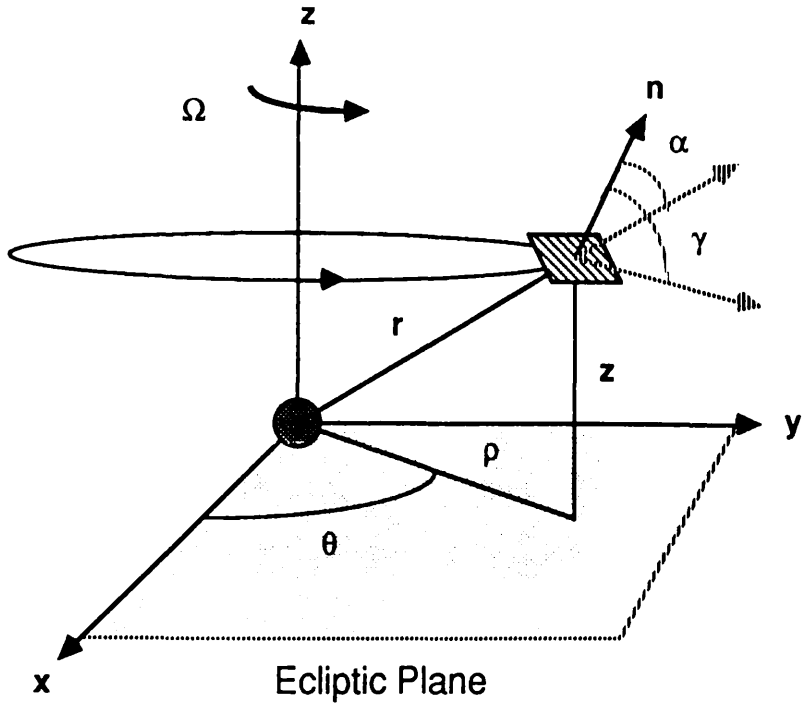
4.1 Introduction

In this chapter the dynamics of a new mode of operation of solar sail spacecraft is discussed, that of heliocentric halo orbits. These orbits are equivalent to circular solar sail orbits, but displaced out of the ecliptic plane. They are achieved by orienting the sail so that a component of the solar radiation pressure force is directed out of the orbital plane, Figure 4.1. The heliocentric halo orbit is unlike other solar sail orbits in that it is essentially a stationary solution to the dynamical equations, unlike transfer trajectories which have a time varying sail attitude and boundary conditions to be satisfied. Since the heliocentric halo orbit is the progenitor of the geocentric halo orbit families the underlying dynamics will be discussed in some detail.

It will be demonstrated that the spacecraft halo orbital period T, halo amplitude ρ and out-of-plane displacement distance z may be chosen independently with a suitable choice of Sun-sail pitch angle α and loading parameter β. The sail orbital period may then be chosen to be synchronous with a Keplerian orbit of the same heliocentric distance or chosen to be some particular fixed value, such as a one year Earth synchronous halo orbit. Perhaps more importantly, the halo orbit period may be chosen to minimise the sail loading required for a given set of halo orbit parameters (ρ,z).

The dynamical stability of the various modes of halo orbit is investigated and families of linearly stable orbits established. It is found that the optimised halo orbits are stable for all orbital parameters, whereas the fixed period halo orbits have a finite region of stability near the orbital plane. The Keplerian synchronous halo

Figure 4.1



Schematic geometry of a heliocentric halo orbit with the solar sail above the ecliptic plane at position $\mathbf{r}=(\rho,\theta,z)$. The sail attitude is defined by a unit vector $\mathbf{n}$ and the reference frame rotates with angular velocity $\boldsymbol{\Omega}$. The axis of the halo need not be normal to the ecliptic plane.

orbits are always unstable. Simple control schemes are then developed for the unstable halo orbit families. It is found that by fixing the Sun-sail pitch angle α all of the halo orbit families become stable, although there is no damping of initial injection errors. A more complex feedback scheme to the sail pitch is developed to include velocity terms, but is found to be unsuitable due the long damping timescale. By allowing small changes in the sail area along with changes in sail attitude a robust and well damped control is obtained.

Lastly, by patching individual halo orbits together, complex and elaborate new trajectories may be generated. By patching four perpendicular halo orbits together the sail may be forced to follow the surface of a cube. Furthermore, by switching the sail attitude to a null configuration with $\alpha=\pi/2$ the halo orbit may be patched to a heliocentric Keplerian ellipse. The applications of such orbits are many and varied both for unique astronomical observations and for communications with spacecraft on interplanetary trajectories, as will be discussed in chapter 7.

4.2 Dynamical Equations and Their Solution

To investigate the dynamics of heliocentric solar sail halo orbits the dynamical equations will be considered in a heliocentric rotating reference frame (see appendix B). Stationary solutions to the dynamical equations will then be found in this co-rotating frame. These stationary solutions correspond to halo type orbits when viewed from an inertial frame. Since the orientation of the co-rotating frame is arbitrary, the axis of the halo may have any desired orientation with respect to the ecliptic plane.

Using an idealised spacecraft model with a planar, perfectly

reflecting sail a solar sail at position $\mathbf{r}$ in the co-rotating frame with angular velocity $\boldsymbol{\Omega}$ will be considered, as shown in Figure 4.1. The sail orientation is defined by the unit vector $\mathbf{n}$, fixed in the co-rotating frame and the ratio of the solar radiation pressure force to the solar gravitational force is given by the dimensionless sail loading parameter $\beta=\sigma^*/\sigma$. Since the sail orientation is fixed in the co-rotating frame the sail must rotate once per orbit with respect to an inertial frame. The vector dynamical equation of the sail in the co-rotating frame may then be written as

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) = \mathbf{a} - \nabla\Phi_2(|\mathbf{r}|) \qquad (4.1)$$

where the terms on the left represent the kinematic, coriolis and centrifugal accelerations respectively. These accelerations are equated to the solar radiation pressure and solar gravitational accelerations exerted on the sail. The two-body gravitational potential $\Phi_2(|\mathbf{r}|)$ and solar radiation pressure acceleration $\mathbf{a}$ are given by

$$\Phi_2(|\mathbf{r}|) = -\frac{\mu}{|\mathbf{r}|} \quad , \quad \mathbf{a} = \beta \frac{\mu}{|\mathbf{r}|^4}(\mathbf{r}.\mathbf{n})^2\,\mathbf{n} \qquad (4.2)$$

as derived in section 2.1. Since the solar radiation pressure acceleration can never be directed sunwards the constraint $\mathbf{r}.\mathbf{n}>0$ is imposed so that the Sun-sail pitch angle $|\alpha|<(\pi/2)$.

It is noted now that the centrifugal term in equation (4.1) is conservative and so may be written in terms of a scalar potential $\Psi(\mathbf{r})$ defined such that

$$\nabla\Psi(\mathbf{r},\Omega) = \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) \quad , \quad \Psi(\mathbf{r},\Omega) = -\frac{1}{2}\,|\boldsymbol{\Omega} \times \mathbf{r}|^2 \qquad (4.3)$$

A new modified potential, $U(\mathbf{r},\Omega)=\Phi_2(|\mathbf{r}|)+\Psi(\mathbf{r},\Omega)$, will now be defined so that a reduced dynamical equation is obtained, viz

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \nabla U(\mathbf{r},\Omega) = \mathbf{a} \qquad (4.4)$$

In the co-rotating frame a stationary solution is required so that the first two terms of equation (4.4) must vanish. Since the vector $\mathbf{a}$ is oriented in direction $\mathbf{n}$, taking the vector product of $\mathbf{n}$ with equation (4.4) it is found that

$$\nabla U(\mathbf{r},\Omega) \times \mathbf{n} = 0 \Rightarrow \mathbf{n} = \lambda \nabla U(\mathbf{r},\Omega) \qquad (4.5)$$

where λ is an arbitrary scalar multiplier. Using the normalisation condition $|\mathbf{n}|=1$, λ is identified as $|\nabla U(\mathbf{r},\Omega)|^{-1}$ so that the required sail attitude is defined by

$$\mathbf{n} = \frac{\nabla U(\mathbf{r},\Omega)}{|\nabla U(\mathbf{r},\Omega)|} \qquad (4.6)$$

Since the spacecraft is to have uniform azimuthal motion there can be no component of the vector $\mathbf{n}$, and so of $\mathbf{a}$, in the azimuthal direction. Therefore, the sail attitude may be conveniently described in terms of a single angle α between $\mathbf{n}$ and $\mathbf{r}$. Taking vector and scalar products of equation (4.6) with $\mathbf{r}$, it is found that

$$\tan \alpha(\mathbf{r},\Omega) = \frac{|\mathbf{r} \times \nabla U(\mathbf{r},\Omega)|}{\mathbf{r}.\nabla U(\mathbf{r},\Omega)} \qquad (4.7)$$

Similarly, the required sail loading parameter is obtained by taking a scalar product of equation (4.4) with $\mathbf{n}$, again requiring a stationary solution, viz

$$\beta(\mathbf{r},\Omega) = \mu^{-1}|\mathbf{r}|^4 \frac{\nabla U(\mathbf{r},\Omega).\mathbf{n}}{(\mathbf{r}.\mathbf{n})^2} \qquad (4.8)$$

Therefore, general vector valued functions for the sail attitude and loading have been obtained in terms of the two-body co-rotating potential $U(\mathbf{r},\Omega)$. For a given halo orbit period and spacecraft position the Sun-sail pitch angle and loading parameter can then be obtained.

If a system of heliocentric cylindrical coordinates, (ρ,θ,z), is now used the co-rotating two-body potential function may be written as

$$U(\rho,z;\Omega) = -\left\{\frac{1}{2}\rho^2\Omega^2 + \frac{\mu}{r}\right\} \quad , \quad r^2 = \rho^2 + z^2 \qquad (4.9)$$

where the sail orbital period is given by $T=(2\pi/\Omega)$. Evaluating the potential gradient in equations (4.7) and (4.8) it is found that

$$\tan\alpha(\rho,z;\Omega) = \frac{(z/\rho)\ (\Omega/\Omega_*)^2}{(z/\rho)^2 + (1 - (\Omega/\Omega_*)^2)} \quad , \quad \Omega_*{}^2=\mu/r^3 \qquad (4.10a)$$

$$\beta(\rho,z;\Omega) = \left\{1 + \left\{\frac{z}{\rho}\right\}^2\right\}^{1/2} \frac{((z/\rho)^2 + (1 - (\Omega/\Omega_*)^2)^2\ )^{3/2}}{((z/\rho)^2 + (1 - (\Omega/\Omega_*)^2)\ )^2} \qquad (4.10b)$$

where Ω_* is the angular velocity of a circular Keplerian orbit at heliocentric distance r. If the parameter μ is now chosen to be unity then the unit of distance becomes the astronomical unit and the unit of time becomes 2π Earth years. These relations may now be used to investigate the various heliocentric halo orbit modes.

4.3 Keplerian Synchronous Mode

For this mode of operation the spacecraft orbital period is chosen to be that of a Keplerian orbit of semi-major axis equal to the sail heliocentric distance r, (ie. $\Omega=r^{-3/2}$). Substituting for this functional form of Ω the required loading and Sun-sail pitch are obtained as

$$\tan \alpha(\rho,z) = \frac{\rho}{z} \tag{4.11a}$$

$$\beta(\rho,z) = \frac{r}{|z|} \tag{4.11b}$$

For a fixed sail loading, equation (4.11b) defines topologically nested surfaces of revolution about the z-axis. A section of these level surfaces of constant sail loading is shown in Figure 4.2 along with the required sail pitch. It can be seen that the sail pitch is such that the unit normal to the sail surface is required to be oriented perpendicular to the plane of the system. Therefore the constraint $\mathbf{r.n}>0$ is always satisfied. The sail loading increases with lower ecliptic latitudes (ie. smaller payloads at lower latitudes for a given sail design). For any set of orbital parameters a sail loading of $\beta>1$ is required so that this mode of operation is of little practical interest.

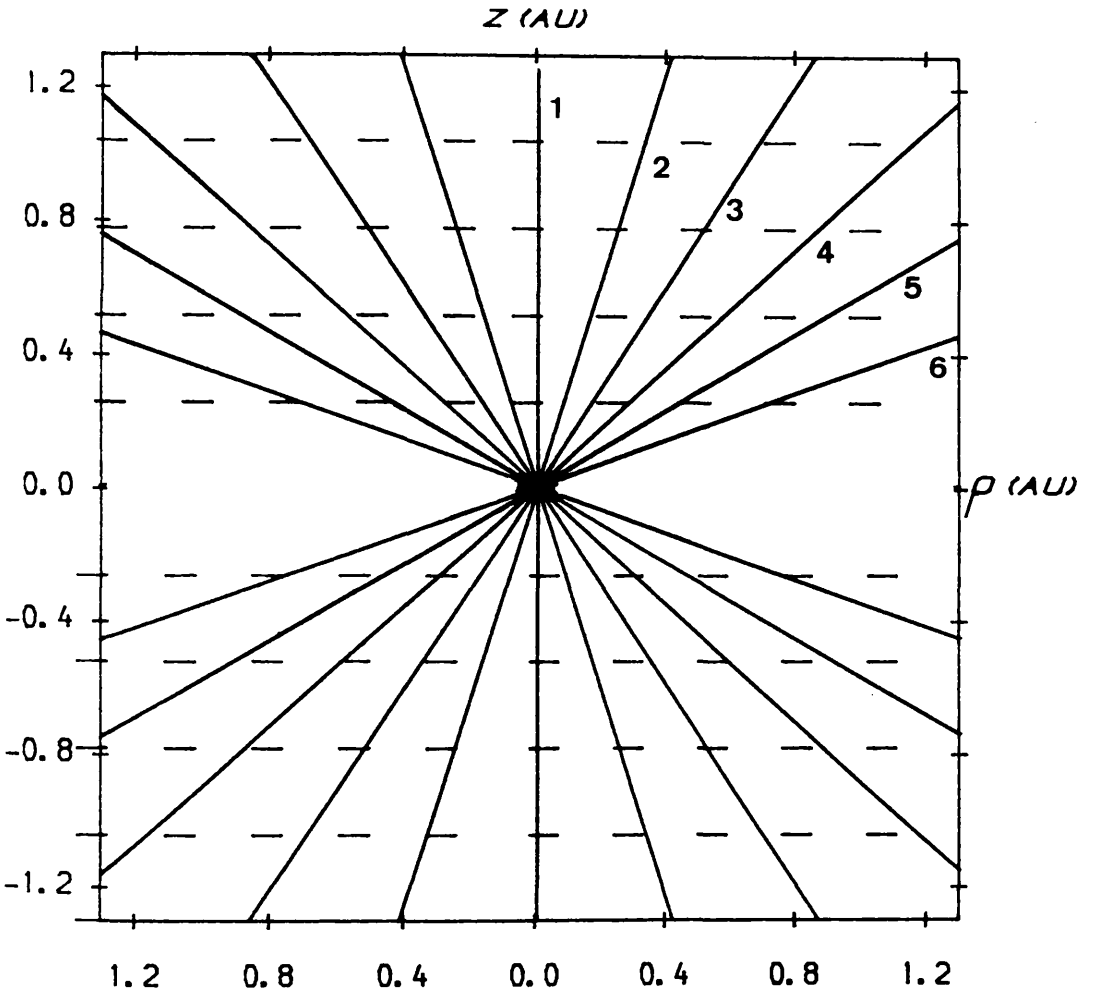
The surfaces of co-rotation (equal orbital period) are defined by spheres of constant radius so that the intersection of these spheres with the constant loading level surfaces defines regions where solar sails will orbit synchronously with each other. The sail loading parameter may be related to the actual sail mass per unit area using the relation $\beta=\sigma^*/\sigma$, (ie. $\sigma=1.53\beta^{-1}$ gm^{-2}).

It is again noted that owing to the symmetry of the problem the axis of the halo need not be normal to the ecliptic plane. In fact a halo orbit may be established about any axis passing through the origin so that off-axis halo orbits, inclined to the ecliptic are possible.

4.4 General Synchronous Mode

For this mode of operation the sail orbital period will be chosen to be some particular fixed value for all halo orbit parameters (ρ,z). A section of the equal loading level surfaces and the required sail pitch

Figure 4.2



Section of surfaces of constant sail loading and the required sail pitch for the Keplerian synchronous mode. The required sail loadings are given by; (1) 1.0 (2) 1.05 (3) 1.2 (4) 1.5 (5) 2.0 (6) 3.0.

is shown in Figure 4.3 for an Earth synchronous one year halo orbit. These surfaces are generated by setting $\Omega=1$ in equation (4.10b). Topologically, the full three-dimensional level surfaces of equal sail loading are a family of tori for $\beta<1$ and cylinders for $\beta>1$, nested around the z-axis. The topology change occurs at $\beta=1$ when the inner radius of the torus vanishes.

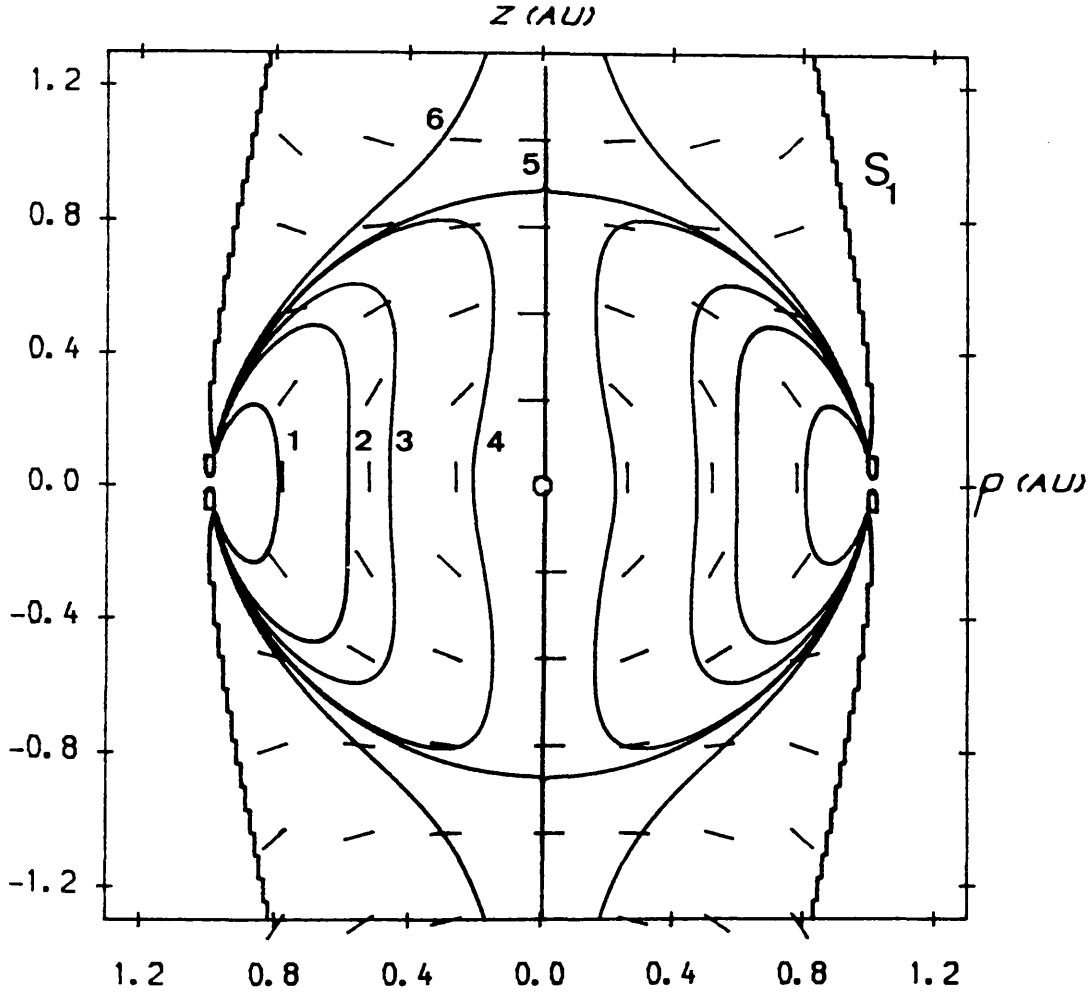
The nested tori intersect the ecliptic plane in a set of circular contours giving the required sail loading for co-rotation in the ecliptic plane with a period of one year. From equation (4.10a) it can be seen that in this case a radial sail attitude with zero pitch angle is required. Furthermore, along the z-axis it is seen from equations (4.10) that $\alpha=0$ and $\beta=1$ corresponding to a stationary solution at any heliocentric distance above the solar poles. This requires a spacecraft mass per unit area of 1.53 gm^{-2}. With the sail deployed on a halo orbit a lower sail loading, and so a larger mass per unit area, is required. This is however at the expense of a decrease in the potential out-of-plane distance. For example, for a halo orbit at $z=0.5$, $\rho=0.7$ AU a sail mass per unit area of 1.91 gm^{-2} is required.

The region of space in the co-rotating frame in which stationary solutions exist is bounded, being defined by the region interior to the surface $\mathbf{r}.\mathbf{n}=0$. It can be seen from equation (4.8) and Figure 4.3 that as this limiting surface is approached the sail loading parameter $\beta\to\infty$. The boundary is defined by

$$S(\rho,z) = \frac{\mu}{r} - \rho^2\Omega^2 = 0 \qquad (4.12)$$

Outside this surface $\mathbf{r}.\mathbf{n}<0$, corresponding to a Sun-sail pitch angle of $\alpha>(\pi/2)$, so that the total gravitational and centrifugal force is

Figure 4.3



Section of surfaces of constant sail loading and the required sail pitch for the Earth synchronous mode. The required sail loadings are given by; (1) 0.5 (2) 0.8 (3) 0.9 (4) 0.99 (5) 1.0 (6) 1.1. The outer contour S_1 represents the boundary $S(\rho,z)=0$.

outward. The solar radiation pressure force will then only augment these forces and so stationary solutions in the co-rotating frame will not be possible. Inside the surface $\mathbf{r}.\mathbf{n}>0$ and stationary solutions are possible with the gravitational, centrifugal and solar radiation pressure forces in equilibrium. The surface $S(\rho,z)=0$ is shown as the outer contour S_1 in Figure 4.3.

4.5 Optimal Halo Mode

To generate an optimal family of halo orbits the spacecraft orbital period will be treated as a free parameter of the system with respect to which the sail loading requirements may be minimized. Therefore, setting the derivative of β with respect to Ω to zero a quadratic in Ω^2 is obtained, viz

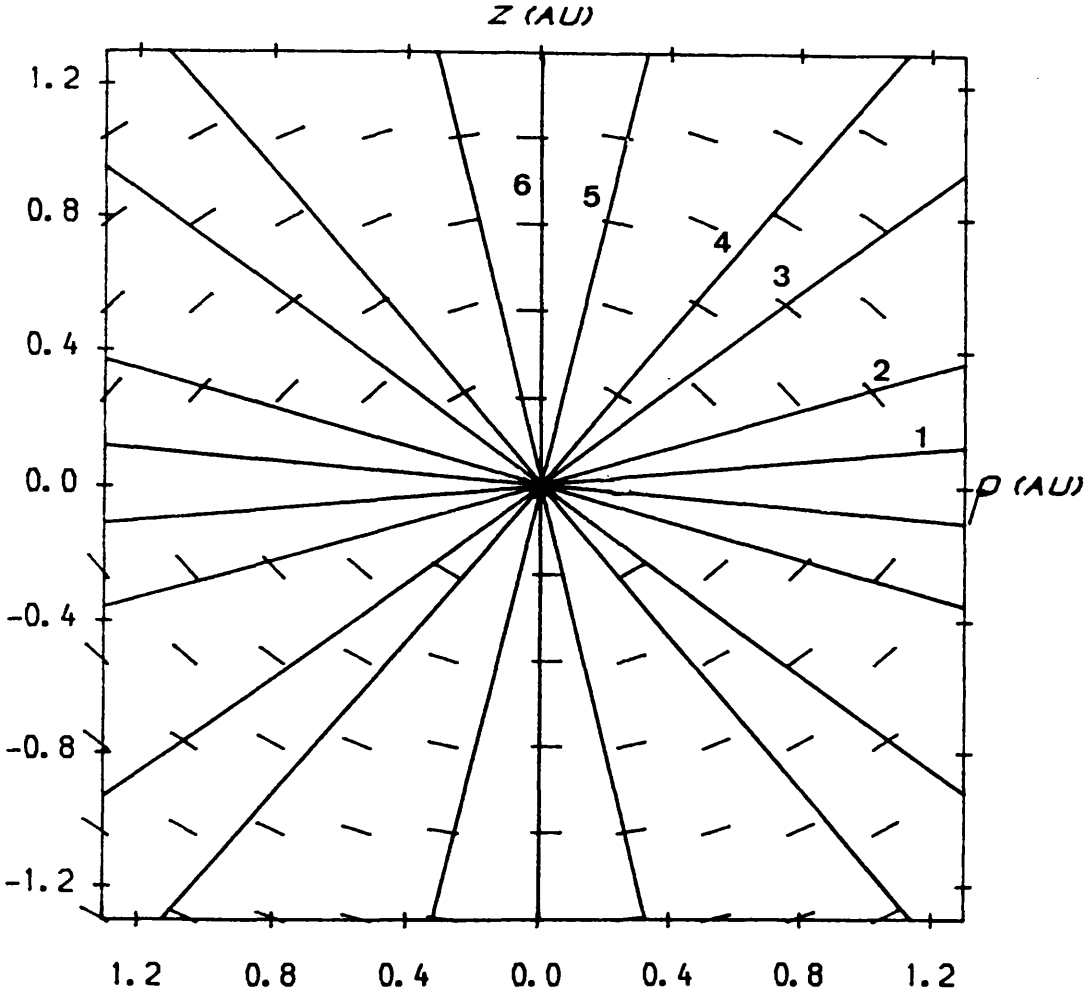
$$\frac{\partial\beta(\rho,z;\Omega)}{\partial\Omega}=0 \;\Rightarrow\; \Omega^4-\Omega^2\Omega_*^2\left\{2+3\left\{\frac{z}{\rho}\right\}^2\right\}+\Omega_*^4\left\{1+\left\{\frac{z}{\rho}\right\}^2\right\}=0 \qquad (4.13)$$

The quadratic gives two solutions for Ω^2, one of which fails to satisfy the condition $\mathbf{r}.\mathbf{n}>0$. The other solution always satisfies this condition. The required solution for the sail orbital angular velocity to give a minimum in the sail loading requirement is then given by

$$\Omega_{opt}=\Omega_*\left\{1+\frac{3}{2}\Lambda\right\}^{1/2}\left\{1-\left\{1-\frac{1+\Lambda}{(1+(3/2)\Lambda)^2}\right\}^{1/2}\right\}^{1/2},\quad \Lambda=\left\{\frac{z}{\rho}\right\}^2 \qquad (4.14)$$

Using this functional form of the sail orbital angular velocity surfaces of constant sail loading may again be generated from equation (4.10b). A section of these surfaces is shown in Figure 4.4 along with the required sail pitch. It can be seen that a halo orbit may be established at all points in space, always with $0<\beta\leqslant 1$. Therefore, by treating Ω as a free parameter of the system, halo orbits with large

Figure 4.4



Section of surfaces of constant sail loading and the required sail pitch for the optimal mode. The required sail loadings are given by; (1) 0.2 (2) 0.5 (3) 0.8 (4) 0.9 (5) 0.99 (6) 1.0.

out-of-plane displacements may be achieved with the required sail loading lower than for the synchronous case. Current sail designs allow sail loading values of up to $\beta \approx 0.3$, as discussed in section 1.4.3, which allows an optimised halo orbit with $z \approx 0.3$ AU at $\rho = 1.0$ AU. A solar sail on such a halo orbit has a period of 1.37 years.

4.6 Heliocentric Halo Orbit Stability

Now that the three halo orbit families have been established their stability characteristics will now be investigated. This is carried out by linearising the dynamical equations about a nominal halo orbit to obtain a variational equation. The variational equation describes the sail motion in the neighbourhood of the nominal halo orbit. Therefore, the stability of the trajectory may be determined by examining the eigenvalues of the variational equation. The non-linear dynamical equations may be linearised by perturbing the sail from its nominal halo orbit with the sail pitch fixed in the (ρ,z) frame. The perturbation is therefore applied with the inertial sail pitch angle γ fixed, as shown in Figure 4.1, so that the Sun-sail pitch angle α is variable.

Applying a perturbation $\boldsymbol{\delta}$ to a sail at an operating point $\mathbf{r}_o=(r_o,\theta_o,z_o)$ such that $\mathbf{r}_o \rightarrow \mathbf{r}_o+\boldsymbol{\delta}$ a linear variational equation is obtained from the non-linear dynamical equation, equation (4.4)

$$\frac{d^2\boldsymbol{\delta}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\boldsymbol{\delta}}{dt} + \nabla U(\mathbf{r}_o+\boldsymbol{\delta}) - \mathbf{a}(\mathbf{r}_o+\boldsymbol{\delta}) = 0 \qquad (4.15)$$

where the vector $\boldsymbol{\delta}=(\xi,\Psi,\eta)$ represents first order displacements in the co-rotating frame in the (ρ,θ,z) directions respectively. The potential gradient and the radiation pressure acceleration may be expanded in

trivariate Taylor series about the operating point $\mathbf{r}_o$ to first order as

$$\nabla U(\mathbf{r}_o+\boldsymbol{\varepsilon}) = \nabla U(\mathbf{r}_o) + \frac{\partial}{\partial \mathbf{r}} \nabla U(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{r}_o} \boldsymbol{\varepsilon} + O(|\boldsymbol{\varepsilon}|^2) \tag{4.16a}$$

$$\mathbf{a}(\mathbf{r}_o+\boldsymbol{\varepsilon}) = \mathbf{a}(\mathbf{r}_o) + \frac{\partial}{\partial \mathbf{r}} \mathbf{a}(\mathbf{r})\Big|_{\mathbf{r}=\mathbf{r}_o,\ \mathbf{n}=\mathbf{n}_o} \boldsymbol{\varepsilon} + O(|\boldsymbol{\varepsilon}|^2) \tag{4.16b}$$

Then, since $\nabla U(\mathbf{r}_o)=\mathbf{a}(\mathbf{r}_o)$ on the nominal halo orbit a linear variational system with constant coefficients is obtained, viz

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} + \mathbf{M}_1 \frac{d\boldsymbol{\varepsilon}}{dt} + (\mathbf{M} - \mathbf{N})\,\boldsymbol{\varepsilon} = \mathbf{0} \tag{4.17}$$

where $\mathbf{M}$ and $\mathbf{N}$, the gravity and radiation gradient tensors in equations (4.16) and the skew symmetric gyroscopic matrix $\mathbf{M}_1$ are given by

$$\mathbf{M}_1 = \begin{bmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{M} = \langle U_{ij} \rangle_{(i,j)\in(\rho,\theta,z)}, \quad \mathbf{N} = \langle a_{ij} \rangle \tag{4.18}$$

The tensor components U_{ij} are the (i,j) partial derivatives of the potential with respect to the cylindrical polar coordinates and a_{ij} is the j^{th} derivative of the i^{th} component of the solar radiation pressure acceleration. Owing to the azimuthal symmetry of the system all derivatives with respect to θ in the matrices $\mathbf{M}$ and $\mathbf{N}$ vanish. In component form the variational equation then becomes

$$\frac{d^2\xi}{dt^2} - 2\Omega\rho_o\frac{d\Psi}{dt} + L_{11}\xi + L_{13}\eta = 0 \tag{4.19a}$$

$$\frac{d^2\Psi}{dt^2} + \frac{2\Omega}{\rho_o}\frac{d\xi}{dt} = 0 \tag{4.19b}$$

$$\frac{d^2\eta}{dt^2} + L_{31}\xi + L_{33}\eta = 0 \qquad (4.19c)$$

where $L_{ij}=M_{ij}-N_{ij}$. This set of three coupled ordinary differential equations may be reduced to two by integrating equation (4.19b) to obtain

$$\frac{d\Psi}{dt} = -\frac{2\Omega}{\rho_0}(\xi - \xi_0) \qquad (4.20)$$

This equation is in effect a linearised version of Kepler's third law, giving the angular velocity of the sail relative to the nominal halo orbit. The equation may then be substituted into equation (4.19a) to eliminate the azimuthal coordinate. However, this leads to a constant term $4\Omega^2\xi_0$ in equation (4.19a) so that the variational system becomes non-homogeneous. The non-homogeneity can be easily removed by re-scaling through a change of variable

$$\xi' = \xi - \frac{4\Omega^2 L_{33}}{L^*_{11}L_{33} - L_{13}L_{31}}\xi_0 \qquad (4.21a)$$

$$\eta' = \eta + \frac{4\Omega^2 L_{13}}{L^*_{11}L_{33} - L_{13}L_{31}}\xi_0 \qquad (4.21b)$$

where $L^*_{11}=L_{11}+4\Omega^2$. Using this transformation a reduced variational system with a set of two coupled equations is obtained, viz

$$\frac{d^2}{dt^2}\begin{Bmatrix}\xi'\\ \eta'\end{Bmatrix} + \left[\begin{array}{c:c} L^*_{11} & L_{13} \\ \hdashline L_{31} & L_{33}\end{array}\right]\begin{Bmatrix}\xi'\\ \eta'\end{Bmatrix} = \begin{Bmatrix}0\\0\end{Bmatrix} \qquad (4.22)$$

By eliminating the azimuthal coordinate the order of the variationai system is reduced, however the sail is free to drift along the halo orbit. Therefore, the orbit cannot have Lyapunov type stability but

may at least be Poincaré stable (see appendix C). The coefficients of the matrix **L** are given by

$$L^*_{11} = 4\Omega^2 - \left\{ \left\{\Omega^2 - \frac{1}{r^3}\right\} + \frac{3\rho^2}{r^5} \right\} - \lambda_1\lambda_2\cos^2\gamma \tag{4.23a}$$

$$L_{13} = -\frac{3\rho z}{r^5} - \lambda_1\lambda_3\sin\gamma\cos\gamma \tag{4.23b}$$

$$L_{31} = -\frac{3\rho z}{r^5} - \lambda_1\lambda_2\sin\gamma\cos\gamma \tag{4.23c}$$

$$L_{33} = \left\{\frac{1}{r^3} - \frac{3z^2}{r^5}\right\} - \lambda_1\lambda_3\sin^2\gamma \tag{4.23d}$$

and the auxiliary coefficients λ_j (j=1,3) are given by

$$\lambda_1 = \frac{2\beta}{r^4}\,(\rho\cos\gamma + z\sin\gamma) \tag{4.23e}$$

$$\lambda_2 = 1 - \frac{2\rho^2}{r^2}\left\{1 + \frac{z}{\rho}\tan\gamma\right\} \tag{4.23f}$$

$$\lambda_3 = 1 - \frac{2z^2}{r^2}\left\{1 + \frac{\rho}{z}\cot\gamma\right\} \tag{4.23g}$$

The stability characteristics of the reduced variational system defined by equation (4.22) may now be investigated by calculating the system eigenvalues. This may be carried out by substituting an exponential solution of the form

$$\begin{Bmatrix} \xi' \\ \eta' \end{Bmatrix} = \begin{Bmatrix} \xi_0 \\ \eta_0 \end{Bmatrix} e^{\omega t} \tag{4.24}$$

Substituting this solution into equation (4.22) gives a matrix equation of the form

$$\begin{bmatrix} \omega^2 + L^*_{11} & \vdots & L_{13} \\ \cdots & \vdots & \cdots \\ L_{31} & \vdots & \omega^2 + L_{33} \end{bmatrix} \begin{Bmatrix} \xi_o \\ \eta_o \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{4.25}$$

For non-trivial solutions a vanishing secular determinant of the matrix equation is required. This then gives the characteristic polynomial of the system, viz

$$\omega^4 + tr(\mathbf{L})\omega^2 + det(\mathbf{L}) = 0 \tag{4.26}$$

with the trace of the system matrix $tr(\mathbf{L})=L^*_{11}+L_{33}$ and its determinant $det(\mathbf{L})=L^*_{11}L_{33}-L_{13}L_{31}$. The fundamental theorem of algebra implies that the characteristic polynomial has four complex roots ω_j (j=1,4), the four frequencies of the eigenmodes of the system. Formally these eigenvalues may be written as

$$\omega_{1..4} = \frac{\pm 1}{\sqrt{2}} \left\{ -tr(\mathbf{L}) \pm \left\{ tr(\mathbf{L})^2 - 4det(\mathbf{L}) \right\}^{1/2} \right\}^{1/2} \tag{4.27}$$

where the positive root gives a long period response and the negative root a short period response. The sail motion in the neighbourhood of the nominal halo orbit is then given by the superposition of the four eigenmodes as

$$\begin{Bmatrix} \xi' \\ \eta' \end{Bmatrix} = \sum_{j=1}^{4} \begin{Bmatrix} \xi_{oj} \\ \eta_{oj} \end{Bmatrix} e^{\omega_j t} \tag{4.28}$$

The stability characteristics of the families of halo orbits may now be investigated by numerically searching for regions with purely imaginary eigenvalues, $\omega_j^2<0$, (j=1,4) giving stable, bound oscillations in the (ρ,z) plane, (ie. $det(\mathbf{L})>0$ and $tr(\mathbf{L})>0$). Firstly, the stable

regions of the general synchronous mode with $\Omega=1$ will be mapped. Using the scale invariance of the system the stability of the Keplerian synchronous and optimal halo modes will then be determined. Setting $\Omega=1$ in equation (4.27) it is found that there exists a stable family of halo orbits with $\omega_j^2<0$, (j=1,4) near the ecliptic plane. This region is bounded by the surface C_1, the section of which is shown in Figure 4.5. Along the z-axis it is found that there exists a marginal instability since equation (4.19c) reduces to $(d^2n/dt^2)=0$.

For the Keplerian synchronous case the surface $\Omega=r^{-3/2}=1$ is a unit sphere, the section of which is shown in Figure 4.5 by the curve C_2. It can be seen that this surface lies outside the stable region of the map so that the $\Omega=1$ Keplerian synchronous halo orbits will be unstable, (it may in fact be easily shown that $\det(L)<0$ for $\Omega=\Omega_*$ in general). Furthermore, the surface defined by $\Omega_{opt}=1$, shown as C_3 in Figure 4.5, may also be generated and is found to lie within the stable region.

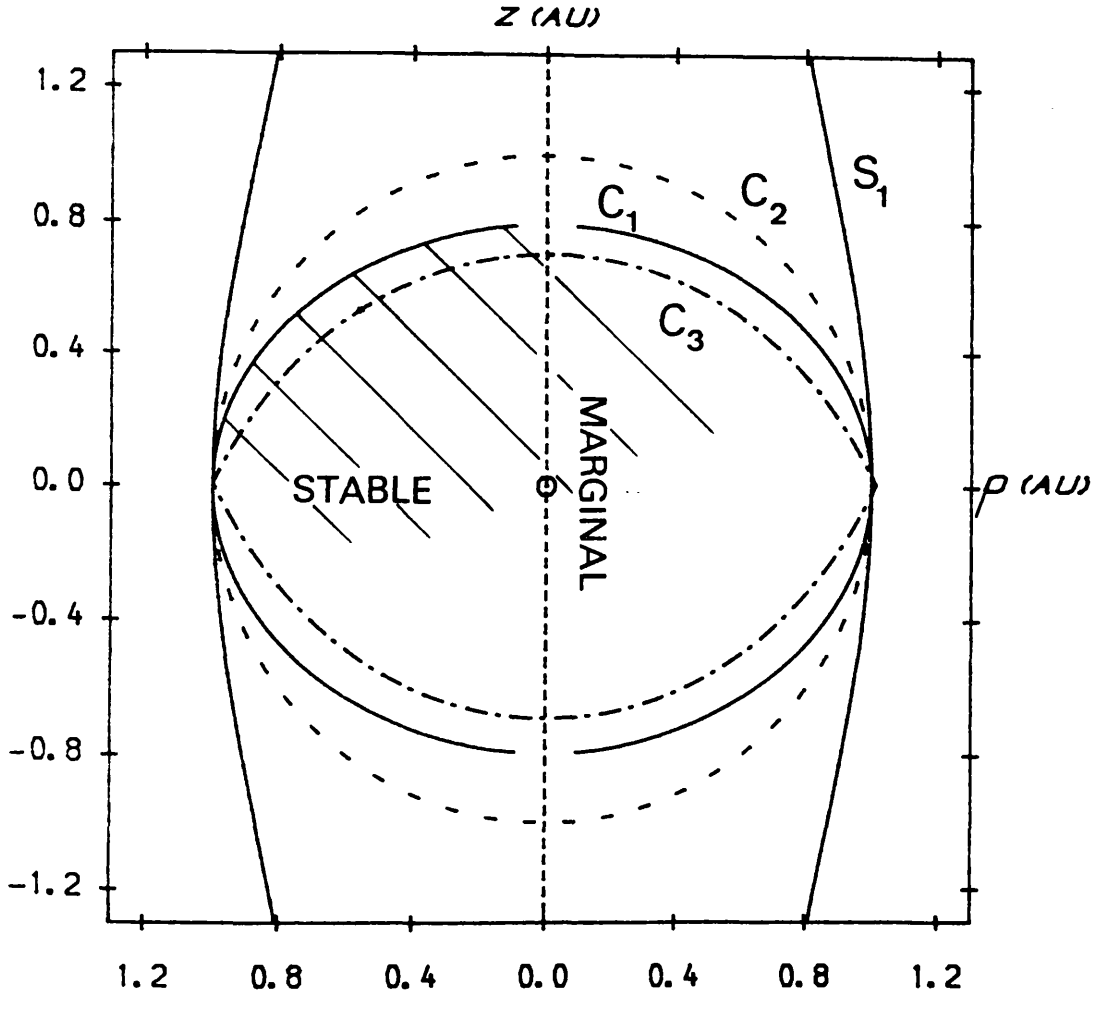
Examining equations (4.10) it can be seen that there exists a scaling law of the form

$$\rho_1 \rightarrow \rho_1 \left\{\frac{\Omega_1}{\Omega_2}\right\}^{2/3}, \quad z_1 \rightarrow z_1 \left\{\frac{\Omega_1}{\Omega_2}\right\}^{2/3} \quad \text{as} \quad \Omega_1 \rightarrow \Omega_2 \tag{4.29}$$

so that the problem is scale invariant to changes in Ω. Therefore, the hierarchy of the surfaces shown in Figure 4.5 will hold for any value of Ω. The Keplerian synchronous mode is therefore always unstable and the optimal mode is always stable.

The coupling of the halo amplitude perturbations to the azimuthal motion may be found by integrating equation (4.20), viz

Figure 4.5



Halo orbit stability map in the ρ-z plane. The contour C_1 defines the stable region for the Ω=1 mode and the contours $C_{2,3}$ define the surfaces given by $\Omega=r^{-3/2}=1$ and $\Omega_{opt}=1$ respectively. The stationary solutions along the z-axis have a marginal instability.

$$\Psi(t) = \Psi_o + \left\{\frac{2\Omega\xi_o}{\rho_o}\right\}t - \frac{2\Omega}{\rho_o}\int_o^t \xi(t')\,dt' \qquad (4.30)$$

where $\xi(t)$ is the response given by equations (4.28) and (4.21a). The first order drift in azimuthal position can then be obtained as

$$\Psi(t) = \Psi_o + \left\{\frac{2\Omega\xi_o}{\rho_o}\right\}\left\{1 - \frac{4\Omega^2 L_{33}}{L^*_{11}L_{33}-L_{13}L_{31}}\right\}t - \frac{2\Omega}{\rho_o}\sum_{j=1}^{4}\frac{\xi_{oj}}{\omega_j}e^{\omega_j t} \qquad (4.31)$$

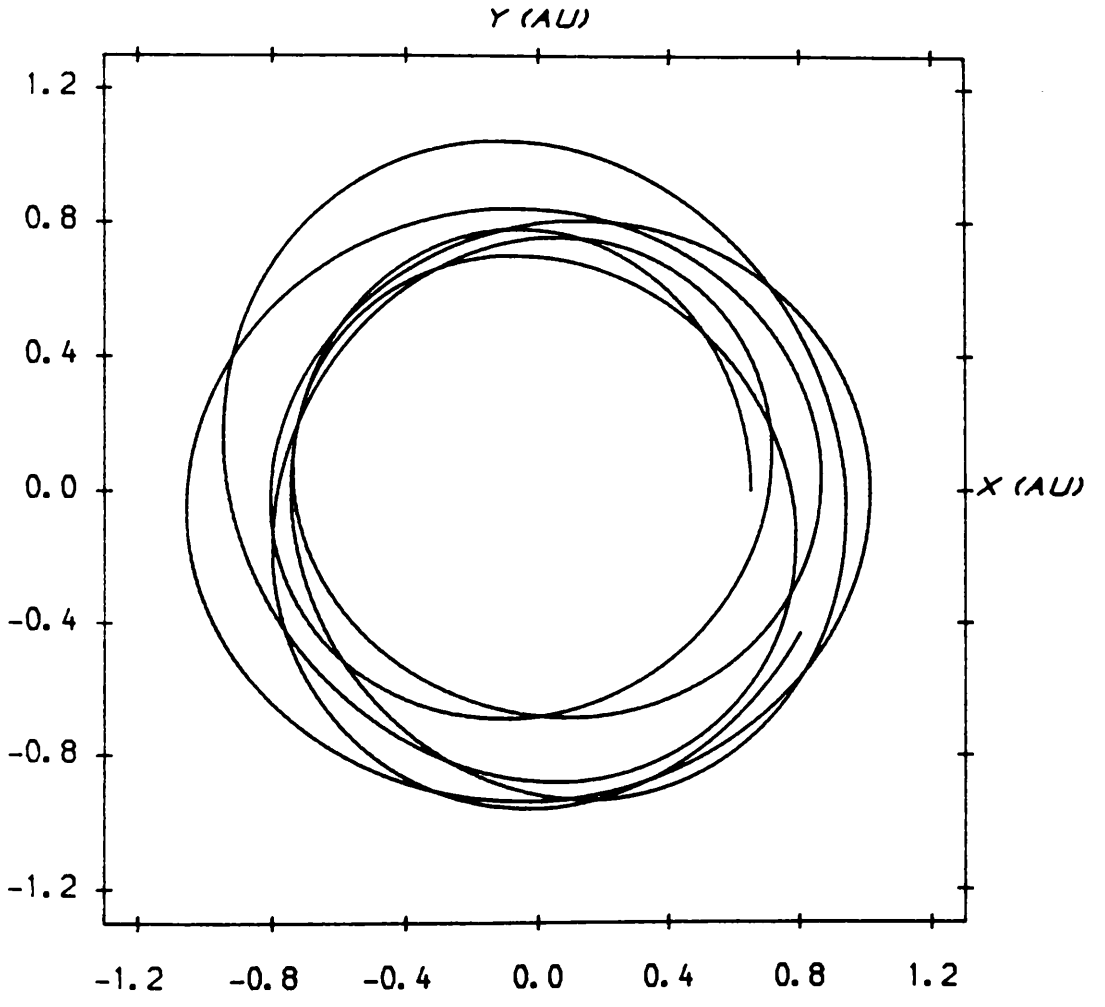
For the stable halo orbit families the azimuthal perturbations are of the form of periodic oscillations with a secular increase. This drift is due to the sail having an initial perturbation in the ρ coordinate so that the sail has a first order difference in orbital period from the nominal halo orbit. The sail motion is therefore constrained to a torus around the nominal orbit and has Poincaré type stability.

Typical stable and unstable responses for a one year Earth synchronous halo orbit are shown in Figures 4.6 and 4.7, using a numerical integration of the full non-linear dynamical equations. It can be seen that the unstable halo orbit has a rather long instability timescale of almost one year. The instability manifests itself in the sail falling sunward. For the stable case it is seen that even for a large initial perturbation outside the linear regime, the halo orbit remains bound with large amplitude in-plane and out-of-plane oscillations.

4.7 Heliocentric Halo Orbit Control

Since the stability analysis has shown that the halo orbit families have regions of instability, simple closed loop control schemes that ensure asymptotic stability will now be investigated. If the spacecraft azimuthal position is unimportant (as would be the case for a passive

Figure 4.6(a)



Stable response for a one year halo orbit with ρ=0.6, z=0.5 and injection errors of $\xi_o=\eta_o=5\times10^{-2}$; (a) x-y projection (b) x-z projection.

Figure 4.6(b)

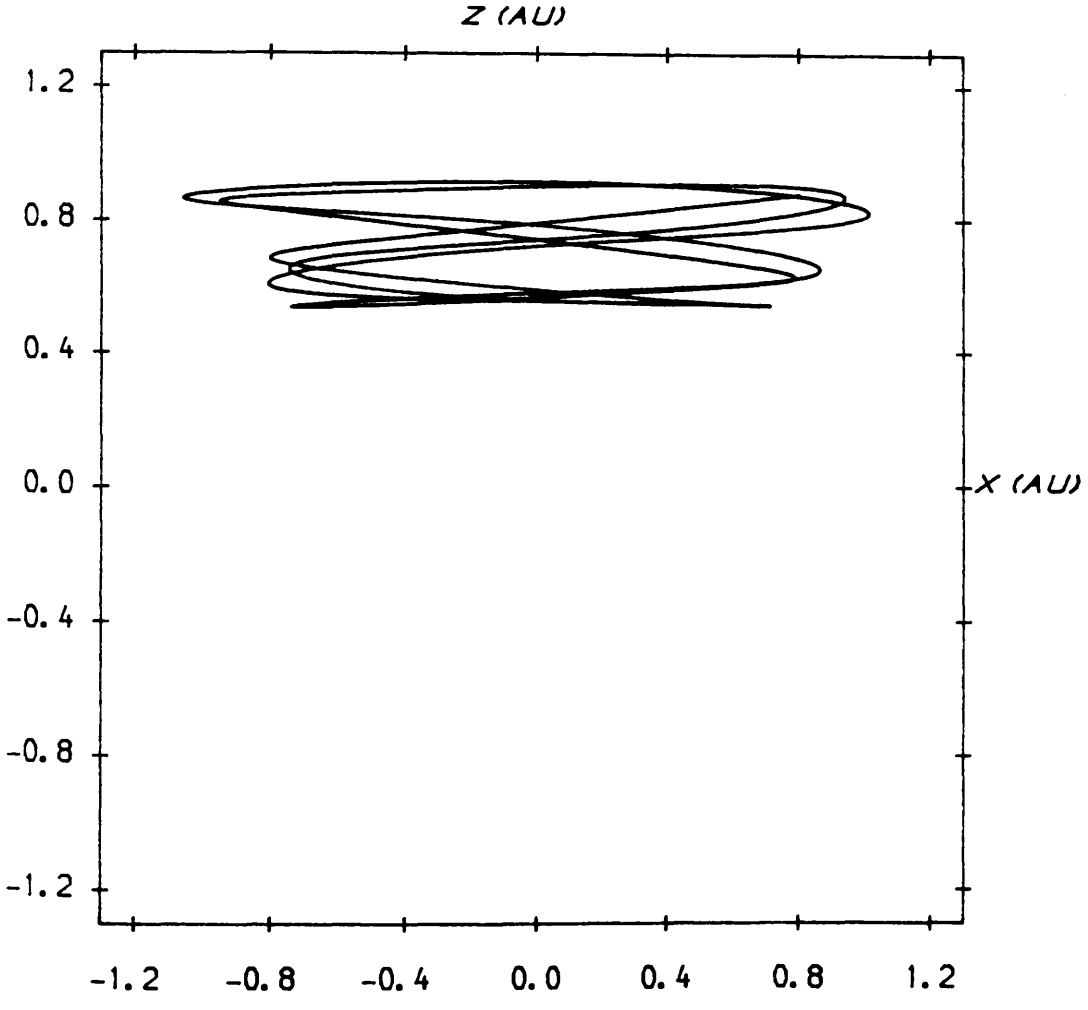
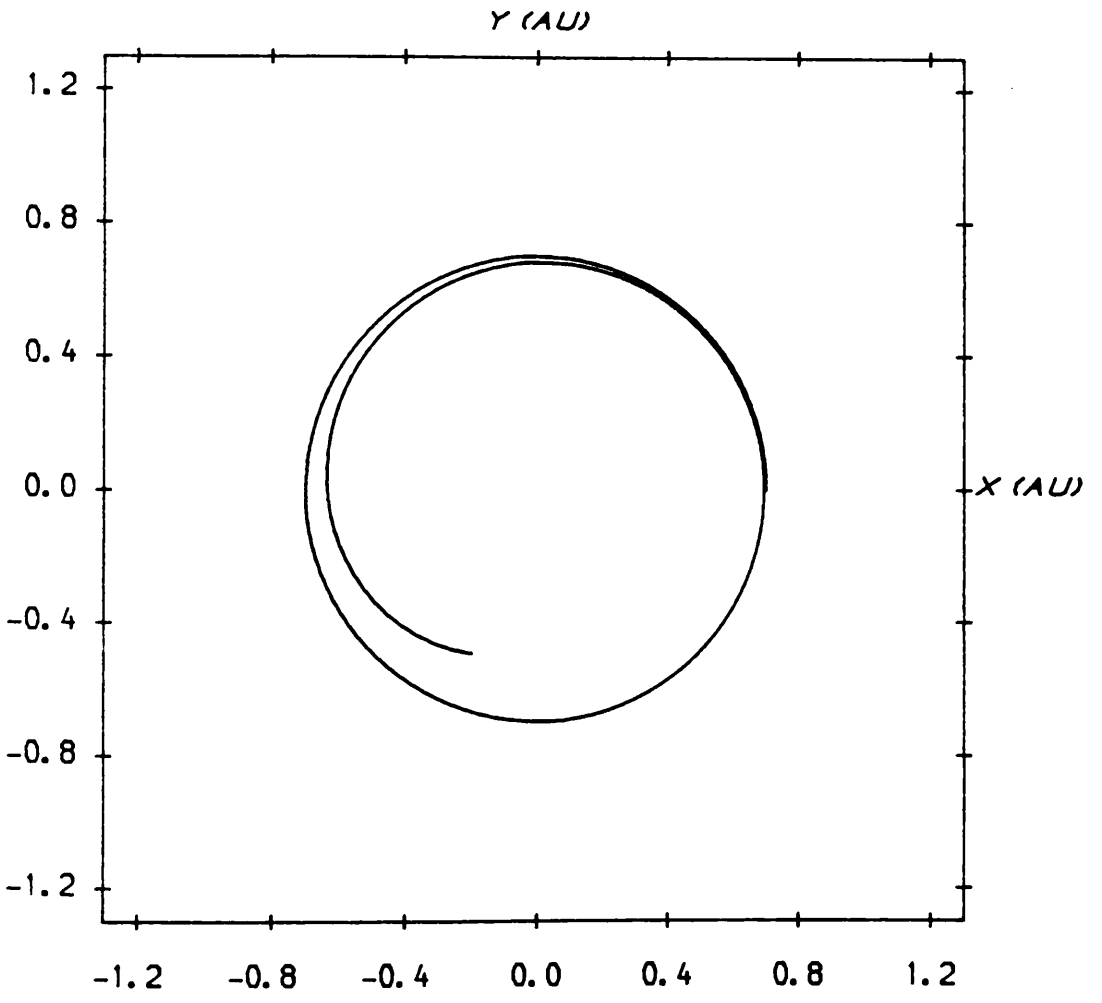
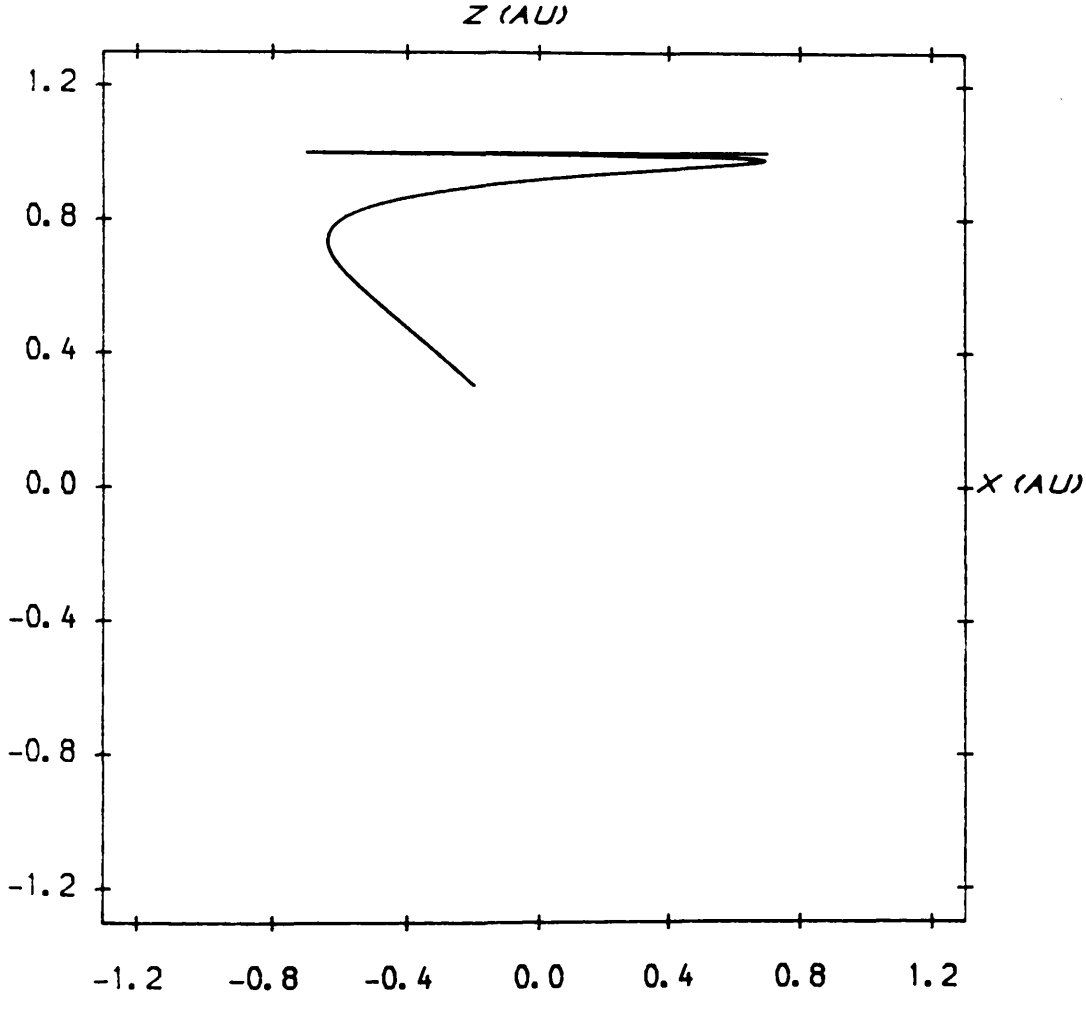


Figure 4.7(a)



Unstable response for a one year halo orbit with ρ=0.7, z=1.0 and injection errors of $\varepsilon_o=n_o=1\times10^{-4}$; (a) x-y projection (b) x-z projection.

Figure 4.7(b)



monitoring platform) then the reduced two-dimensional system may be considered. It will be shown that the two-dimensional system is in principle controllable using feedback to the sail pitch. However, simulations show that the damping timescale is too long for practical purposes. Therefore, a general feedback scheme, with variable sail loading is developed. This scheme is found to produce a suitable control.

4.7.1 Controllability by Sail Pitch

The first type of control to be investigated is a proportional and derivative feedback to the inertial sail pitch angle γ. However, it must firstly be established that the system is in fact controllable using the sail pitch alone.

If the vector dynamical equation is linearised by allowing first order changes in the sail position $\mathbf{r}_o \rightarrow \mathbf{r}_o + \delta$ and also the sail attitude $\mathbf{n}_o \rightarrow \mathbf{n}_o + \delta\mathbf{n}$ then a modified variational equation is obtained, viz

$$\frac{d^2\delta}{dt^2} + \mathbf{M}_1 \frac{d\delta}{dt} + (\mathbf{M} - \mathbf{N})\,\delta = \mathbf{K}\,\delta\mathbf{n} \quad , \quad \mathbf{K} = \left.\frac{\partial \mathbf{a}}{\partial \mathbf{n}}\right|_{\mathbf{r}=\mathbf{r}_o,\ \mathbf{n}=\mathbf{n}_o} \tag{4.32}$$

where the matrix $\mathbf{K}$ gives the first order variation in the solar radiation pressure acceleration with changes in sail attitude. The azimuthal coordinate may again be eliminated using equation (4.20). The variable transformations defined by equations (4.21) are then used to reduce the modified variational equation to the variables $\delta' = (\xi', \eta')$. The variational equation then becomes

$$\frac{d^2\delta'}{dt^2} + \mathbf{L}\,\delta' = \mathbf{K}\,\delta\gamma \quad , \quad \mathbf{L} = \begin{bmatrix} L^*_{11} & \vdots & L_{13} \\ \cdots & \cdots & \cdots \\ L_{31} & \vdots & L_{33} \end{bmatrix} \quad , \quad \mathbf{K} = \begin{bmatrix} \frac{\partial a_\rho}{\partial \gamma} \\ \\ \frac{\partial a_z}{\partial \gamma} \end{bmatrix} \tag{4.33}$$

where the first order attitude change $\delta\mathbf{n}$ now becomes the change in inertial sail pitch angle $\delta\gamma$. The radiation pressure acceleration partial derivatives given by the components of $\mathbf{K}=(K_1,K_2)^T$ are

$$K_1 = \frac{1}{2}\frac{\beta z}{r^4}(\rho\cos\gamma + z\sin\gamma)\left\{3\cos 2\gamma\left\{1 - \frac{\rho}{z}\tan 2\gamma\right\} + 1\right\} \qquad (4.34a)$$

$$K_2 = \frac{1}{2}\frac{\beta \rho}{r^4}(\rho\cos\gamma + z\sin\gamma)\left\{3\cos 2\gamma\left\{1 - \frac{z}{\rho}\tan 2\gamma\right\} - 1\right\} \qquad (4.34b)$$

The system is now reduced further by transforming to a set of four first order equations in the state variable $\mathbf{x}=(\delta',d\delta'/dt)$. The variational equation may then be written in standard form as

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}^*\mathbf{x} + \mathbf{K}^*\delta\gamma \quad , \quad \mathbf{L}^* = \left\{\begin{array}{c|c} \mathbf{0} & \mathbf{I} \\ \hline -\mathbf{L} & \mathbf{0} \end{array}\right\} \quad , \quad \mathbf{K}^* = \left\{\begin{array}{c} \mathbf{0} \\ \hline \mathbf{K} \end{array}\right\} \qquad (4.35)$$

To determine the controllability of system the rank of the 4x4 controllability matrix $\mathbf{C}=(\mathbf{K}^*,\mathbf{L}^*\mathbf{K}^*,\mathbf{L}^{*2}\mathbf{K}^*,\mathbf{L}^{*3}\mathbf{K}^*)$, formed from the system matrix $\mathbf{L}^*$ and the input distribution matrix $\mathbf{K}^*$, is calculated. For the system to be fully controllable the matrix $\mathbf{C}$ must have full rank, (see appendix C). Evaluating the controllability matrix it is found that

$$\mathbf{C} = \left\{\begin{array}{cccc} 0 & K_1 & 0 & -L^*_{11}K_1 - L_{13}K_2 \\ 0 & K_2 & 0 & -L_{31}K_1 - L_{33}K_2 \\ K_1 & 0 & -L^*_{11}K_1 - L_{13}K_2 & 0 \\ K_2 & 0 & -L_{31}K_1 - L_{33}K_2 & 0 \end{array}\right\} \qquad (4.36)$$

It is clear then that $r(\mathbf{C})=4$ if $K_{1,2}\neq 0$ (ie. $\rho\cos\gamma+z\sin\gamma\neq 0$). However, $\mathbf{r}.\mathbf{n}=\rho\cos\gamma+z\sin\gamma$ is always non-zero for the Keplerian synchronous and optimal halo orbit modes and is non-zero for the Earth synchronous

mode, provided that the orbit is within the allowed region of the (ρ,z) plane. There are also several other discrete uncontrollable cases found by setting the terms in brackets in equations (4.34) to zero or equating $K_1=K_2$. Finally, it is found that for the full three-dimensional variational equation the system is uncontrollable using sail pitch alone. That is, although equation (4.19b) shows a coupling between the azimuthal and radial motion it is not sufficient to give full rank to the controllability matrix.

4.7.2 Control by Variable Sail Pitch

Now that the controllability of the system has been established a closed loop feedback scheme to the sail pitch will be investigated. The sail pitch will be related to the state variables through a general feedback expression

$$\varepsilon\gamma = \sum_{j=1}^{4} g_j x_j \quad , \quad \mathbf{x} = (x_j) \quad (j=1,4) \tag{4.37}$$

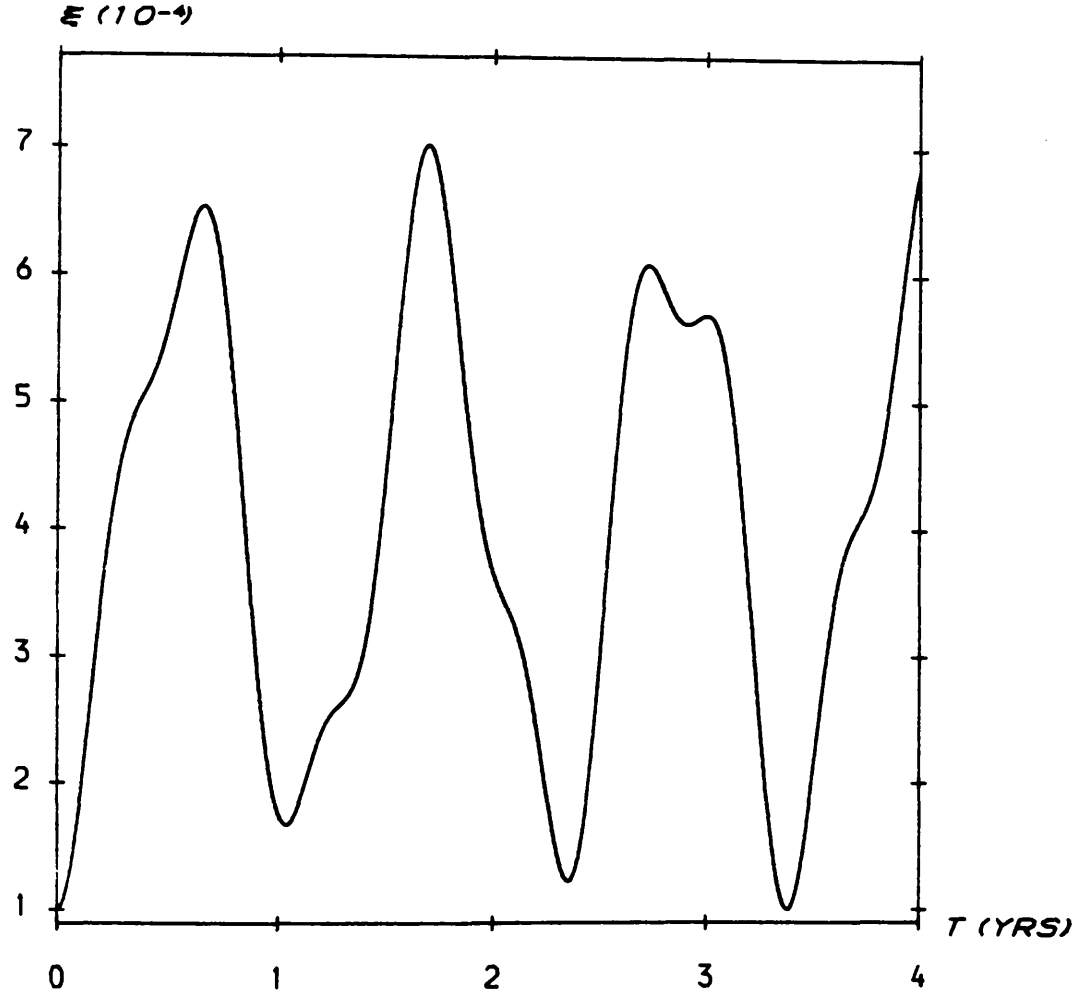
where x_j $(j=1,4)$ are the components of the state vector. The four feedback gains g_j $(j=1,4)$ are chosen to ensure all four of the eigenvalues of the system are in the left hand complex plane so that the system has asymptotic stability. Since the reference state of the system is $\mathbf{x}=0$ the controller is a state regulator. Equation (4.37) can be substituted into equation (4.33) and a new characteristic polynomial obtained. The Routh-Hurwitz criterion (see for example Barnett and Cameron (1985)) then defines limits to the range of values the gains may take so that all eigenvalues are in the left hand complex plane. Since there are four gains it is not useful to design the control scheme by plotting the positions of the closed loop eigenvalues in the

complex plane, (ie. root locus technique, see for example Barnett and Cameron (1985)). Instead, a simulation using a numerical integration of the full non-linear dynamical equations was used. The Routh-Hurwitz conditions were implicit in the simulation to ensure the values of the gains were within the correct bounds.

The performance of the variable sail pitch control is shown for a low displacement halo orbit, Figures 4.8 and 4.9. The free, open loop response for a stable Earth synchronous halo orbit is shown in Figure 4.8. It can be seen that there is a combination of long and short period responses, corresponding to the two sets of eigenvalues of the stable system. The azimuthal drift corresponding to equation (4.31) is also seen. The closed loop response for the same halo orbit with feedback to the sail pitch is shown in Figure 4.9. The gains were chosen to attempt to minimise the damping time while avoiding excessive overshoot. It can be seen that the damping timescale is extremely long. This is due to the coupling between the magnitude and direction of the solar radiation pressure force. It will later be shown in section 4.7.4 that if first order variations in the sail area are allowed to overcome this coupling, much shorter damping timescales are possible. The azimuthal drift is again present, but with the superimposed oscillations being damped out.

It can also be seen from Figure 4.9 that the injection errors damp out to non-zero values. This is due to the variable transformations, defined by equations (4.21), used to obtain the reduced variational equation. That is, the control scheme ensures that $(\xi',\eta')\rightarrow 0$ leaving the constant terms in equations (4.21) as residual errors. Physically these residual errors are due to the sail having excess orbital angular momentum due to the initial error ξ_0 along the

Figure 4.8(a)



Open loop response for a one year halo orbit with ρ=0.6, z=0.2 and injection errors of $\xi_o=\eta_o=\Psi_o=1x10^{-4}$; (a) ξ ($x10^{-4}$) response (b) η ($x10^{-4}$) response (c) Ψ response.

Figure 4.8(b)

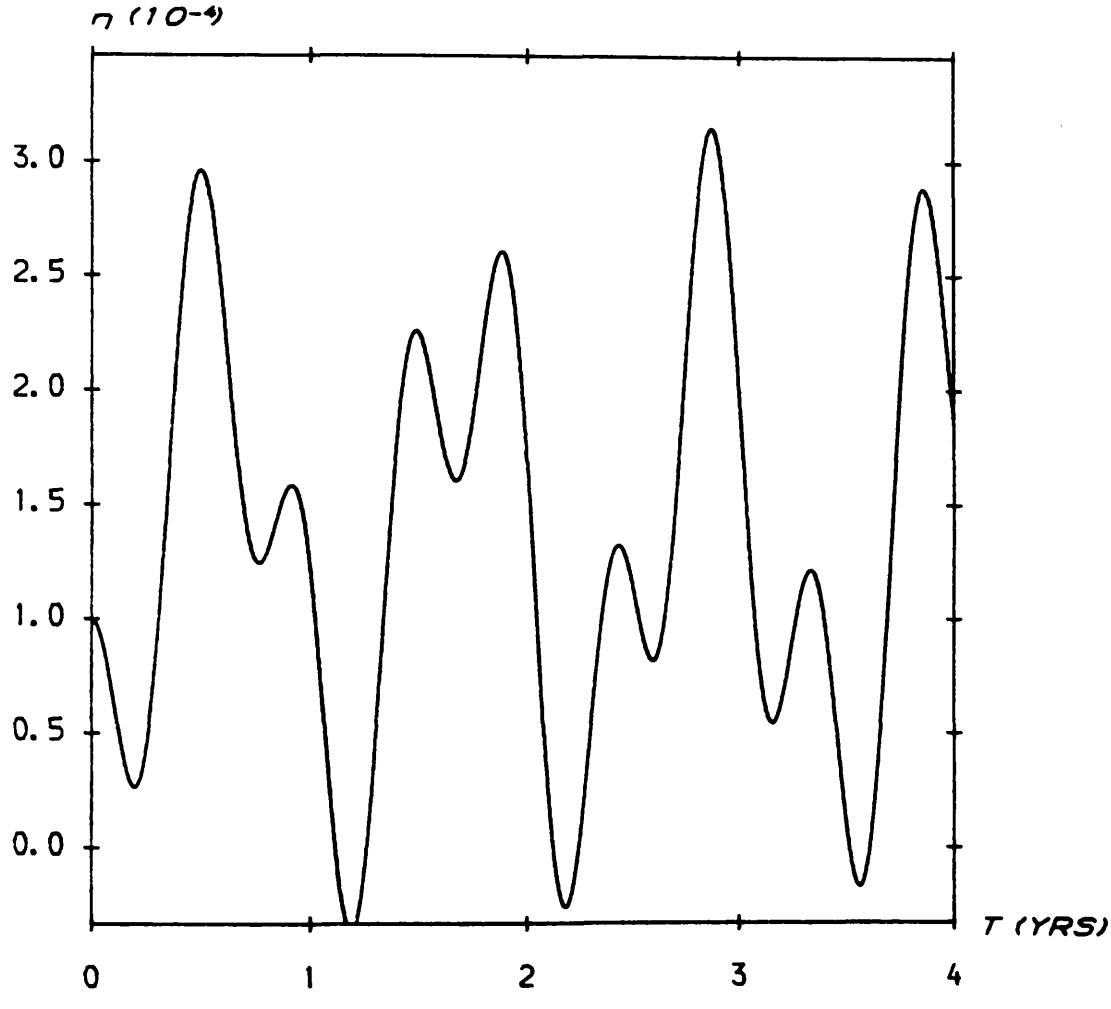


Figure 4.8(c)

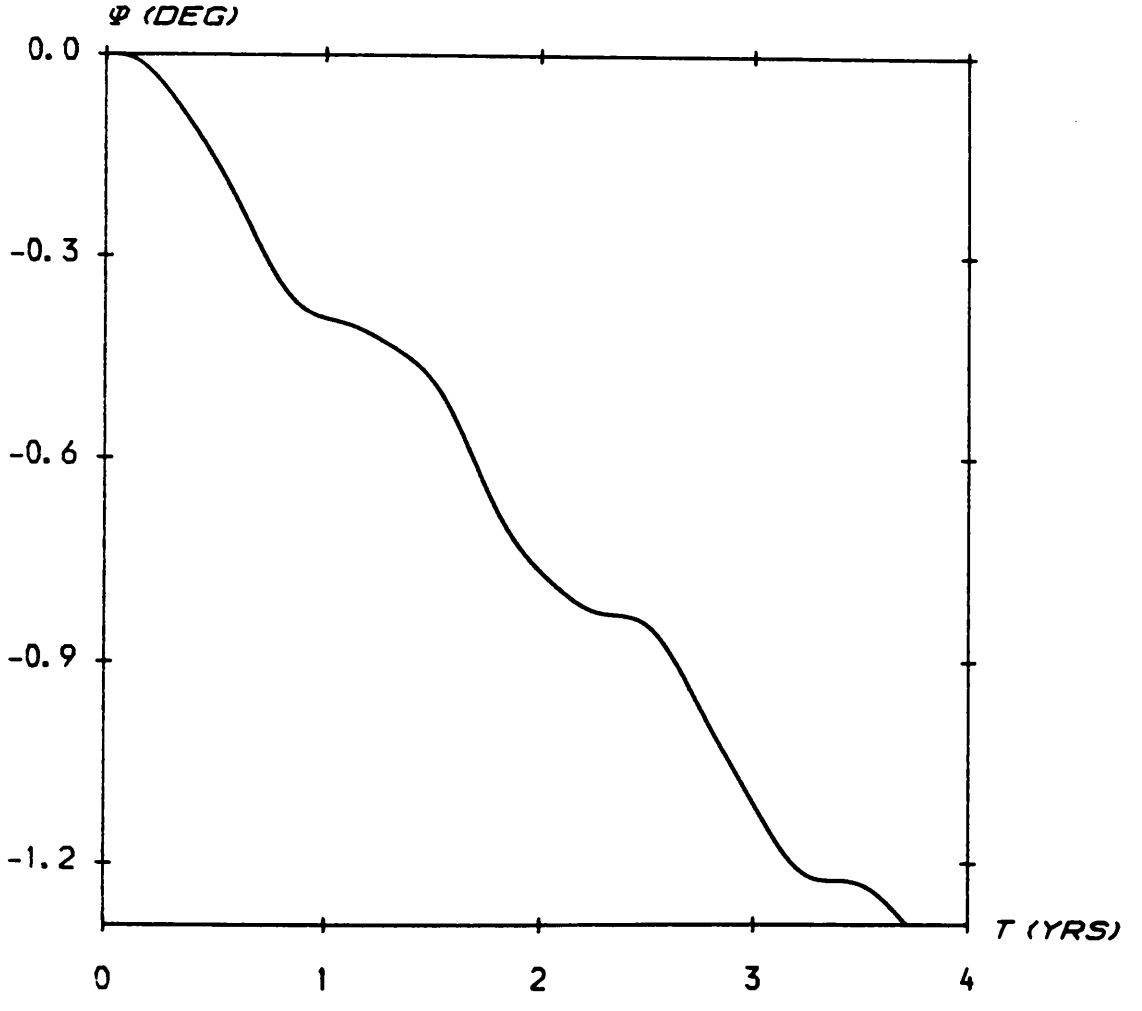
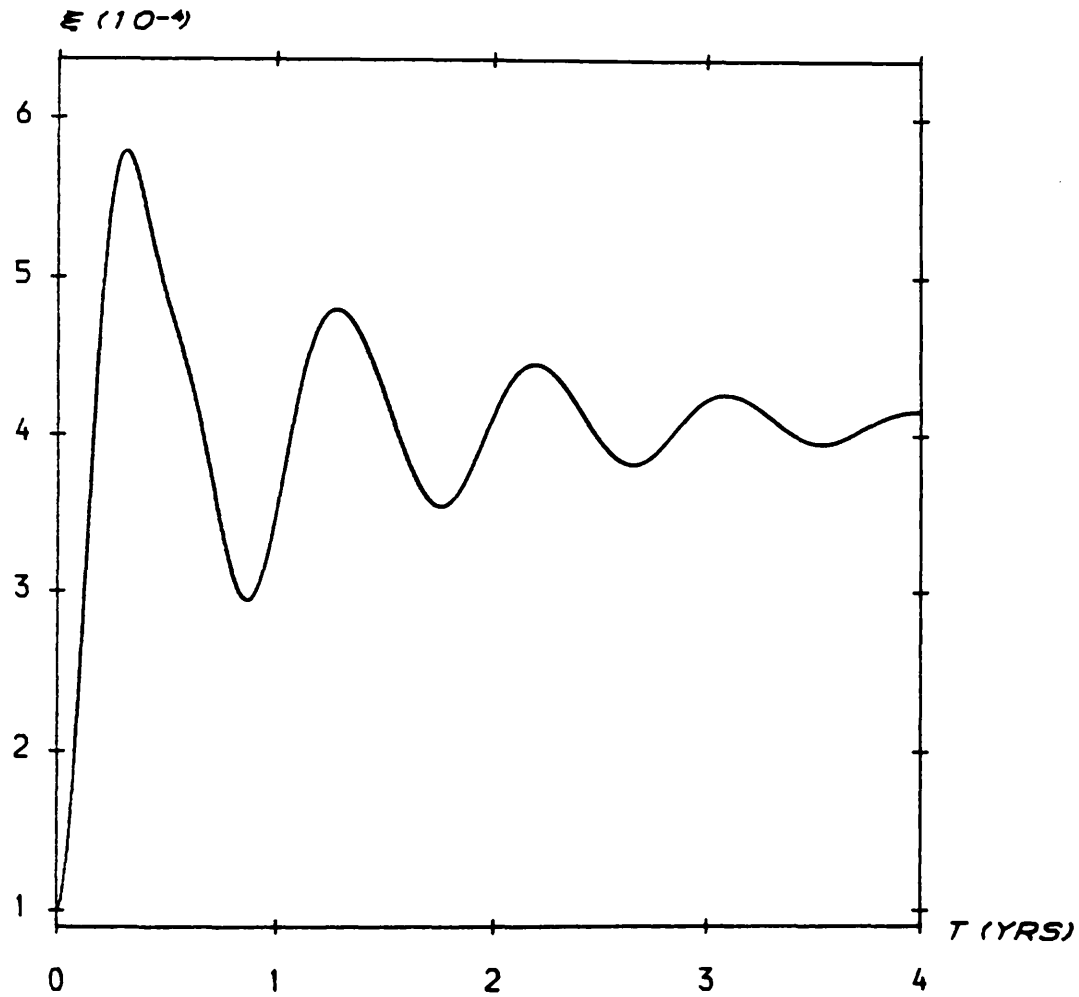


Figure 4.9(a)



Closed loop response for a one year halo orbit with ρ=0.6, z=0.2, injection errors of $\varepsilon_o=\eta_o=\Psi_o=1\times10^{-4}$ and a variable sail pitch angle control. The feedback gains are g_1=-1.5, g_2=-1.2, g_3=-2.0 and g_4=-1.5; (a) ε ($\times10^{-4}$) response (b) η ($\times10^{-4}$) response (c) Ψ response (d) sail inertial pitch angle control ($\times10^{-2}$).

Figure 4.9(b)

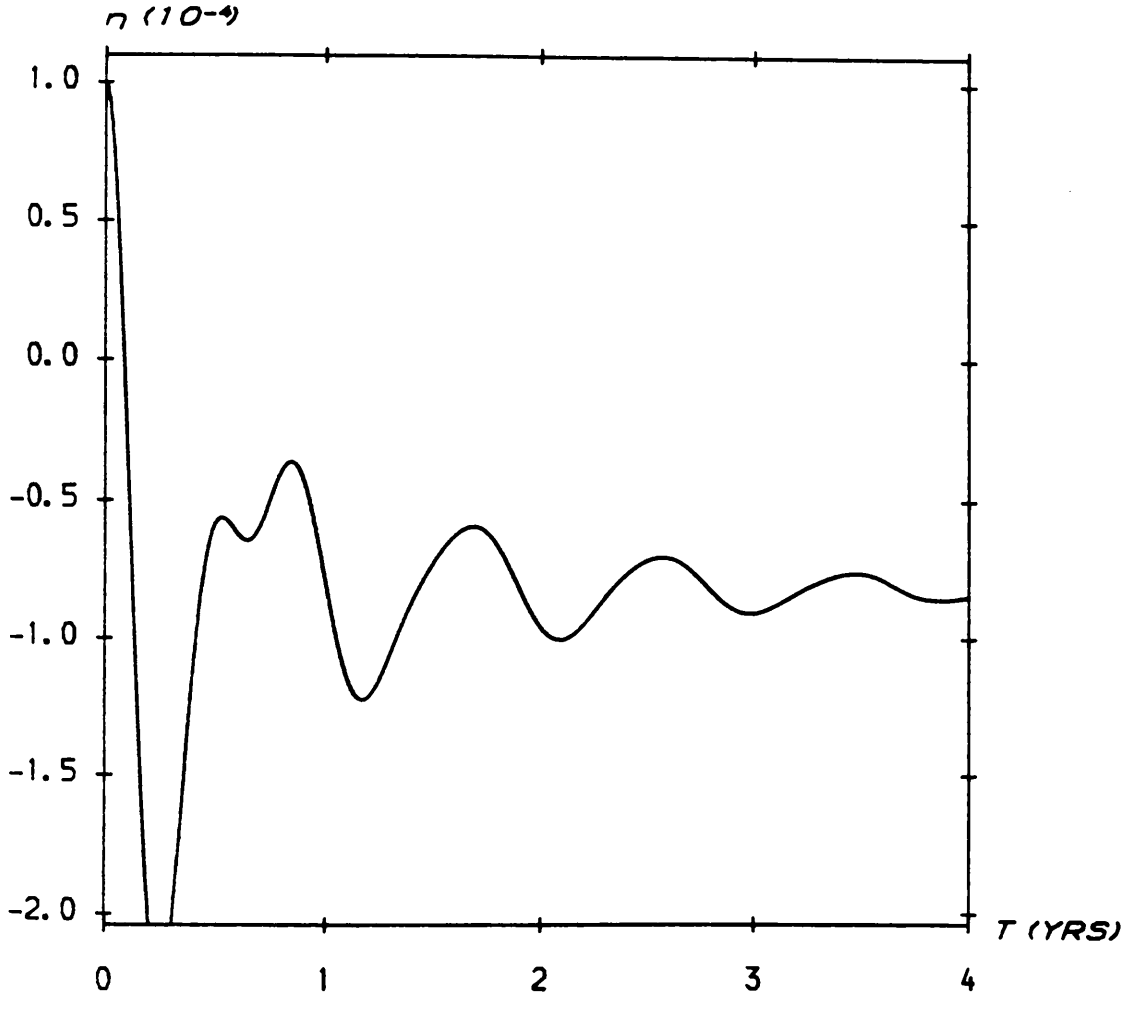


Figure 4.9(c)

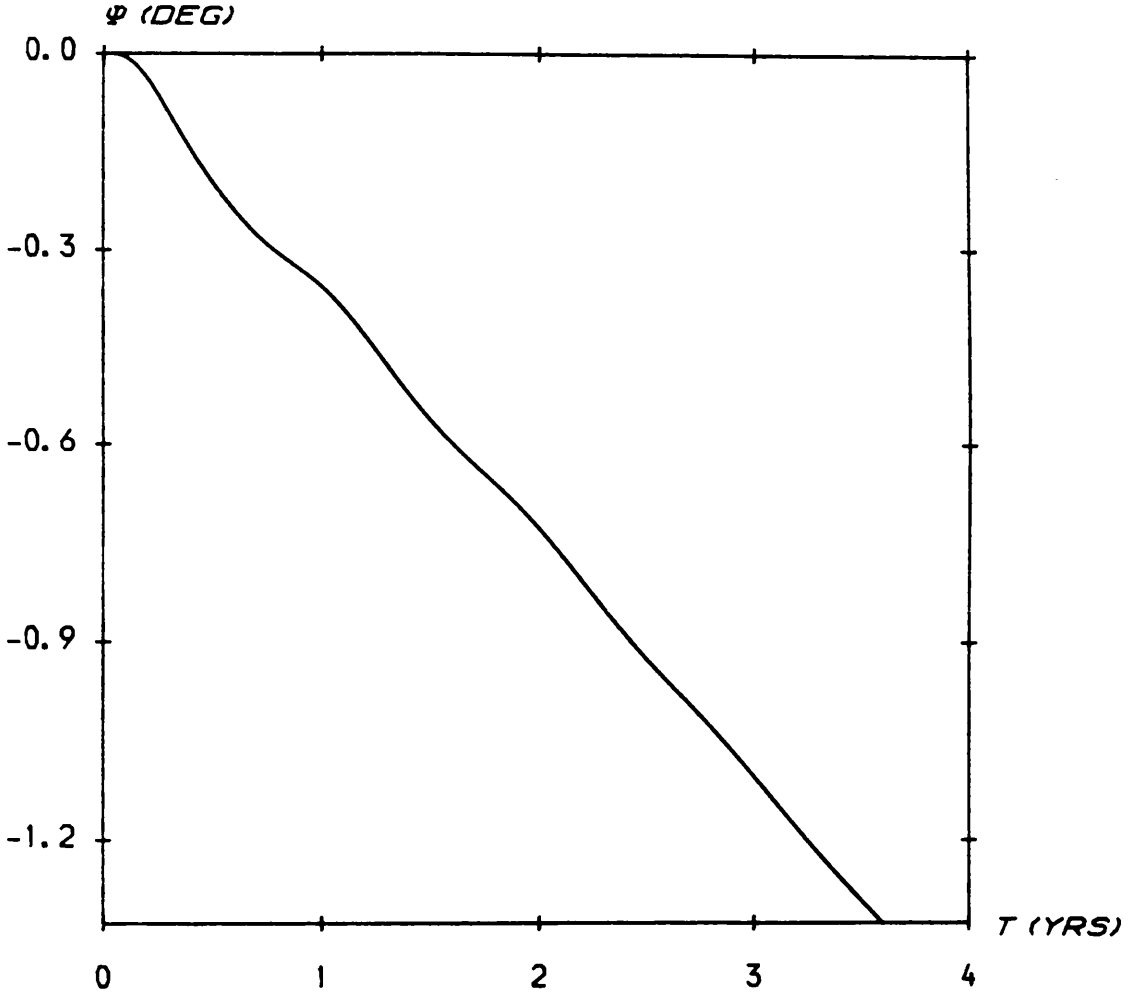
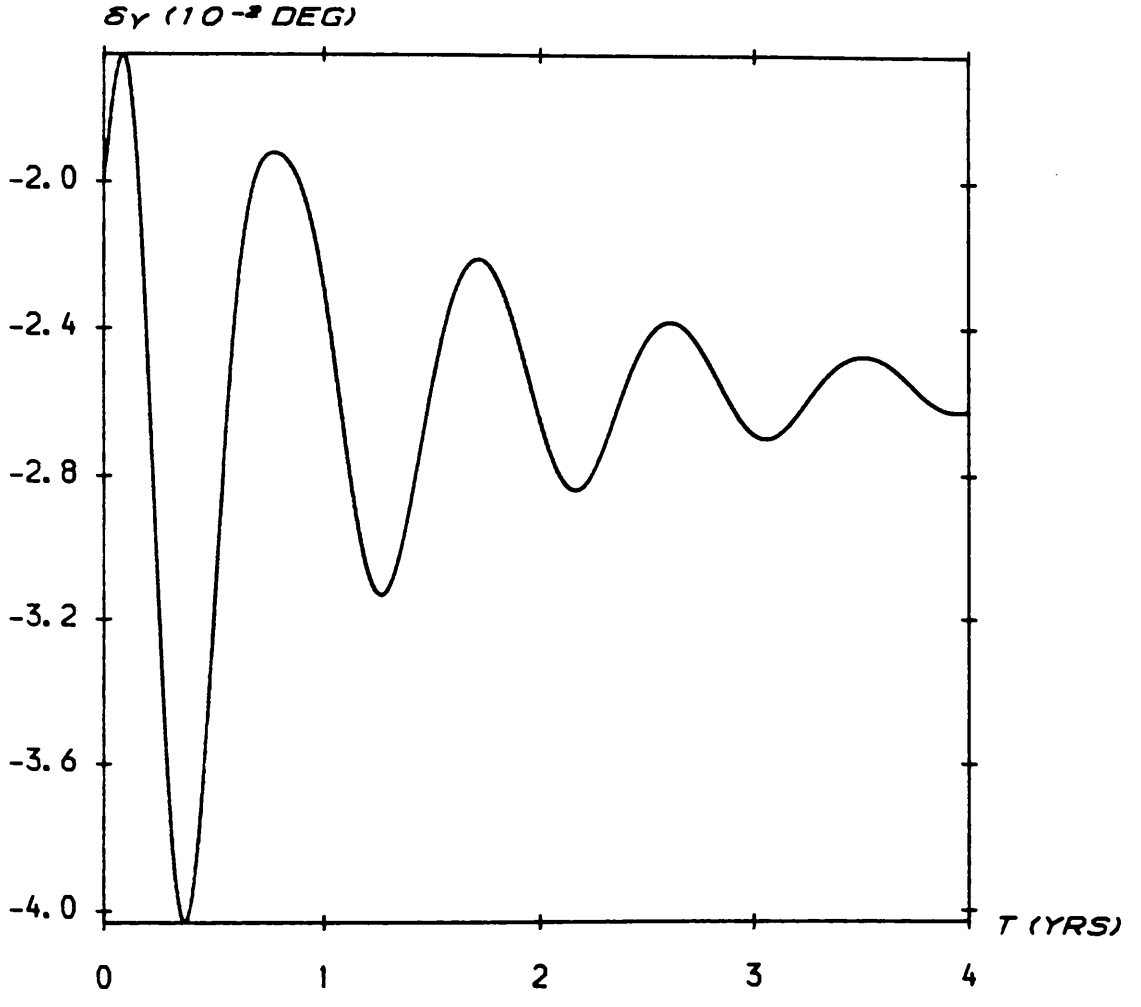


Figure 4.9(d)



ρ-axis. This excess momentum can in fact be removed through an open loop control manoeuvre. The manoeuvre uses a sail rotation about the yaw axis to generate a control acceleration in the azimuthal direction. This technique will be discussed in detail in section 5.8.3.

4.7.3 Control by Fixed Sun-Sail Pitch

An extremely simple closed loop scheme is now investigated which ensures Lyapunov stability for all of the halo orbit families. The scheme does not damp out injection errors, but is used only to stabilize unstable orbits, such as Earth synchronous halo orbits high above the ecliptic plane. The scheme requires that the Sun-sail pitch angle α remains fixed. Therefore, as the sail moves from the nominal halo orbit the inertial sail pitch angle γ must vary.

In terms of the fixed Sun-sail pitch angle α the components of the radiation pressure acceleration (a_ρ, a_z) may be written as

$$a_\rho = \frac{\beta}{r^2} \cos^2\alpha \left\{ \cos\alpha \frac{\rho}{r} - \sin\alpha \frac{z}{r} \right\} \tag{4.38a}$$

$$a_z = \frac{\beta}{r^2} \cos^2\alpha \left\{ \sin\alpha \frac{\rho}{r} + \cos\alpha \frac{z}{r} \right\} \tag{4.38b}$$

These expressions may then be used to evaluate the new radiation gradient tensor, in a similar procedure to section 4.6. Therefore, a new system matrix may be formed and its stability characteristics examined. It is found that the reduced variational equation becomes

$$\frac{d^2}{dt^2} \begin{bmatrix} \xi' \\ \eta' \end{bmatrix} + \begin{bmatrix} P^*_{11} & \vdots & P_{13} \\ \cdots & \cdots & \cdots \\ P_{31} & \vdots & P_{33} \end{bmatrix} \begin{bmatrix} \xi' \\ \eta' \end{bmatrix} = 0 \tag{4.39}$$

where the coefficients of the matrix **P** are given by

$$P^*_{11} = 4\Omega^2 - \left\{ \left\{\Omega^2 - \frac{1}{r^3}\right\} + \frac{3\rho^2}{r^5} \right\} - \Gamma_1 \, (\cos\alpha - \Gamma_2\Gamma_4) \qquad (4.40a)$$

$$P_{13} = - \frac{3\rho z}{r^5} + \Gamma_1 \, (\sin\alpha + \Gamma_3\Gamma_4) \qquad (4.40b)$$

$$P_{31} = - \frac{3\rho z}{r^5} - \Gamma_1 \, (\sin\alpha - \Gamma_2\Gamma_5) \qquad (4.40c)$$

$$P_{33} = \left\{\frac{1}{r^3} - \frac{3z^2}{r^5}\right\} - \Gamma_1 \, (\cos\alpha - \Gamma_3\Gamma_5) \qquad (4.40d)$$

The auxiliary coefficients Γ_j (j=1,5) are given by

$$\Gamma_1 = \frac{\beta}{r^3} \cos^2\alpha \qquad (4.41a)$$

$$\Gamma_2 = \frac{3\rho}{r^2} \qquad (4.41b)$$

$$\Gamma_3 = \frac{3z}{r^2} \qquad (4.41c)$$

$$\Gamma_4 = \rho\cos\alpha - z\sin\alpha \qquad (4.41d)$$

$$\Gamma_5 = \rho\sin\alpha + z\cos\alpha \qquad (4.41e)$$

To determine the stability of the system the new characteristic polynomial is obtained and the eigenvalues found, viz

$$\omega^4 + tr(\mathbf{P})\omega^2 + det(\mathbf{P}) = 0 \qquad (4.42)$$

The conditions for purely imaginary eigenvalues and Lyapunov stability are again that $tr(\mathbf{P})>0$ and $det(\mathbf{P})>0$. However, due to the scale invariance of the system, as discussed in section 4.6, only the stability characteristics of the Keplerian synchronous mode need be examined. Substituting for $\Omega=\Omega_*$ it is found that

$$\mathrm{tr}(\mathbf{P}) = \Omega_*^2\left\{2 + \left\{\frac{z}{r}\right\}^2\right\} > 0 \quad , \quad \det(\mathbf{P}) = \Omega_*^4\left\{\frac{\rho}{r}\right\}^2 > 0 \tag{4.43}$$

It is clear then that the Keplerian synchronous mode has Lyapunov stability with the fixed Sun-sail pitch angle control. Therefore, due to the scale invariance, all halo orbit modes become stable. There will of course be an azimuthal drift due to the excess orbital angular momentum of the sail owing to the initial error along the ρ-axis.

The change inertial pitch angle $\delta\gamma$ can be related to a feedback control by evaluating the change in Sun-sail pitch angle with first order changes in ρ and z. The angle α is given by

$$\alpha = \gamma - \tan^{-1}\left\{\frac{z}{\rho}\right\} \tag{4.44}$$

so that $\delta\alpha$ may be formed from $\partial\alpha/\partial\rho$ and $\partial\alpha/\partial z$. The required control to maintain a fixed Sun-sail pitch angle is then $\delta\gamma=-\delta\alpha$, viz

$$\delta\gamma = \frac{1}{\rho}\,\frac{1}{1 + (z/\rho)^2}\left\{\eta - \frac{z}{\rho}\,\xi\right\} \tag{4.45}$$

This control scheme is appealing in practice since no state variable information is required. A Sun tracking sensor would measure the change in Sun-sail pitch angle $\delta\alpha$ which would then be used directly to command the change in sail inertial pitch angle $\delta\gamma$. It is this simplicity of the control scheme which makes it so attractive. A series of open loop commands would be used to remove any initial injection errors. The fixed Sun-sail pitch angle control may then be used to ensure the orbits remain Lyapunov stable.

4.7.4 Control by Variable Sail Attitude and Loading

The variable sail pitch control investigated in section 4.7.2

showed that the coupling between the magnitude of the solar radiation pressure acceleration and the sail attitude lead to long damping timescales. To overcome this coupling first order changes in the sail loading (ie. changes in the sail area) will be allowed. Then, arbitrary control accelerations may be generated and the required variation in the sail attitude and loading obtained.

The full three-dimensional variational equation with a control acceleration Δ may be written as

$$\frac{d^2\delta}{dt^2} + M_1 \frac{d\delta}{dt} + (M - N)\,\delta = \Delta \quad , \quad \Delta = \Lambda_1\delta + \Lambda_2 \frac{d\delta}{dt} \qquad (4.46)$$

where the control acceleration is related to the state variables by the feedback gain matrices $\Lambda_{1,2}$. It would in principle be possible to choose the elements of these matrices so that the off-diagonal terms in the variational equation, which lead to dynamic coupling, were eliminated, (cf. the elements L_{13} and L_{31} in equation (4.22)). However, since these terms are of the same magnitude as the principal diagonal terms the system is not suitable for artificial de-coupling.

The variations in the sail attitude and loading required to generate the control acceleration may be obtained from the expression for the solar radiation pressure acceleration given by equation (4.2). Allowing first order variations in the sail attitude $\delta\mathbf{n}$ and the sail loading $\delta\beta$ it is required that

$$(\beta+\delta\beta)\,\frac{\mu}{|\mathbf{r}|^4}\,(\mathbf{r}.(\mathbf{n}+\delta\mathbf{n}))^2\,(\mathbf{n}+\delta\mathbf{n}) = \beta\,\frac{\mu}{|\mathbf{r}|^4}\,(\mathbf{r}.\mathbf{n})^2\mathbf{n} + \Delta \qquad (4.47)$$

By noting that $|\mathbf{n}+\delta\mathbf{n}|=1$ this vector equation may be solved by taking vector products to obtain the required sail attitude control as

$$\delta \mathbf{n} = \frac{\mathbf{a} + \Delta}{|\mathbf{a} + \Delta|} - \mathbf{n} \tag{4.48}$$

where $\mathbf{a}$ is the nominal solar radiation pressure acceleration given by equation (4.2). This expression can then be related to the inertial sail pitch angle control $\delta\gamma$ and the yaw angle control $\delta\chi$, (rotation about the z-axis). The loading control is also obtained as

$$\delta\beta = \mu^{-1}|\mathbf{r}|^4 \frac{|\mathbf{a} + \Delta|^3}{(\mathbf{r}.(\mathbf{a} + \Delta))^2} - \beta \tag{4.49}$$

The required sail pitch and yaw controls along with the required loading control can therefore be obtained as a function of the control acceleration.

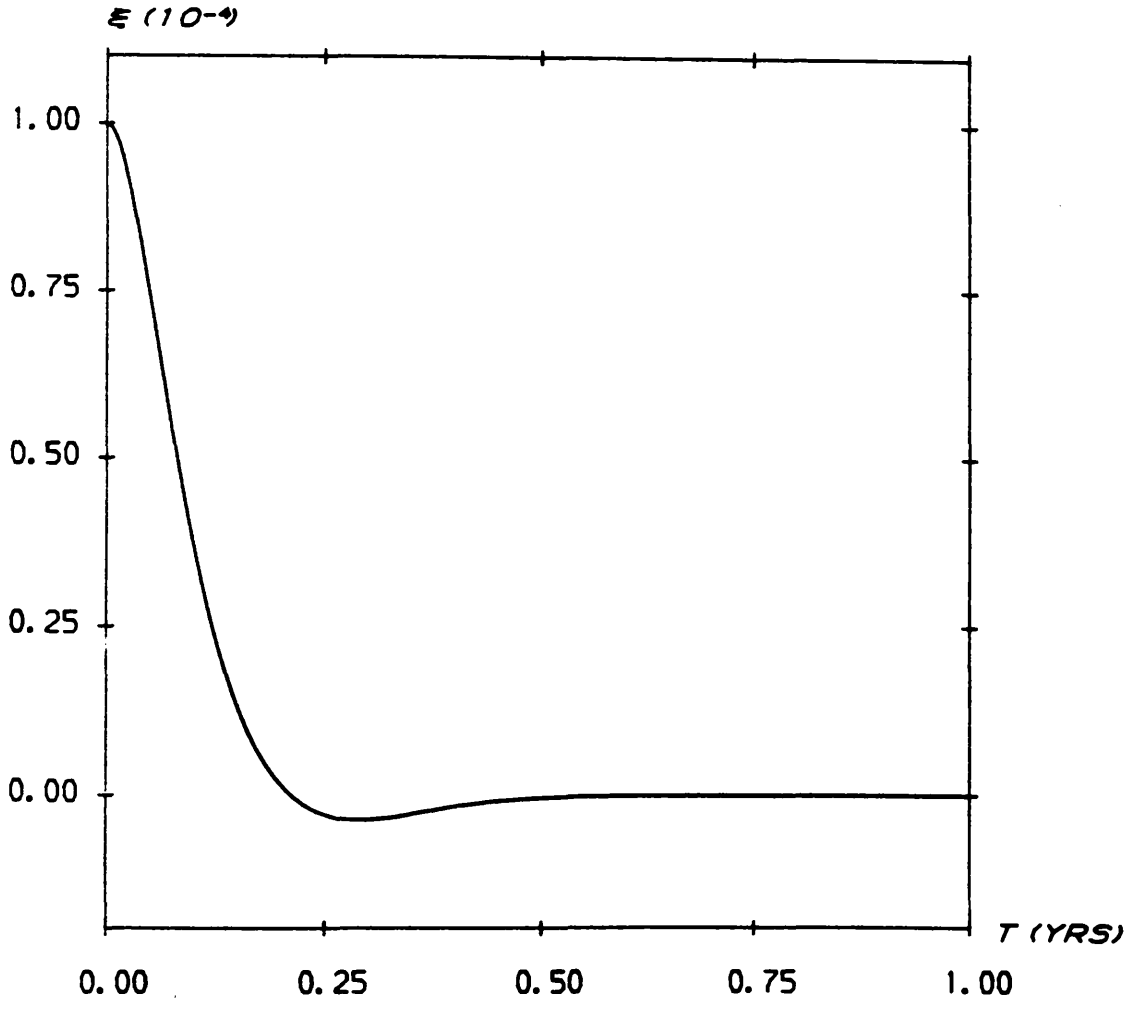
The 3x3 gain matrices will be chosen to be diagonal so that each component of the control acceleration is a simple function of the position and velocity along that axis, viz

$$\Lambda_1 = (\delta_{ij} g_{1j}) \quad , \quad \Lambda_2 = (\delta_{ij} g_{2j}) \quad , \quad (i,j=1,3) \tag{4.50}$$

where δ_{ij} is the kronecker delta function. Using this form for the gain matrices a simulation of the control scheme is used to choose the individual gains to achieve a suitable time response, Figure 4.10. It can be seen that the injection errors are damped out in approximately 60° of orbit, with a maximum loading variation of less than 10^{-2}. Shorter damping timescales are possible, but at the expense of a larger overshoot and larger variations in the sail loading.

It has been shown then that the unstable families of halo orbit may be well controlled by using a combination of sail attitude and loading variations to give adequate damping timescales. However, the simple fixed Sun-sail pitch angle control is appealing in that no state

Figure 4.10(a)



Closed loop response for a one year halo orbit with variable sail pitch, yaw and loading control. The orbit parameters are ρ=0.5, z=1.5 with injection errors of $\varepsilon_o=\eta_o=\Psi_o=1\times10^{-4}$. The feedback gains are g_{1j}=-10 and g_{2j}=-5, (j=1,3); (a) ε ($\times10^{-4}$) response (b) η ($\times10^{-4}$) response (c) Ψ ($\times10^{-3}$) response (d) sail inertial pitch angle control (e) sail yaw angle control (f) sail loading control ($\times10^{-2}$).

Figure 4.10(b)

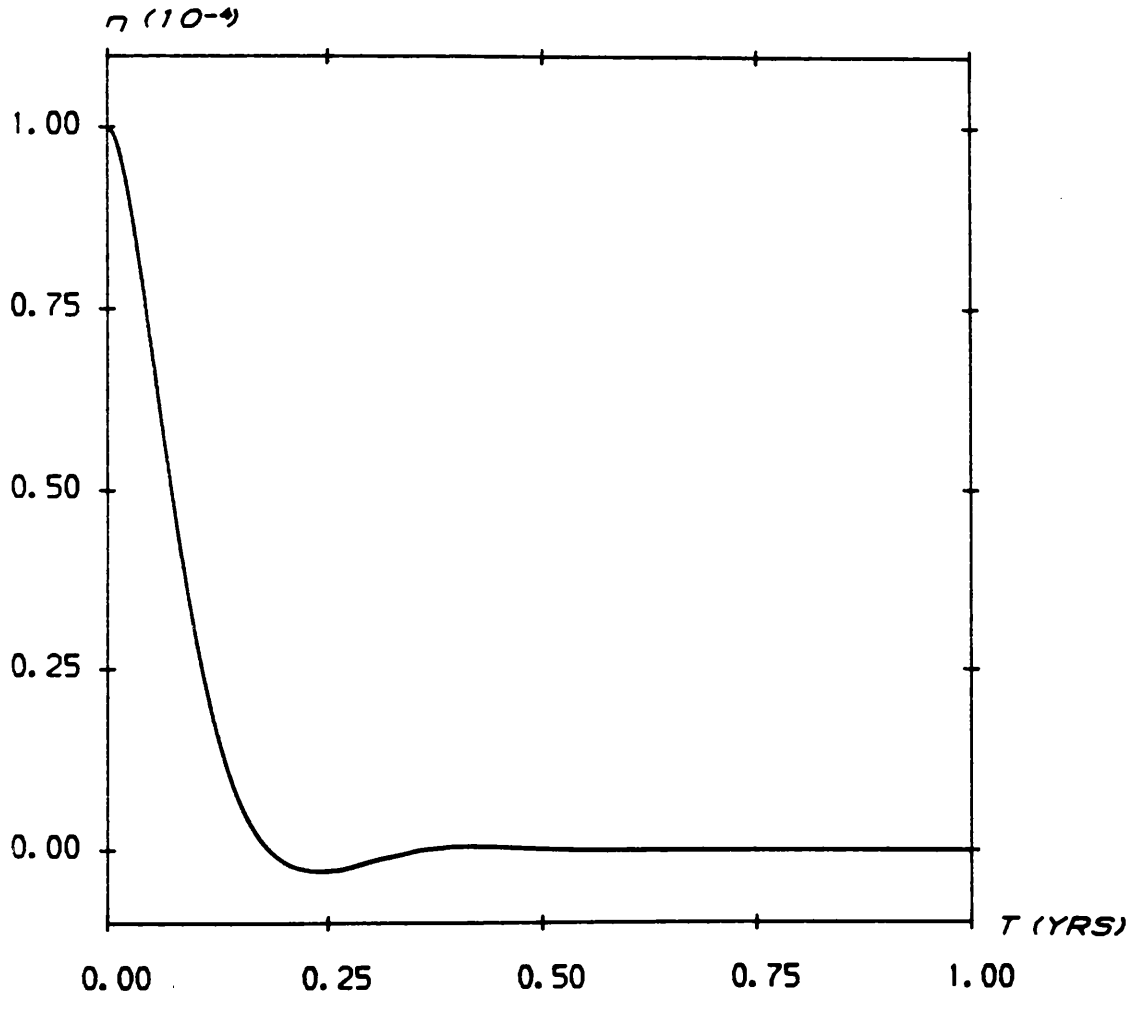


Figure 4.10(c)

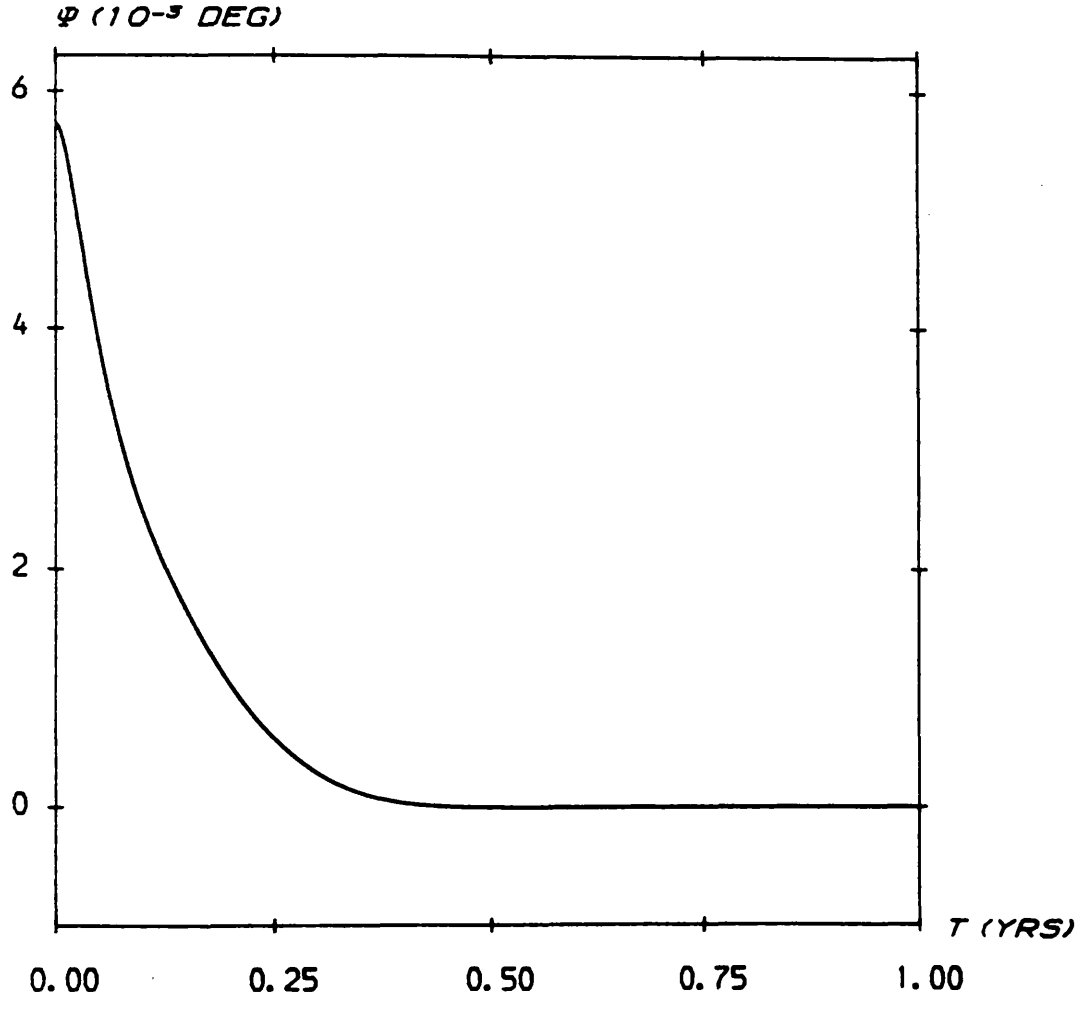


Figure 4.10(d)

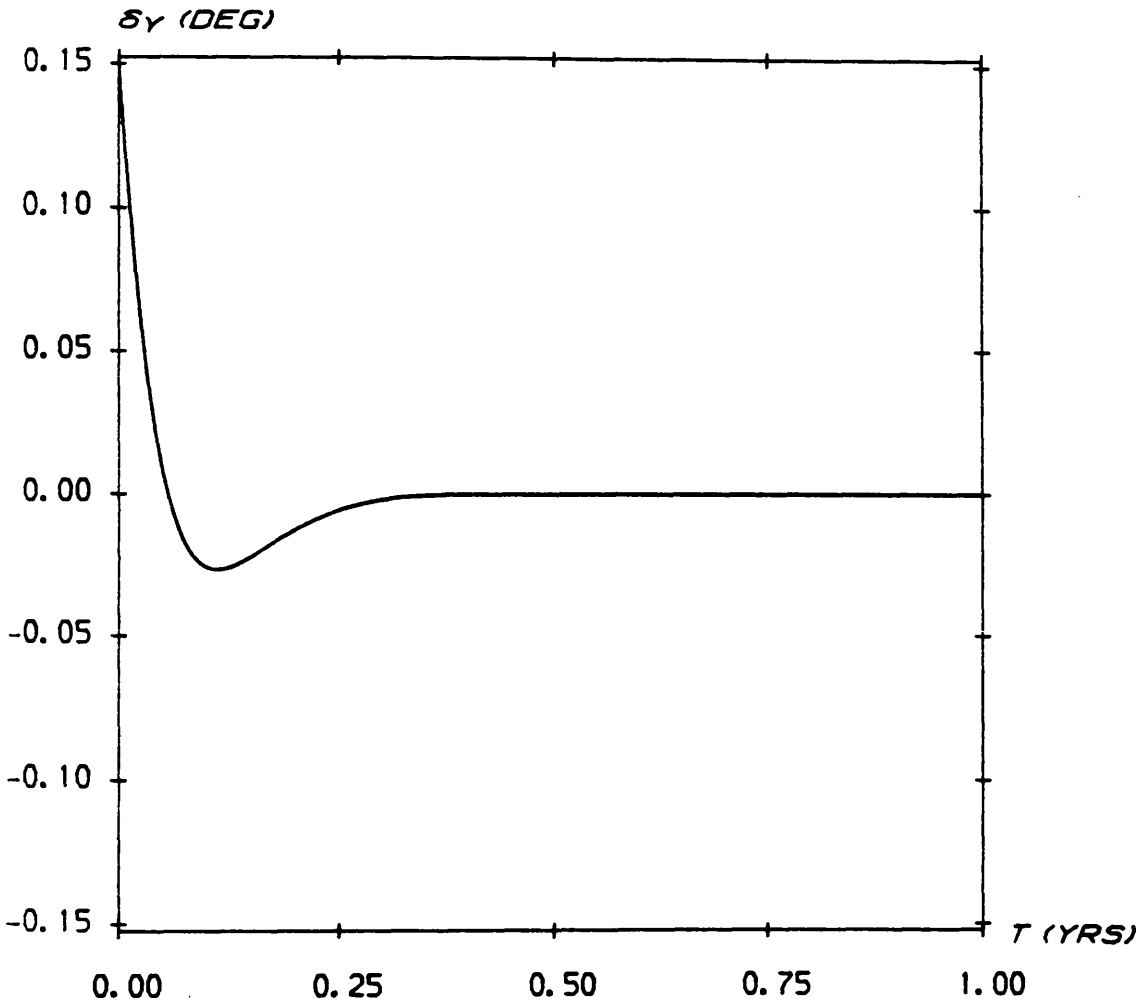


Figure 4.10(e)

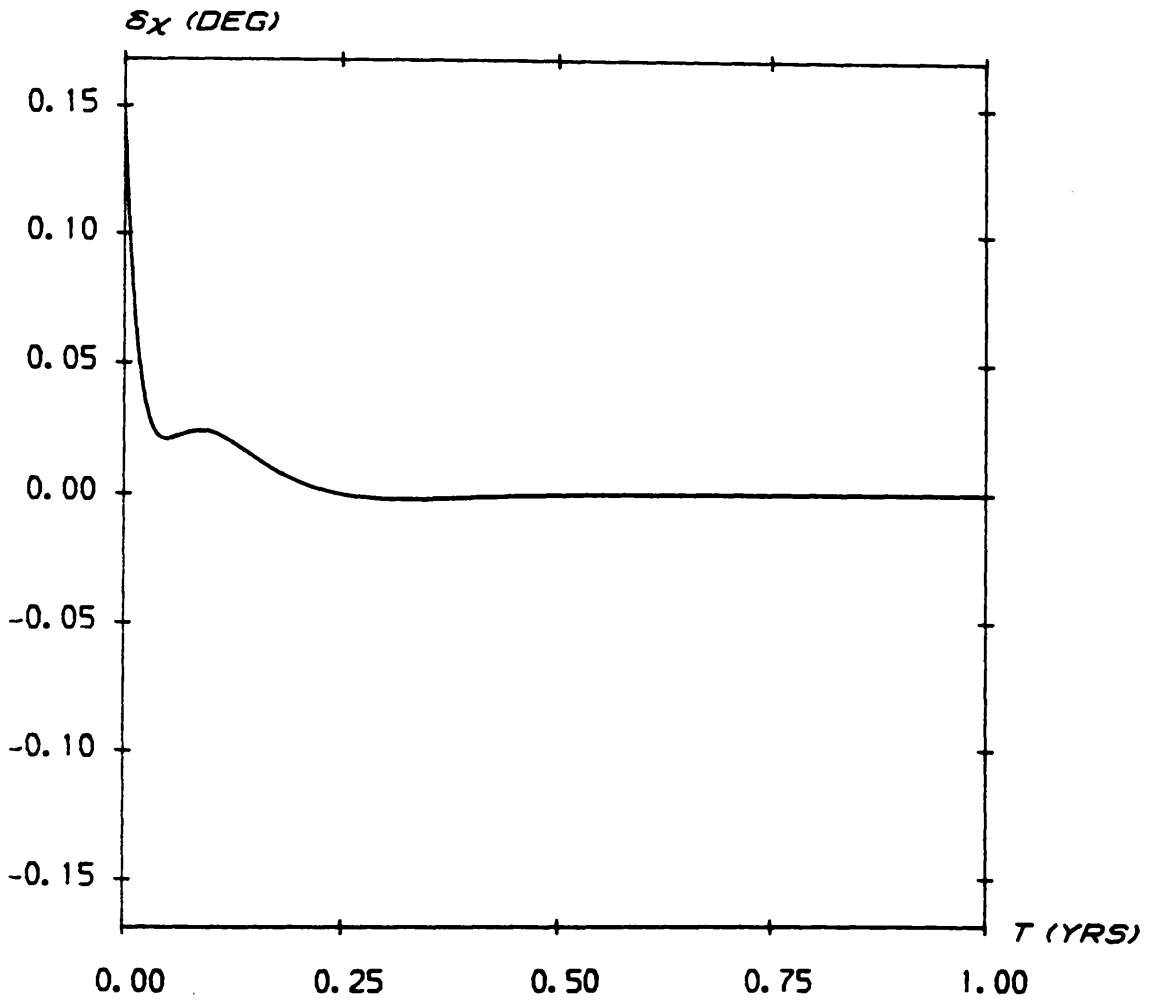
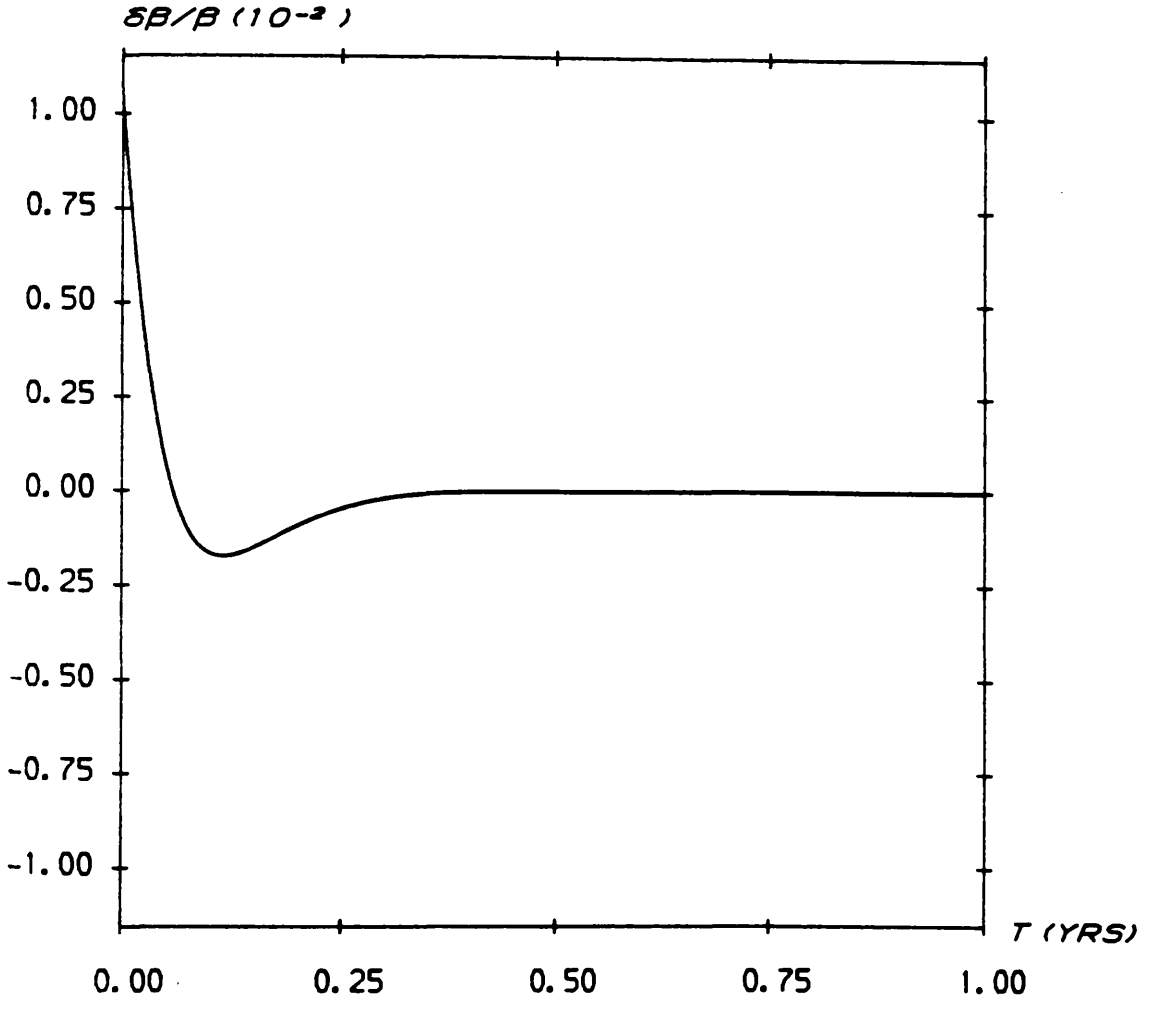


Figure 4.10(f)


δB/B (10⁻²)
1.00
0.75
0.50
0.25
0.00
-0.25
-0.50
-0.75
-1.00
T (YRS)
0.00
0.25
0.50
0.75
1.00

variable information is required onboard for the control system. All that is required is a Sun tracking sensor to generate inertial sail pitch angle commands. Open loop commands may then be used to limit the sail motion and to remove the initial injection errors.

4.8 Patched Heliocentric Halo Orbits

In this final section the possibility of patching halo orbits together will be investigated. By a simple, assumed instantaneous, switching operation on the sail attitude at discrete points along the trajectory complex and elaborate new trajectories may be generated. In order that the operation involves only a change in the sail attitude several boundary conditions must be satisfied. Denoting the initial halo (I) by the subscript 1 and the final halo (II) by the subscript 2 these conditions may be written as

(i) $r_1=r_2$; Intersection of the halo orbits.

(ii) $\mathbf{v}_1=\mathbf{v}_2$; No velocity impulse required at the switching point.

(iii) $E_1=E_2$; Sail energy is continuous across the switching operation.

(iv) $\beta_1=\beta_2$; Sail loading is continuous across the switching operation.

4.8.1 Halo-Halo Transfer

Since the switching of the sail attitude is assumed to take place instantaneously condition (i) implies that the sail gravitational potential is continuous across the operation. Furthermore, since the radiation field is non-conservative with respect to sail rotations (ie. the sail does no work against the field by being re-oriented) it is only required that the sail kinetic energy is continuous across the switching operation. Condition (iii) therefore reduces to $|\mathbf{v}_1|=|\mathbf{v}_2|$,

(ie. $\rho_1\Omega_1=\rho_2\Omega_2$). The period of halo II is then determined explicitly by $\Omega_2=\Omega_1(\rho_1/\rho_2)$.

Using this relation and equating the sail loadings to satisfy condition (iv) it is found that the required conditions are $\rho_1=\rho_2$ and so $z_1=z_2$. Therefore, the spacecraft must transfer to an identical halo orbit II, but with an axis perpendicular to halo I. Using this procedure four halo orbits may be patched together, with the sail attitude being switched at each of the intersections to form an elaborate new trajectory in which the sail orbits over the surface of a cube, Figure 4.11. It should be noted that halo III is retrograde with respect to halo I. Furthermore, the entire patched trajectory is symmetric to rotations and so the cube may be oriented in any way with respect to the ecliptic plane.

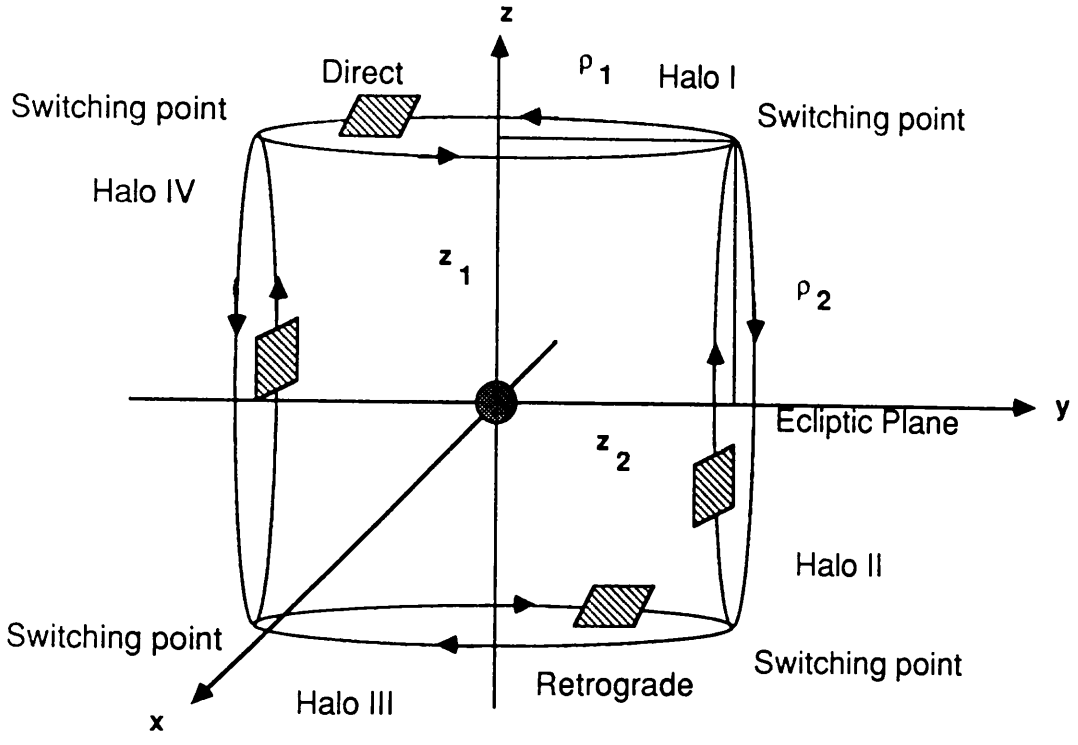
4.8.2 Keplerian Transfer

Transfer to and from other halo orbits is not the only means of patching trajectories. The sail attitude may be switched into a null orientation with $\alpha=\pi/2$, so that the sail will be transferred onto a Keplerian ellipse, Figure 4.12. This transfer may take place at the perihelion or aphelion of the ellipse. The perihelion and aphelion velocities to be matched are given by, (see for example Roy (1982))

$$v_p = \left\{ \frac{1}{a} \left\{ \frac{1+e}{1-e} \right\} \right\}^{1/2}, \quad v_a = \left\{ \frac{1}{a} \left\{ \frac{1-e}{1+e} \right\} \right\}^{1/2} \tag{4.51}$$

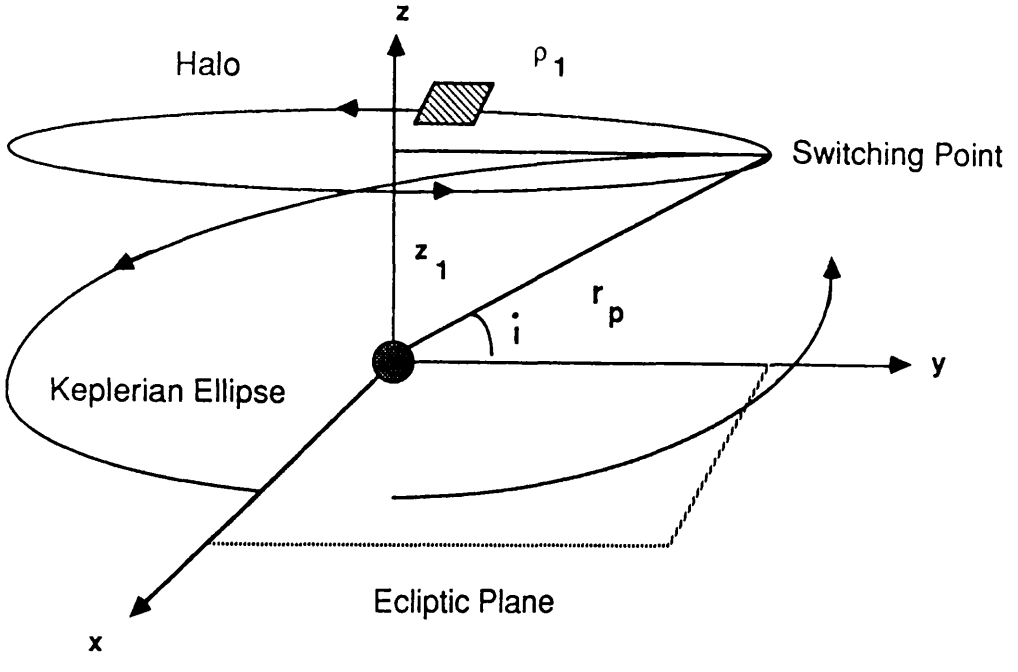
where a and e are the semi-major axis and eccentricity of the ellipse. If a transfer at the perihelion point is considered the required halo amplitude is given by $\rho_1=r_p\cos(i)$, where $r_p=a(1-e)$ is the perihelion distance and i is the ecliptic inclination of the ellipse. Therefore,

Figure 4.11



Patched 'cubic' trajectory formed from four perpendicular halo orbits. The switching of the sail attitude occurs at the intersection points.

Figure 4.12



Halo orbit patched to a Keplerian ellipse. The transfer may occur at either the perihelion or aphelion points and to or from an off-axis halo.

applying condition (iii) the required angular velocity of the inital halo orbit is given in terms of the orbital elements of the ellipse by

$$\Omega_1(a,e;i) = a^{-3/2} \cos^{-1}(i) \frac{(1 + e)^{1/2}}{(1 - e)^{3/2}} \qquad (4.52)$$

For an aphelion transfer an angular velocity of $\Omega_1(a,-e;i)$ is required and for a transfer to a circular orbit an angular velocity of $\Omega_1(a,0;i)$ is required. Again, due to the symmetry of the system the sail may be transferred to or from an off-axis halo orbit. In particular the spacecraft may be transferred from an off-axis halo orbit to a solar polar orbit.

4.9 Conclusions

It has be shown in this chapter that solar sail spacecraft may be used to establish heliocentric halo type orbits. Three distinct families of heliocentric halo orbit exist with different requirements on the sail loading parameter and with different stability characteristcs. The fixed period orbits, such as the Earth synchronous one year halo orbit, are the most dynamically complex. However, the minimal loading halo orbits are of more interest for practical applications due to the less demanding requirements on the total spacecraft mass per unit area.

Although the Keplerian synchronous halo orbits and some of the fixed period halo orbits are unstable a simple feedback control scheme using variable sail attitude and loading gives a suitable control for the unstable modes. The simpler fixed Sun-sail pitch angle control is easier to implement, although open loop control would be required at regular intervals to compensate for any drift in the sail position.

Finally, it has been demonstrated that the individual halo orbits may be patched together to form complex new trajectories, such as the 'cubic' trajectory. The patching of a halo orbit to a Keplerian ellipse is also possible by switching the sail attitude into a null orientation.

5. SOLAR SAIL GEOCENTRIC HALO ORBITS

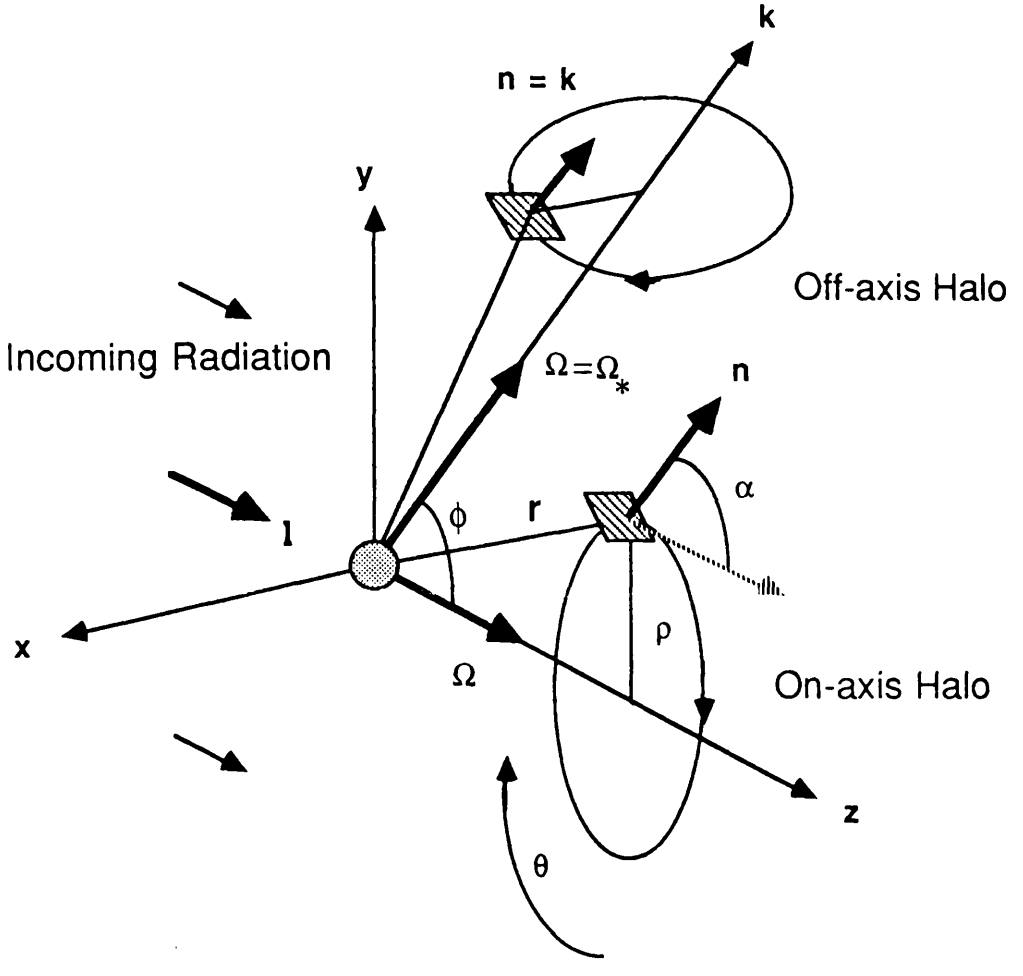
5.1 Introduction

Following the investigation of the heliocentric halo orbit families in chapter 4 another new mode of operation of solar sail spacecraft is now discussed, that of geocentric halo orbits. These orbits are similar to the heliocentric halo orbit families in that they are achieved by orienting the sail such that a component of the solar radiation pressure force is directed out of the orbital plane. Therefore, the geocentric halo orbit is a circular geocentric orbit, but displaced in the anti-Sun direction, Figure 5.1.

By suitably choosing the sail pitch angle α and loading parameter β it will be demonstrated that the spacecraft orbital period T, halo amplitude ρ and out-of-plane displacement distance z may be chosen at will. The sail orbital period may be chosen to be synchronous with a Keplerian near polar orbit of semi-major axis equal to the halo amplitude, fixed at some particular value for all orbit parameters (ρ,z) or, chosen to minimise the sail loading requirements. For the minimal loading family of halo orbits it is found that the axis of the halo need not lie along the Sun-Earth line, Figure 5.1.

The dynamical model assumes a radiation field that is uniform over the scale of the problem (tens of planetary radii) so that the ratio of the solar radiation pressure acceleration to the local gravitational acceleration increases with increasing distance from the planetary centre. It is this relation that leads to interesting new dynamics. The region of space over which the assumptions remain valid will be limited due to solar and lunar gravitational perturbations. Given that the lunar sphere of influence has a radius of order 10 R_O

Figure 5.1



Schematic geometry of a geocentric halo orbit with the solar sail at position **r**=(ρ,θ,z) displaced in the anti-Sun direction. The sail attitude is defined by a unit vector **n** and the solar radiation is incident along the Sun-line direction **l**. The reference frame rotates with angular velocity **Ω** directed along **l** or, for the off-axis case, directed along **k**.

(planetary radii) halo orbits with a geocentric distance of up to 40-50 R_O will be considered. Over this distance the solar radiation pressure varies by only $4x10^{-3}$ from its value at 1 AU. The analysis is of course invalid along the planetary shadow near $\rho=0$. In the initial analysis the annual rotation of the Sun-line will also be neglected. The inertial forces introduced by a further rotation of the coordinate system with the Sun-line are relatively small with respect to the large solar radiation pressure force required to establish the halo orbits. The effect of these perturbing forces will be considered in section 5.7. It will be assumed that these effects may be corrected for through active control, as discussed in section 5.8.

The dynamical stability of the various modes of halo orbit will be investigated and linearly stable families identified. For the unstable families simple control schemes are developed by using a feedback to the sail pitch. However, long damping timescales are obtained which renders this control unsuitable. A well damped control is obtained by including first order variations in the sail loading parameter in the feedback loop. This control is suitable for stabilising the unstable halo orbit families against perturbations, such as the annual rotation of the Sun-line.

Lastly, by patching together individual halo orbits it will be demonstrated that complex new trajectories may be formed by a simple switching operation on the sail attitude at discrete points along the orbit. The patching may be between individual halo orbits or between a halo orbit and a Keplerian ellipse. Geocentric halo orbit families have potentially useful applications for near Earth space science missions, as will be discussed in chapter 7.

5.2 Dynamical Equations and Their Solution

Following the analysis of section 4.2 the dynamical equations will be considered for a planar, perfectly reflecting solar sail in a co-rotating reference frame with the origin centred on a point mass Earth. The axis of rotation $\boldsymbol{\Omega}$ will be directed along the Sun-Earth line, Figure 5.1. The sail attitude is again defined by a unit vector $\mathbf{n}$ fixed in the co-rotating frame and the magnitude of the solar radiation pressure acceleration is given by the parameter $\beta=P_o/\sigma$, where P_o is the solar radiation pressure acceleration at 1 AU and σ is the total spacecraft mass per unit area. Since the sail attitude is fixed in the co-rotating frame the sail must rotate once per orbit with respect to an inertial frame.

The vector dynamical equation for a solar sail in the co-rotating frame under the action of a point mass potential and superimposed uniform radiation field is given by

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) = \mathbf{a} - \nabla\Phi_2(|\mathbf{r}|) \qquad (5.1)$$

The two-body gravitational potential $\Phi_2(|\mathbf{r}|)$ and the solar radiation pressure acceleration $\mathbf{a}$ are given by

$$\Phi_2(|\mathbf{r}|) = -\frac{\mu}{|\mathbf{r}|} \quad , \quad \mathbf{a} = \beta\,(\mathbf{l}.\mathbf{n})^2\,\mathbf{n} \qquad (5.2)$$

where the unit vector $\mathbf{l}=(0,0,1)$ is directed along the (assumed) fixed Sun-line. The requirement $\mathbf{l}.\mathbf{n}\geqslant 0$ is imposed to ensure that the normal to the sail always points away from the Sun which consequently constrains the sail motion to the planetary night-side (+z).

Equation (5.1) may be simplified by introducing the scalar potential $\Psi(\mathbf{r})$ to represent the conservative centrifugal term. Similarly

the solar radiation pressure acceleration is, in this case, conservative and may be written in terms of a scalar potential $\Gamma(\mathbf{r},\mathbf{n})$, viz

$$\nabla\Psi(\mathbf{r},\Omega) = \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) \quad , \quad \Psi(\mathbf{r},\Omega) = -\frac{1}{2} |\boldsymbol{\Omega} \times \mathbf{r}|^2 \tag{5.3a}$$

$$\nabla\Gamma(\mathbf{r},\mathbf{n}) = \beta\, (\mathbf{l}.\mathbf{n})^2\, \mathbf{n} \quad , \quad \Gamma(\mathbf{r},\mathbf{n}) = \beta\, (\mathbf{l}.\mathbf{n})^2\, (\mathbf{r}.\mathbf{n}) \tag{5.3b}$$

Defining new potentials $U(\mathbf{r},\Omega)=\Phi_2(|\mathbf{r}|)+\Psi(\mathbf{r},\Omega)$ and $V(\mathbf{r},\mathbf{n};\Omega)=U(\mathbf{r},\Omega)+\Gamma(\mathbf{r},\mathbf{n})$, equation (5.1) then becomes the reduced dynamical equation

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \nabla V(\mathbf{r},\mathbf{n};\Omega) = 0 \tag{5.4}$$

In the co-rotating frame stationary solutions are again required so that the first two terms of equation (5.4) vanish. Therefore, since the vector $\nabla\Gamma(\mathbf{r},\mathbf{n})$ is oriented in direction $\mathbf{n}$, taking the vector product of $\mathbf{n}$ with equation (5.4) it is found that

$$\nabla U(\mathbf{r},\Omega) \times \mathbf{n} + \nabla\Gamma(\mathbf{r},\mathbf{n}) \times \mathbf{n} = 0 \;\Rightarrow\; \mathbf{n} = \lambda \nabla U(\mathbf{r},\Omega) \tag{5.5}$$

where λ is an arbitrary scalar multiplier. The normalisation condition $|\mathbf{n}|=1$ is then used to identify λ as $|\nabla U(\mathbf{r},\Omega)|^{-1}$. The sail attitude required for a stationary solution in the co-rotating frame is therefore defined by

$$\mathbf{n} = \frac{\nabla U(\mathbf{r},\Omega)}{|\nabla U(\mathbf{r},\Omega)|} \tag{5.6}$$

Since the spacecraft is to be in uniform co-rotation there can be no azimuthal component of the solar radiation pressure acceleration. Therefore, there can be no component of the vector $\mathbf{a}$ in the azimuthal

direction. The sail attitude may then be described by a single angle α between **l** and **n** defined by

$$\tan \alpha(\mathbf{r},\Omega) = \frac{|\mathbf{l} \times \nabla U(\mathbf{r},\Omega)|}{\mathbf{l}.\nabla U(\mathbf{r},\Omega)} \qquad (5.7)$$

Similarly, the required solar radiation pressure acceleration may be obtained by taking a scalar product of **n** with equation (5.4). Again requiring a stationary solution in the co-rotating frame it is found that

$$\beta(\mathbf{r},\Omega) = \frac{\nabla U(\mathbf{r}).\mathbf{n}}{(\mathbf{l}.\mathbf{n})^2} \qquad (5.8)$$

In geocentric cylindrical polar coordinates (ρ,θ,z) the co-rotating potential may be written as

$$U(\rho,z;\Omega) = -\left\{ \frac{1}{2} (\Omega\rho)^2 + \frac{\mu}{r} \right\} \quad , \quad r^2 = \rho^2 + z^2 \qquad (5.9)$$

where the period of the halo orbit is given by $T=2\pi/\Omega$. Therefore, evaluating the potential gradient it is found that the scalar expressions for the Sun-sail pitch angle and the solar radiation pressure acceleration required for a halo orbit of amplitude ρ, displacement z and period T are given by

$$\tan \alpha(\rho,z;\Omega) = \left\{\frac{\rho}{z}\right\} \left\{1 - \left\{\frac{\Omega}{\Omega_*}\right\}^2 \right\} \quad , \quad \Omega_*{}^2 = \mu/r^3 \qquad (5.10a)$$

$$\beta(\rho,z;\Omega) = \Omega_*{}^2 \left\{1 + \left\{\frac{\rho}{z}\right\}^2 \left\{1 - \left\{\frac{\Omega}{\Omega_*}\right\}^2 \right\}^2 \right\}^{3/2} z \qquad (5.10b)$$

where Ω_* is the angular velocity of a circular Keplerian orbit at geocentric distance r. The case of $\Omega=0$ corresponds to the static

equilibrium 'statite' concept for low bit rate communications with high latitude regions, as discussed in section 1.5.3. The sail is then stationary above the night-side of the Earth with the solar radiation pressure acceleration exactly balancing the local gravitational acceleration. The required sail acceleration is therefore μz^{-2}, as discussed by Forward (1989b) using the same dynamical model. If the unit of length is chosen to be the planetary radius R_o and $\mu=1$ then the parameter β will be the solar radiation pressure acceleration made dimensionless with respect to the gravitational acceleration at R_o. This will now be termed the sail loading parameter. The total spacecraft mass per unit area is then obtained from the relation $\sigma=9.31\times10^{-4}\beta^{-1}$ gm^{-2}.

5.3 Polar Synchronous Mode

For this mode of operation the spacecraft orbital period will be chosen to be equal to that of a Keplerian near polar orbit above the planetary terminator with a geocentric distance equal to the halo amplitude ρ, (ie. $\Omega=\rho^{-3/2}$). Therefore, the sail will maintain polar synchronism at all displacement distances z so that there will be cylindrical surfaces of co-rotation extending in the anti-Sun direction. Solar sails with the same halo amplitude will then orbit synchronously with each other at differing displacement distances. The equatorial inclination of the Keplerian orbit will vary from 66.5° at the solstices to 90° at the equinoxes. From equations (5.10) the required Sun-sail pitch angle and sail loading parameter are obtained as

$$\tan\alpha(\rho,z) = \left\{\frac{\rho}{z}\right\}\left\{1-\left\{1+\left\{\frac{z}{\rho}\right\}^2\right\}^{3/2}\right\} \qquad (5.11a)$$

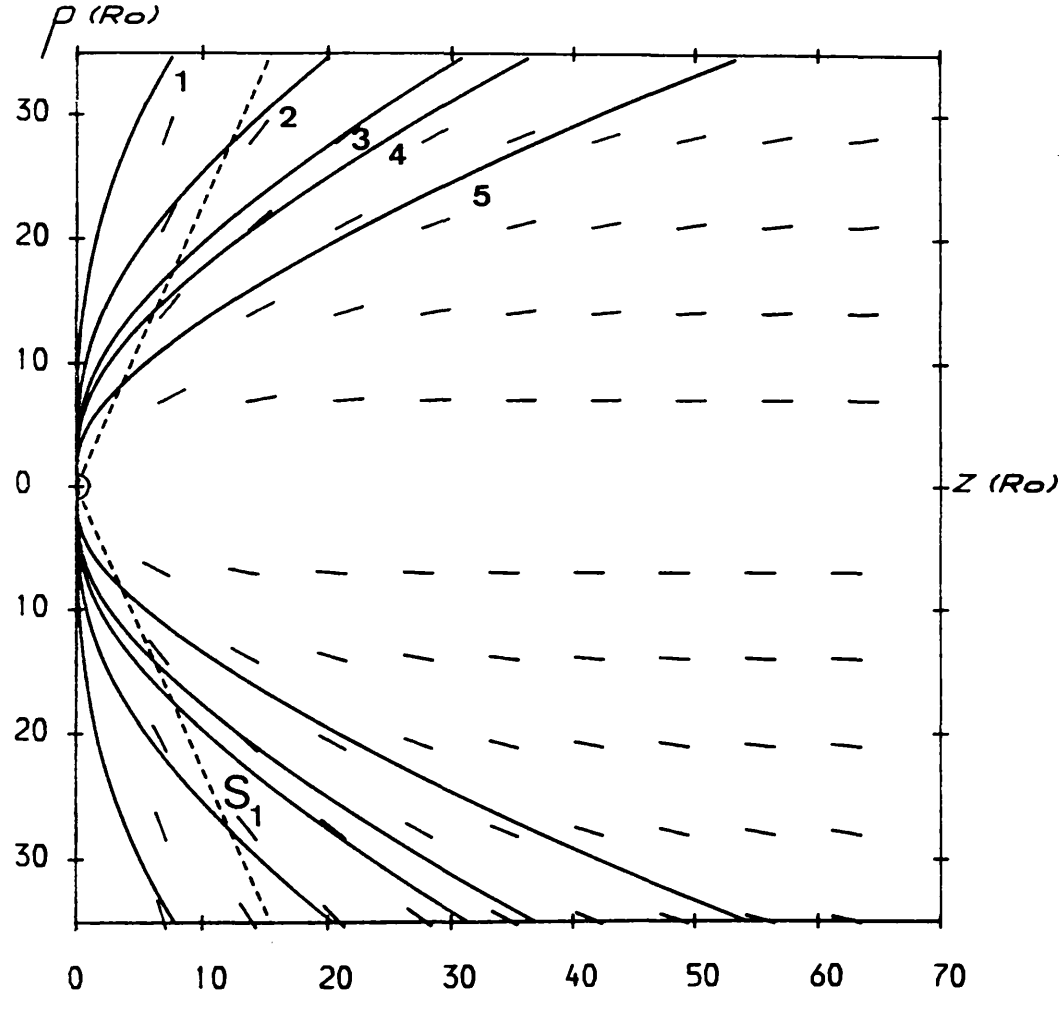
$$\beta(\rho,z) = \frac{z}{r^3}\left\{1 + \left\{\frac{\rho}{z}\right\}^2 \left\{1 - \left\{1 + \left\{\frac{z}{\rho}\right\}^2\right\}^{3/2}\right\}^2\right\}^{3/2} \qquad (5.11b)$$

A section of the level surfaces of constant sail loading generated by equation (5.11b), along with the required sail pitch for polar synchronism is shown in Figure 5.2. It can be seen that for reasonably low sail loadings halo orbits of large amplitude with respect to their displacement distance are required. For example a 5.24 day halo orbit with an amplitude of 20 R_o and a displacement distance of 5 R_o requires a demanding spacecraft mass per unit area of 1.33 gm^{-2}. The surfaces of constant sail loading approach the z=0 plane as $\beta \to 0$ corresponding to near polar Keplerian orbits.

5.4 General Synchronous Mode

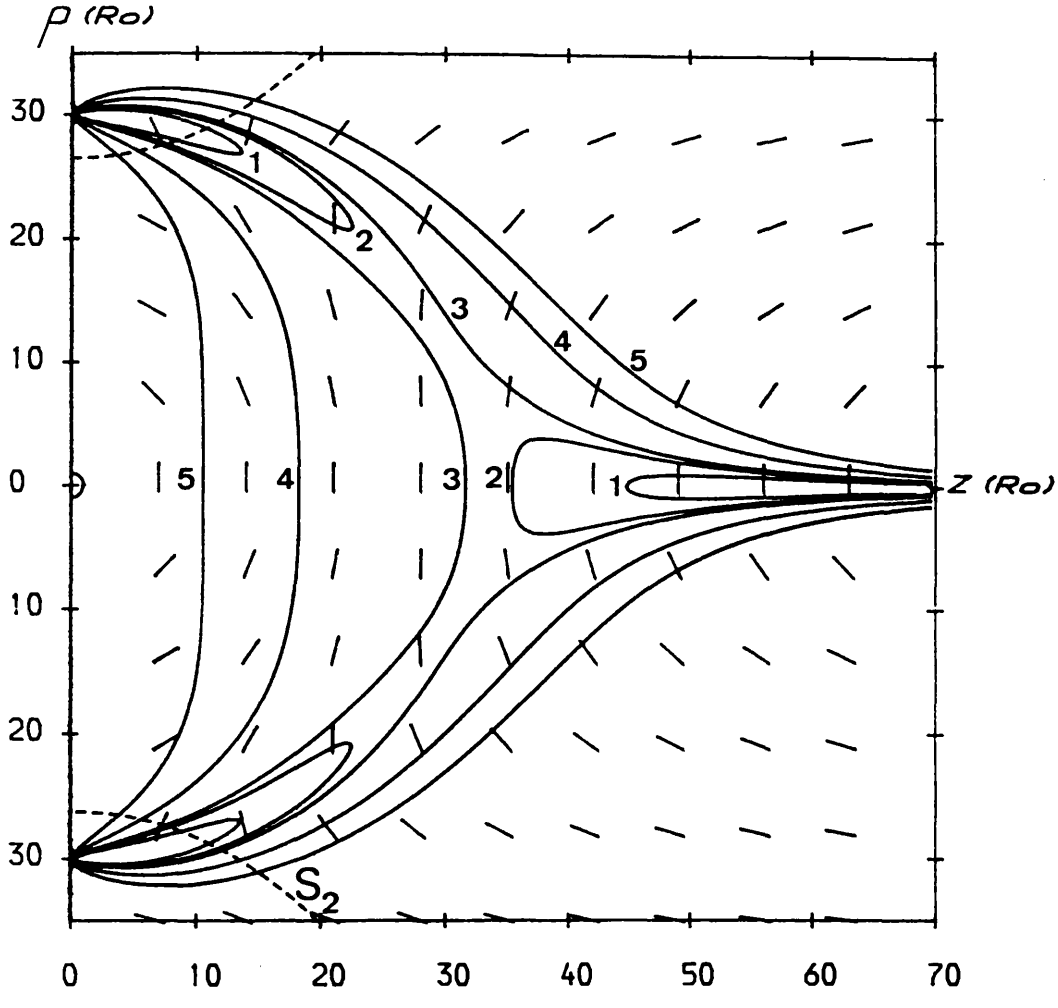
For this mode of operation the sail orbital period will be chosen to be fixed at some particular value for all halo orbit parameters (ρ,z). This is equivalent to choosing the period to be synchronous with some particular Keplerian orbit with a geocentric distance r_o, (ie. $\Omega = r_o^{-3/2}$). From equation (5.10b) surfaces of constant sail loading may be generated with Ω chosen to be some fixed value. Figure 5.3 shows a section of these level surfaces with the sail orbital period chosen to be synchronous with a Keplerian orbit of radius 30 R_o, (period≈9.6 days in the geocentric case). It can be seen that for the lower values of sail loading there are two topologically disconnected surfaces. These surfaces correspond to large amplitude halo orbits near the 30 R_o synchronous region and large displacement, low amplitude halo orbits with a low iocal gravitational acceleration. As the sail loading increases these surfaces expand and connect. It can be seen that along the ρ=0 axis the sail pitch angle is zero corresponding to

Figure 5.2



Section of surfaces of constant sail loading and the required sail pitch for the polar synchronous case. The required sail loadings are given by; (1) $2x10^{-4}$ (2) $8x10^{-4}$ (3) $2x10^{-3}$ (4) $3x10^{-3}$ (5) $9x10^{-3}$. The contour S_1 represents the partitioning of the stable and unstable halo orbit families.

Figure 5.3



Section of surfaces of constant sail loading and the required sail pitch for the general synchronous case. The required sail loadings are given by; (1) 5×10^{-4} (2) 8×10^{-4} (3) 1×10^{-3} (4) 3×10^{-3} (5) 9×10^{-3}. The contour S_2 represents the partitioning of the stable and unstable halo orbit families.

stationary solutions.

5.5 Optimal Halo Mode

The sail orbital period will now be treated as a free parameter of the system so that the sail loading requirements may be minimized with respect to it, to obtain an optimal family of halo orbits. Therefore, setting the derivative of β with respect to Ω to zero it is found that

$$\frac{\partial\beta(\rho,z;\Omega)}{\partial\Omega} = 0 \quad \Rightarrow \quad \Omega = \Omega_* \tag{5.12}$$

For the minimisation of the sail loading it is therefore required that the sail orbital period is equal to the orbital period of a Keplerian orbit of geocentric distance r. With this orbital period the minimized sail loading and the required sail pitch angle are given by equations (5.10) as

$$\tan\alpha = 0 \tag{5.13a}$$

$$\beta = \frac{z}{r^3} \tag{5.13b}$$

The sail attitude is therefore such that the normal to the sail surface is directed along the Sun-line. In this case the general three-dimensional motion of the sail can be obtained in closed form, as discussed in section 5.9.

Level surfaces of constant sail loading may be generated by fixing $\beta=\beta_0$. Then, equation (5.13b) can be inverted to give

$$\rho(z) = \left\{ \frac{z^{2/3}}{\Lambda} - z^2 \right\}^{1/2}, \quad \Lambda = \beta_0^{2/3} \tag{5.14}$$

which defines a surface of revolution about the z-axis. Making the substitution $u=z^{2/3}$ to clear the radical equation (5.14) may be written as

$$\rho^2(u) = u\,(\Lambda^{-1/2} + u)(\Lambda^{-1/2} - u) \tag{5.15}$$

so that $\rho=0$ when $u=0$ or $u=\pm\Lambda^{-1/2}$, (ie. $z=0$ or $z=\beta_o^{-1/2}$). The point $z=\beta_o^{-1/2}$ corresponds to the 'statite' solution with the solar radiation pressure acceleration balancing the local gravitational acceleration. For a fixed sail loading the point of maximum halo amplitude may be obtained by setting $(d\rho/du)=0$, viz

$$\frac{d\rho(u)}{du} = \frac{1}{2}\,\frac{(\Lambda^{-1} - 3u^2)}{(\Lambda^{-1}u - u^3)^{1/2}} \;\Rightarrow\; u = (3\Lambda)^{-1/2} \tag{5.16}$$

The maximum displacement along the Sun-line is therefore $z_m=(3)^{-3/4}\beta_o^{-1/2}$. Furthermore, at this extremal value of z it is found from equation (5.14) that

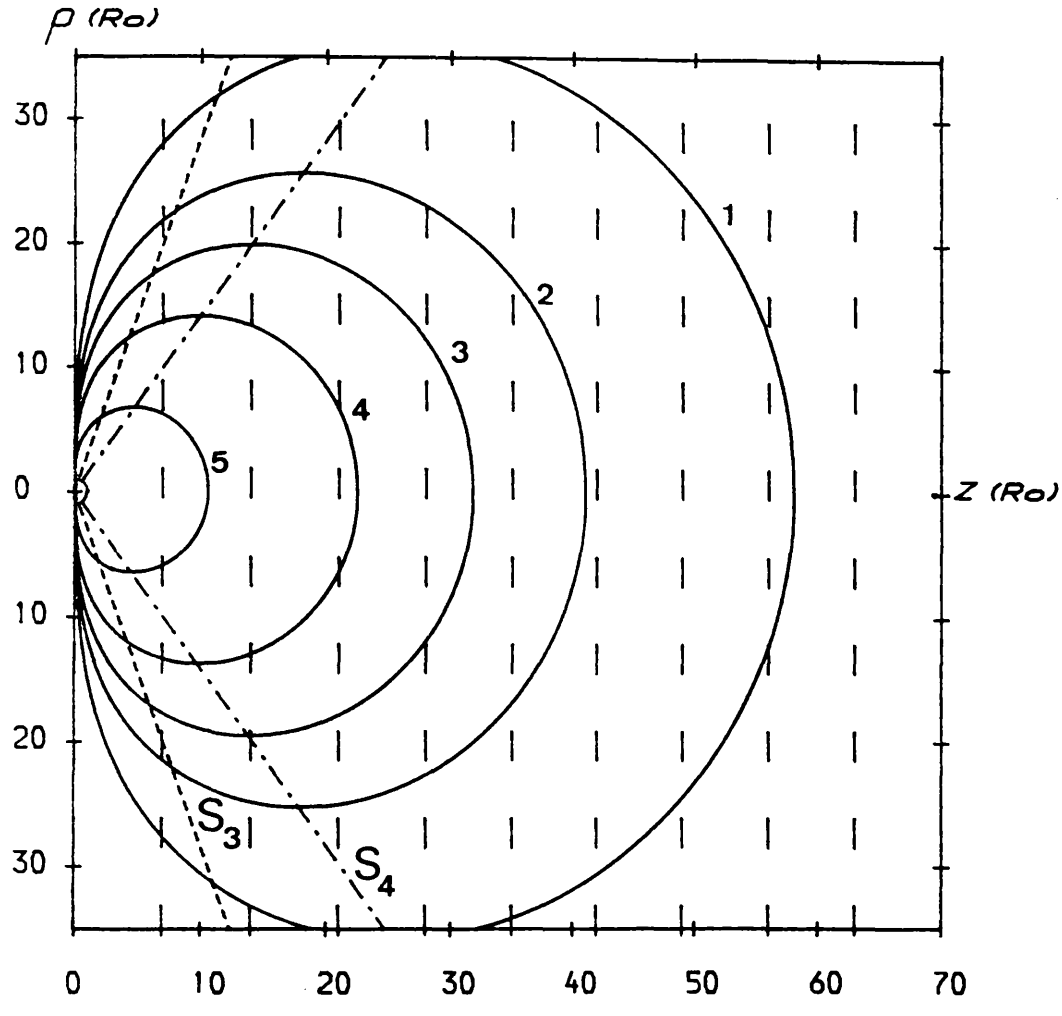
$$\rho_m = 2^{1/2}\,(3)^{-3/4}\,\beta_o^{-1/2} \tag{5.17}$$

Therefore, the locus of the maximum of halo amplitudes is given by a cone defined by

$$\rho_m = \sqrt{2}\,z_m \tag{5.18}$$

Sections of level surfaces of constant sail loading are shown in Figure 5.4. It can be seen that for small amplitudes and displacements the required minimized sail loading varies rapidly. The line S_4 gives the locus of the maximum halo amplitudes, as defined by equation (5.18). Statite type solutions are again shown along the $\rho=0$ axis. The

Figure 5.4



Section of surfaces of constant sail loading and the required sail pitch for the optimal case. The required sail loadings are given by; (1) $3x10^{-4}$ (2) $6x10^{-4}$ (3) $1x10^{-3}$ (4) $2x10^{-3}$ (5) $9x10^{-3}$. The contour S_3 represents the partitioning of the stable and unstable halo orbit families and the contour S_4 represents the locus of maximum amplitude halo orbits.

surfaces of co-rotation are defined by spheres of constant Ω_*. The intersection of these co-rotation surfaces with the surfaces of constant sail loading then defines regions where solar sails deployed on minimal loading halo orbits will orbit synchronously with each other.

It can be seen from Figure 5.4 that for reasonable values of spacecraft mass per unit area of 2-10 gm^{-2} large amplitude halo orbits are required. A 15.38 day optimised geocentric halo orbit with an amplitude of 30 R_O and a displacement distance of 40 R_O requires a mass per unit area of 2.91 gm^{-2}, which is near current attainable values. For halo orbits around other bodies the requirements on the spacecraft mass per unit area are not so great, even for Mars and especially for Mercury. For halo orbits at Mercury however the Sun-planet line rotates rapidly so that large corrective manoeuvres would be required to maintain the halo orbit. Furthermore, as the halo amplitude and displacement distance increase solar perturbations become of increasing importance as do lunar perturbations for the geocentric case. However, with advanced sail materials and designs much smaller and less perturbed halo orbits would be possible.

Since, for the optimal case, the sail normal is directed along the axis of symmetry of the system, the axis of the halo orbit need not in fact lie along the Sun-line. Any axis $\mathbf{k}$ passing through the origin, such that $\mathbf{l}.\mathbf{k}>0$, will generate new off-axis halo orbits, as shown in Figure 5.1. The sail loading requirements are however increased by a factor $(\mathbf{l}.\mathbf{k})^{-2}$ due to the oblique incidence of the photons on the sail as the normal to the sail surface is directed along the new axis $\mathbf{n}=\mathbf{k}$.

5.6 Geocentric Halo Orbit Stability

The dynamical stability of the geocentric halo orbit modes will

now be investigated through a linear perturbation analysis similar to section 4.6 and stable families of halo orbits identified. Since the dynamical model is conservative there are no dissipatve terms so that asymptotic stability will not be possible. Assuming the sail is at some operating point $\mathbf{r}_o=(\rho_o,\theta_o,z_o)$, a perturbation $\mathbf{r}_o\rightarrow\mathbf{r}_o+\boldsymbol{\delta}$ is applied to equation (5.4) to obtain a variational equation, viz

$$\frac{d^2\boldsymbol{\delta}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\boldsymbol{\delta}}{dt} + \nabla V(\mathbf{r}_o+\boldsymbol{\delta}) = 0 \tag{5.19}$$

where $\boldsymbol{\delta}=(\xi,\psi,\eta)$ represents small displacements in the co-rotating frame along the (ρ,θ,z) directions. Expanding the potential gradient in a trivariate Taylor series about the point $\mathbf{r}_o$ it is found that

$$\nabla V(\mathbf{r}_o+\boldsymbol{\delta}) = \nabla V(\mathbf{r}_o) + \frac{\partial}{\partial \mathbf{r}} \nabla V(\mathbf{r}) \Big|_{\substack{\mathbf{r}=\mathbf{r}_o,\ \boldsymbol{\eta}=\boldsymbol{\eta}_o}} \boldsymbol{\delta} + 0\ (|\boldsymbol{\delta}|^2) \tag{5.20}$$

Then, since $\nabla V(\mathbf{r}_o)=0$ and the radiation field is uniform, so that $(\partial\nabla\Gamma/\partial\mathbf{r})=0$, a linear variational system is obtained, viz

$$\frac{d^2\boldsymbol{\delta}}{dt^2} + \mathbf{M}_1 \frac{d\boldsymbol{\delta}}{dt} + \mathbf{M}\,\boldsymbol{\delta} = \mathbf{0} \tag{5.21}$$

where $\mathbf{M}=\partial\nabla U/\partial\mathbf{r}$, the gravity gradient tensor, and the skew symmetric gyroscopic matrix $\mathbf{M}_1$ are given by

$$\mathbf{M}_1 = \begin{Bmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{Bmatrix}, \quad \mathbf{M} = (U_{ij})_{(i,j)\in(\rho,\theta,z)} \tag{5.22}$$

where U_{ij} is the (i,j) partial derivative of the potential. Since the potential is azimuthally symmetric all partial derivatives with respect

to θ vanish. Furthermore, due to the potential being conservative, $\nabla x \nabla U = 0$ so that $M_{31} = M_{13}$. In component form the variational equations are then

$$\frac{d^2\xi}{dt^2} - 2\Omega\rho_o\frac{d\Psi}{dt} + M_{11}\xi + M_{13}\eta = 0 \tag{5.23a}$$

$$\frac{d^2\Psi}{dt^2} + \frac{2\Omega}{\rho_o}\frac{d\xi}{dt} = 0 \tag{5.23b}$$

$$\frac{d^2\eta}{dt^2} + M_{13}\xi + M_{33}\eta = 0 \tag{5.23c}$$

This set of three coupled ordinary differential equations may be reduced by integrating equation (5.23b), viz

$$\frac{d\Psi}{dt} = -\frac{2\Omega}{\rho_o}(\xi - \xi_o) \tag{5.24}$$

which can then be substituted into equation (5.23a). This then leads to a constant term $4\Omega^2\xi_o$ in equation (5.23a) which may easily be removed by re-scaling through a change of variable

$$\xi' = \xi - \frac{4\Omega^2 M_{33}}{M^*_{11}M_{33} - M_{13}^2}\xi_o \tag{5.25a}$$

$$\eta' = \eta + \frac{4\Omega^2 M_{13}}{M^*_{11}M_{33} - M_{13}^2}\xi_o \tag{5.25b}$$

where $M^*_{11} = M_{11} + 4\Omega^2$. A reduced variational equation is then obtained, viz

$$\frac{d^2}{dt^2}\begin{Bmatrix}\xi' \\ \eta'\end{Bmatrix} + \begin{bmatrix} M^*_{11} & \vdots & M_{13} \\ \cdots & \vdots & \cdots \\ M_{13} & \vdots & M_{33}\end{bmatrix}\begin{Bmatrix}\xi' \\ \eta'\end{Bmatrix} = \begin{Bmatrix}0 \\ 0\end{Bmatrix} \tag{5.26}$$

As with the heliocentric case the sail is free to drift along the halo

orbit. Therefore, the orbit may only be Poincaré stable and cannot be Lyapunov stable. The coefficients of the matrix **M** are given by

$$M^*_{11} = 4\Omega^2 - \left\{ \left\{\Omega^2 - \frac{1}{r^3}\right\} + \frac{3\rho^2}{r^5} \right\} \tag{5.27a}$$

$$M_{13} = - \frac{3\rho z}{r^5} \tag{5.27b}$$

$$M_{33} = \left\{\frac{1}{r^3} - \frac{3z^2}{r^5}\right\} \tag{5.27c}$$

The stability characteristics of the system may now be investigated by substituting a solution of the form

$$\begin{Bmatrix} \xi' \\ \eta' \end{Bmatrix} = \begin{Bmatrix} \xi_0 \\ \eta_0 \end{Bmatrix} e^{\omega t} \tag{5.28}$$

to calculate the system eigenvalues. Substituting this solution into equation (5.26) yields a matrix equation

$$\begin{Bmatrix} \omega^2 + M^*_{11} & \vdots & M_{13} \\ \cdots\cdots & \vdots & \cdots\cdots \\ M_{13} & \vdots & \omega^2 + M_{33} \end{Bmatrix} \begin{Bmatrix} \xi_0 \\ \eta_0 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{5.29}$$

For non-trivial solutions it is required that the secular determinant of this matrix equation vanish. The characteristic polynomial of the system is then found to be

$$\omega^4 + tr(\mathbf{M})\omega^2 + det(\mathbf{M}) = 0 \tag{5.30}$$

where the eigenvalues ω_j (j=1,4) are the four frequencies of the eigenmodes of the system. Formally the eigenvalues may be written as

$$\omega_{1..4} = \frac{\pm 1}{\sqrt{2}} \left\{ -tr(\mathbf{M}) \pm \left\{ tr(\mathbf{M})^2 - 4det(\mathbf{M}) \right\}^{1/2} \right\}^{1/2} \qquad (5.31)$$

where the trace and determinant of the matrix **M** are given by

$$tr(\mathbf{M}) = 3\Omega^2 - \Omega_*^2 \quad , \quad det(\mathbf{M}) = 3\Omega_*^2\Omega^2 \left\{ 1 - \left\{ \frac{z}{r} \right\}^2 \right\} - 2\Omega_*^4 \qquad (5.32)$$

Therefore, substituting into equation (5.31) the four eigenvalues of the system may be written as

$$\omega_{1..4} = \pm \frac{\Omega_*}{\sqrt{2}} \left\{ \left\{ 1-3\left\{ \frac{\Omega}{\Omega_*} \right\}^2 \right\} \pm \left\{ 9\left\{ 1-\left\{ \frac{\Omega}{\Omega_*} \right\}^2 \right\}^2 + 36\left\{ \frac{\Omega}{\Omega_*} \right\}^2 \left\{ \frac{z_o}{r_o} \right\}^2 \right\}^{1/2} \right\}^{1/2} \qquad (5.33)$$

The sail motion in the neighbourhood of the nominal halo orbit is then given by the superposition of the long period and short period eigenmodes as

$$\begin{Bmatrix} \xi' \\ \eta' \end{Bmatrix} = \sum_{j=1}^{4} \begin{Bmatrix} \xi_{oj} \\ \eta_{oj} \end{Bmatrix} e^{\omega_j t} \qquad (5.34)$$

The sail azimuthal motion may then be obtained by integrating equation (5.24), viz

$$\Psi(t) = \Psi_o + \left\{ \frac{2\Omega\xi_o}{\rho_o} \right\} t - \frac{2\Omega}{\rho_o} \int_o^t \xi(t')\, dt' \qquad (5.35)$$

where $\xi(t)$ is given by equations (5.34) and (5.25a). The first order drift in azimuthal position is then be obtained as

$$\Psi(t) = \Psi_o + \left\{ \frac{2\Omega\xi_o}{\rho_o} \right\} \left\{ 1 - \frac{4\Omega^2 M_{33}}{M^*_{11}M_{33} - M_{13}^2} \right\} t - \frac{2\Omega}{\rho_o} \sum_{j=1}^{4} \frac{\xi_{oj}}{\omega_j} e^{\omega_j t} \qquad (5.36)$$

The stable halo orbit families will therefore have a secular increase in

azimuthal position and so will be constrained to a torus around the nominal halo orbit. The stable families are therefore Poincaré stable.

The stability characteristics of each halo orbit mode may now be investigated by substituting for the required functional form of Ω and determining the regions where the roots of the characteristic polynomial are purely imaginary $\omega_j^2<0$ (j=1,4) giving bound, stable oscillations in the (ρ,z) plane (ie. tr(**M**)>0 and det(**M**)>0).

5.6.1 Polar Synchronous Mode

For this mode $\Omega=\rho^{-3/2}$ so that the first condition for stability tr(**M**)>0 may be written as

$$3\left\{\frac{r}{\rho}\right\}^3 - 1 > 0 \qquad (5.37)$$

Therefore, since $\rho \ll r$, the inequality in equation (5.37) always holds so that there will always be at least one pair of purely imaginary eigenvalues. For stability, with four purely imaginary eigenvalues it is required that the additional condition det (**M**)>0 holds. It may be shown that this condition reduces to

$$\frac{2}{3}\left\{\frac{\rho}{r}\right\}^3 + 3\left\{\frac{z}{r}\right\}^2 - 1 < 0 \qquad (5.38)$$

with an equality in equation (5.38) defining the boundary between regions of stability and instability. With an equality the solution of equation (5.38) is of the form $z=\lambda\rho$, where λ is a constant. Substituting for z, λ can be obtained as the solution of

$$\frac{2}{3} + 3\lambda^2(1+\lambda^2)^{1/2} - (1+\lambda^2)^{3/2} = 0 \qquad (5.39)$$

Numerically it is found that λ=0.442 so that the (ρ,z) plane is partitioned into distinct regions of stability and instability with the condition for stability

$$\rho > 2.26\ z \tag{5.40}$$

The partitioning is therefore defined by a cone, the section of which is shown in Figure 5.2 as the line S_1.

5.6.2 General Synchronous Mode

For the general synchronous mode the halo orbit period is fixed with $\Omega=r_o^{-3/2}$, where r_o is a constant. The condition tr(**M**)>0 may then be written as

$$3\left\{\frac{r}{r_o}\right\}^3 - 1 > 0 \tag{5.41}$$

so that the region defined by $r<(1/3)^{1/3}r_o$ will necessarily be unstable. The determination of the overall stability map requires a numerical solution of the condition det(**M**)>0 which reduces to

$$\frac{2}{3}\left\{\frac{r}{r_o}\right\}^3 + 3\left\{\frac{z}{r}\right\}^2 - 1 < 0 \tag{5.42}$$

The section of the resulting boundary surface is shown as S_2 in Figure 5.3 for the 30 day halo orbit. The intersection of the surface and the ρ-axis is at the point $\rho=(1/3)^{1/3}r_o$.

5.6.3 Optimal Halo Mode

For this mode of halo orbit with a minimized sail loading the spacecraft angular velocity is given by equation (5.12) as $\Omega=\Omega_*$. It is

therefore found that $tr(\mathbf{M})=2\Omega_*^2$ which is of course strictly positive. The condition $det(\mathbf{M})>0$ then reduces to

$$\left\{1 - 9\left\{\frac{z}{r}\right\}^2\right\} > 0 \tag{5.43}$$

This condition is satisfied provided that

$$\rho > 2\sqrt{2}\, z \tag{5.44}$$

so that there are again two distinct regions of stability and instability. The partitioning is therefore defined by a cone, whose section is shown in Figure 5.4 as S_3. For a given sail loading parameter the maximum stable halo displacement distance and corresponding amplitude are given by

$$z_* = \frac{1}{3\sqrt{3}}\,\beta^{-1/2} \quad , \quad \rho_* = \left\{\frac{2}{3}\right\}^{3/2}\beta^{-1/2} \tag{5.45}$$

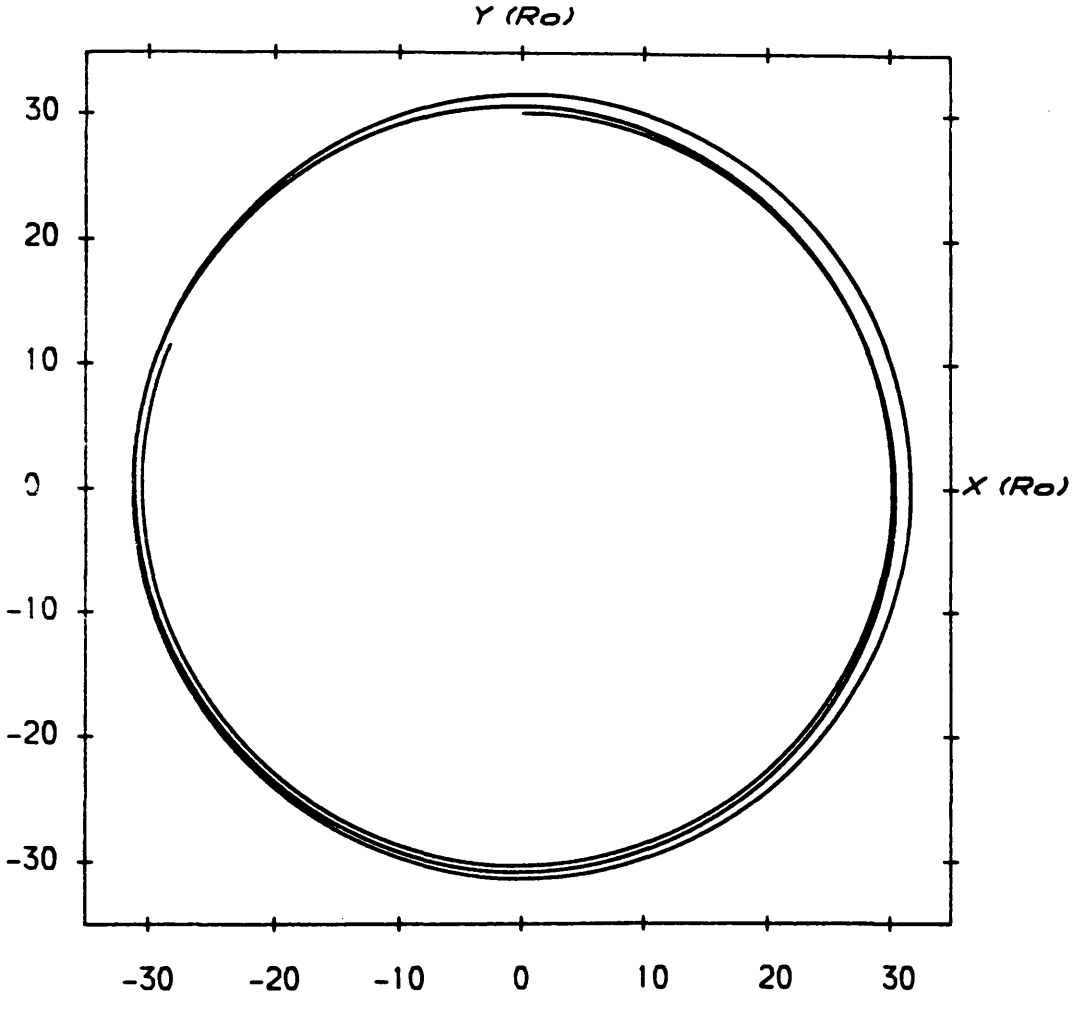
which corresponds to an operating point on the partitioning cone. On the cone of maximum halo amplitude, where $\rho=\sqrt{2}z$, the shortest timescale of instability is given by

$$\tau = \frac{2\pi}{\Omega_*}\,(\sqrt{3} - 1)^{-1/2} \tag{5.46}$$

so that the maximum amplitude, minimal loading halo orbits are unstable on a timescale approximately equal to their orbital period.

Typical stable and unstable responses for 10.13 day and 11.38 day optimal halo orbits are shown in Figures 5.5 and 5.6 using a numerical integration of the full non-linear dynamical equations. It can be seen that the 10.13 day halo orbit is bound, even for large injection errors of $\xi=\eta=0.1$. However, the 11.38 day halo orbit becomes

Figure 5.5(a)



Stable response for a 10.13 day optimal halo orbit with ρ=30, z=8 and injection errors of $\xi_o=\eta_o=0.1$; (a) x-y projection (b) y-z projection.

Figure 5.5(b)

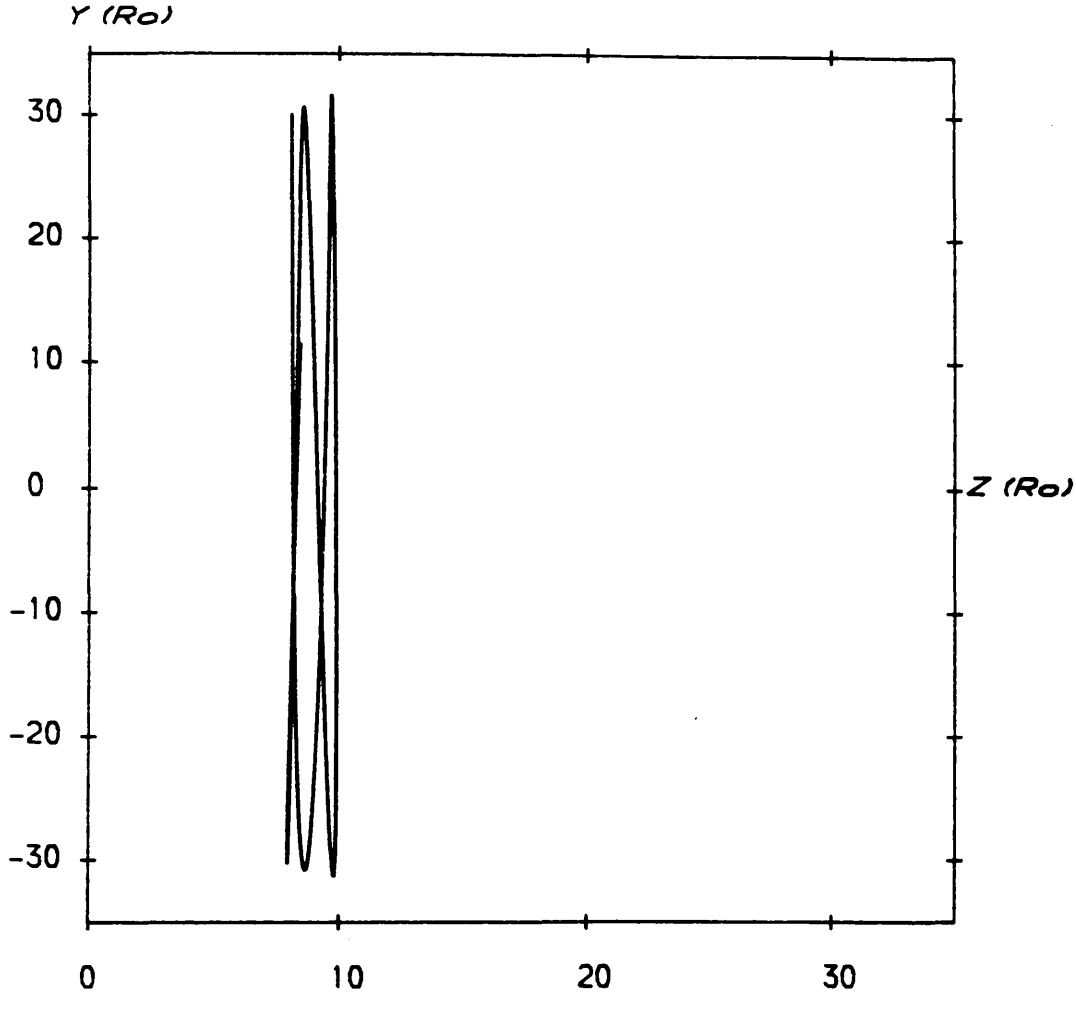
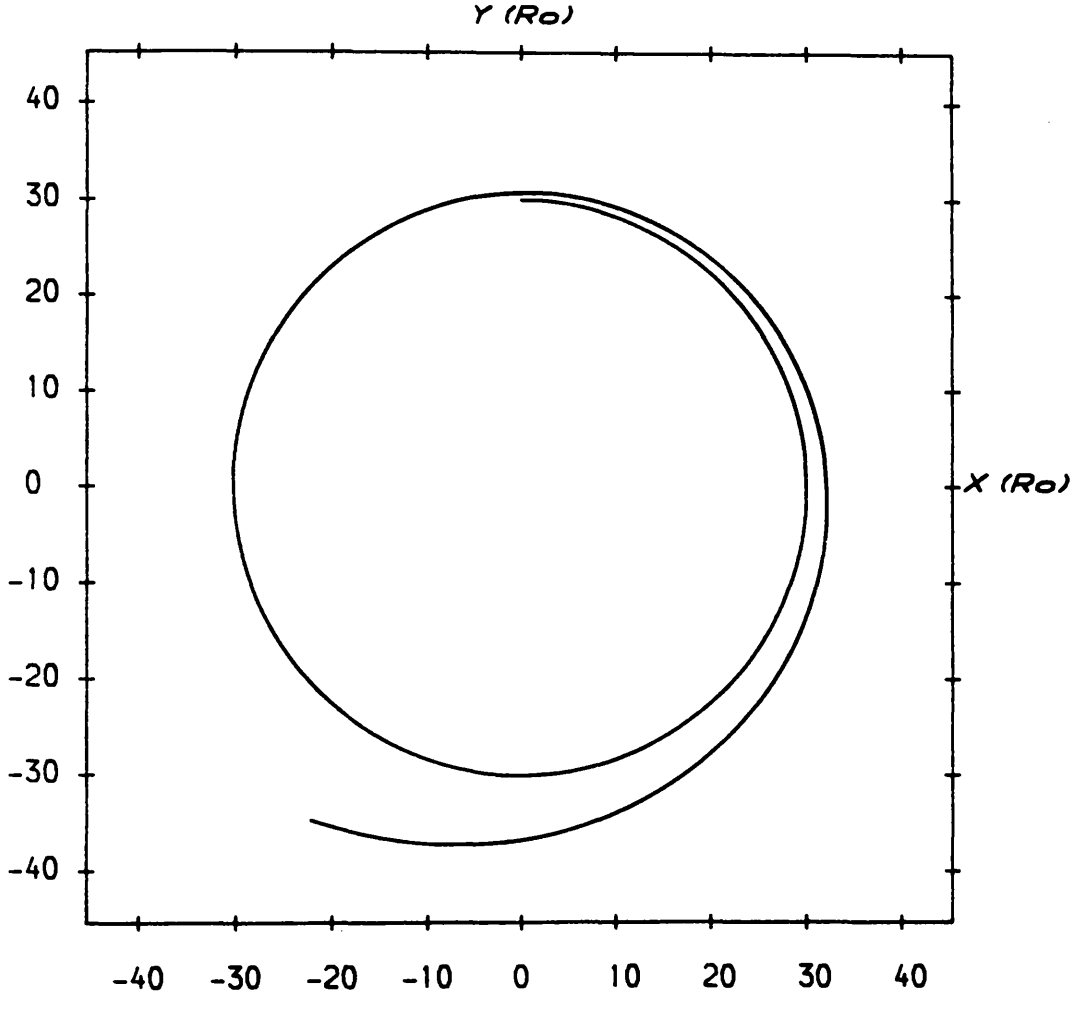
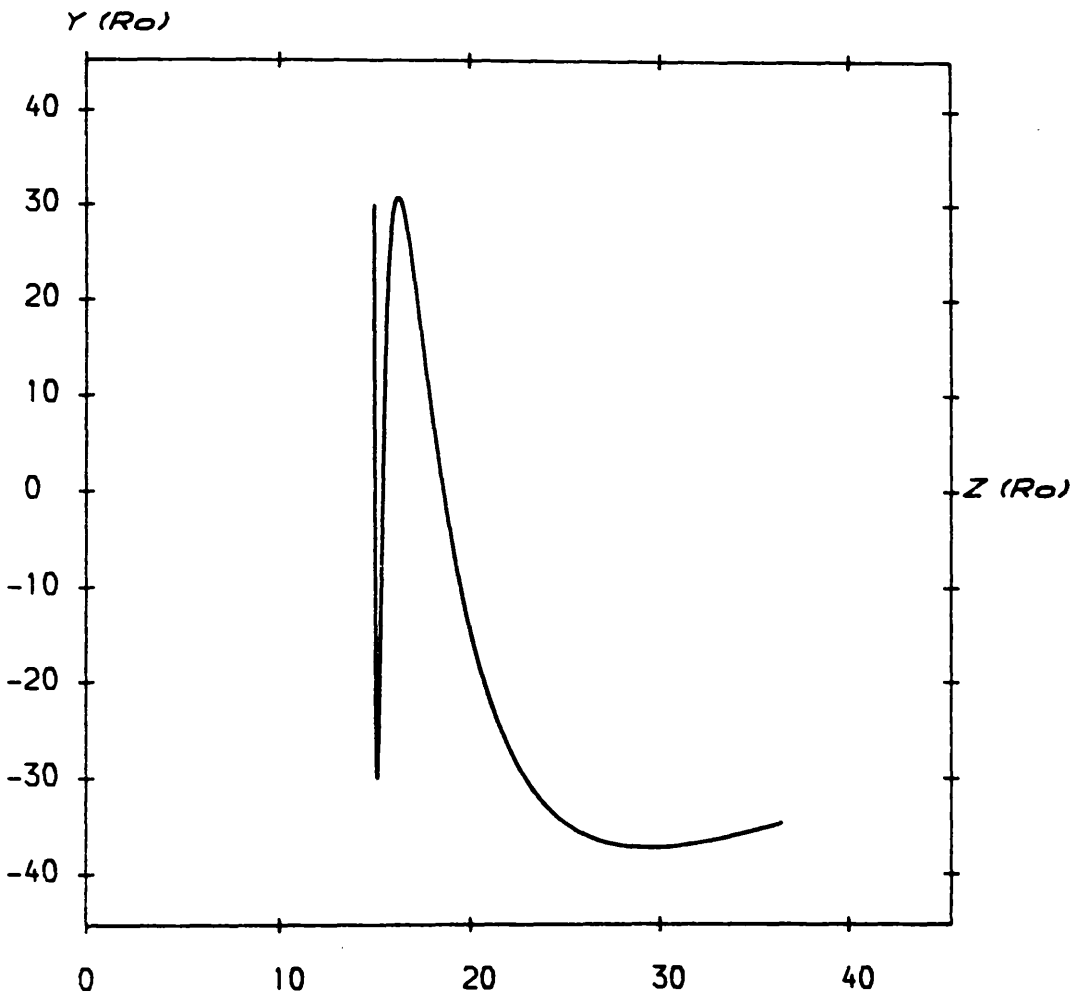


Figure 5.6(a)



Unstable response for a 11.38 day optimal halo orbit with ρ=30, z=15 and injection errors of $\xi_0=\eta_0=10^{-2}$; (a) x-y projection (b) y-z projection.

Figure 5.6(b)



unbound with the solar radiation pressure acceleration driving the spacecraft out of the gravitational potential well.

5.7 Rotating Sun-line Perturbation

Now that the dynamical stability of the various halo orbit modes has been investigated it is necessary to evaluate the effect of perturbations on the nominal halo orbits. These perturbations are due to the assumptions of the dynamical model used, as discussed in section 5.1.

The main perturbation arises from the assumption of a fixed Sun-line. In fact the Sun-line rotates at 0.986^{o} per day due to the heliocentric motion of the Earth. The further rotation of the reference frame to follow the Sun-line therefore results in additional coriolis and centrifugal accelerations. Rotating the reference frame about the y-axis with an angular velocity $\boldsymbol{\Omega}_o=\Omega_o(0,1,0)^T$ the coriolis and centrifugal accelerations are obtained as

$$\mathbf{a}_1 = -2\boldsymbol{\Omega}_o \times \left\{\frac{d\mathbf{r}}{dt}\right\} \quad , \quad \mathbf{a}_2 = -\boldsymbol{\Omega}_o \times(\boldsymbol{\Omega}_o \times \mathbf{r}) \tag{5.47}$$

where the sail position vector $\mathbf{r}=(\rho\cos\theta,\rho\sin\theta,z)$. In the units of the system $\Omega_o=1.604\times10^{-4}$ so that the second order centrifugal term is ignorable. Therefore, evaluating the coriolis acceleration in cylindrical polar coordinates the perturbative acceleration components are obtained as

$$a_\rho = -2\Omega_o \sin\theta\left\{\frac{dz}{dt}\right\} \tag{5.48a}$$

$$a_\theta = -2\Omega_o \cos\theta\left\{\frac{d\rho}{dt}\right\} \tag{5.48b}$$

$$a_z = 2\Omega_o \sin\theta\left\{\frac{d\rho}{dt}\right\} + 2\Omega_o \rho\cos\theta\left\{\frac{d\theta}{dt}\right\} \qquad (5.48c)$$

These accelerations are included in the non-linear dynamical equations in the numerical integration of the sail dynamics. A typical response of a stable, optimal 10.13 day halo orbit is shown in Figure 5.7. It can be seen the the Sun-line rotation gives a periodic response, as would be expected from the periodic nature of the perturbation. For the unstable halo orbit families the perturbation will excite the unstable eigenmodes of the system leading to unbound motion.

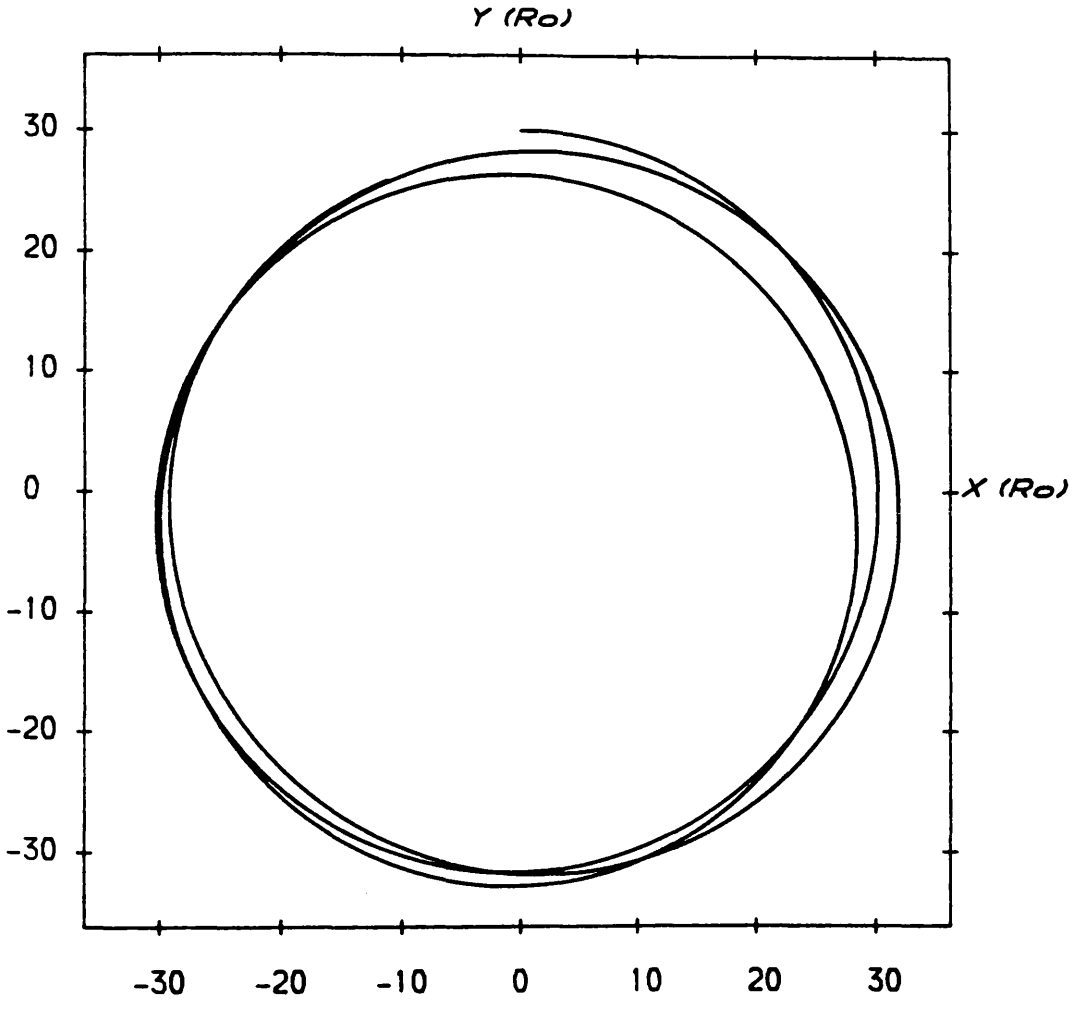
The other major disturbance to the nominal halo orbit, for the geocentric case, is that of lunar perturbations. Given that the lunar mean sphere of influence is of order 10 R_o lunar perturbations will be small for halo orbits with a geocentric distance of less than 40-50 R_o. It is found in general that lunar perturbations are approximately an order of magnitude less than those due to the rotating Sun-line.

5.8 Geocentric Halo Orbit Control

The stability analysis of section 5.6 has shown that unstable families of geocentric halo orbits exist. Furthermore, for the stable families the Sun-line rotation induces periodic oscillations about the nominal orbit. It is therefore necessary to develop simple control schemes to stabilize the unstable families and to damp the sail response to the Sun-line rotation.

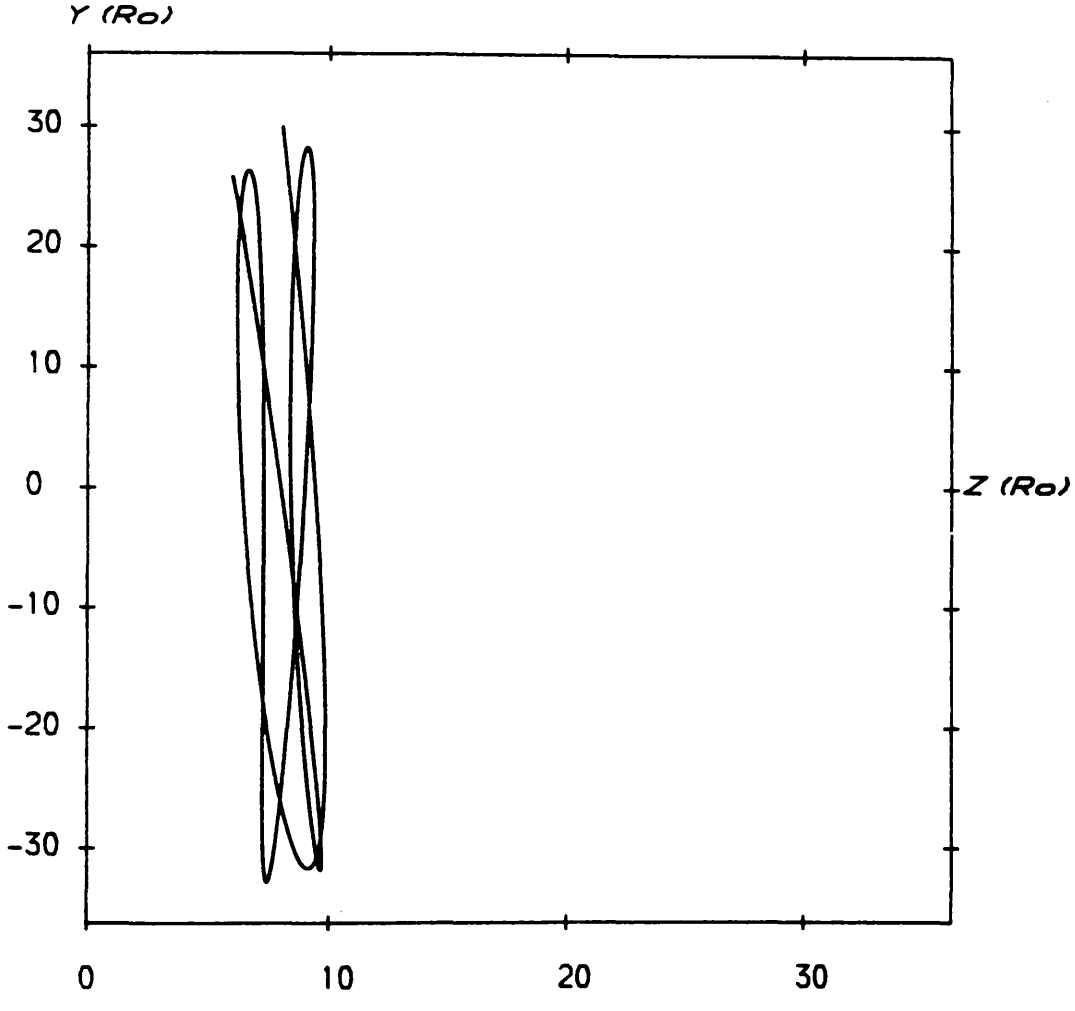
As in section 4.7 the reduced two-dimensional system will be used so that, in the first instance, the azimuthal motion of the spacecraft will be ignored. It will be shown that the system is controllable using a feedback to the sail pitch, but that the associated damping timescale is too long. Therefore, a well damped control using

Figure 5.7(a)



Perturbed response for a stable 10.13 day optimal halo orbit with ρ=30, z=8 and a rotating Sun-line; (a) x-y projection (b) y-z projection.

Figure 5.7(b)



variations in sail attitude and loading is developed.

5.8.1 Controllability by Sail Pitch

It will now be demonstrated that the reduced two-dimensional system is controllable using sail pitch alone. Making use of the variational equation (5.21) and allowing first order changes in sail attitude $\delta\mathbf{n}$ a modified variational equation is obtained

$$\frac{d^2\boldsymbol{\delta}}{dt^2} + \mathbf{M}_1 \frac{d\boldsymbol{\delta}}{dt} + \mathbf{M}\,\boldsymbol{\delta} = \mathbf{K}\,\delta\mathbf{n} \quad , \quad \mathbf{K} = \frac{\partial \mathbf{a}}{\partial \mathbf{n}}\Big|_{\mathbf{r}=\mathbf{r}_o,\ \mathbf{n}=\mathbf{n}_o} \tag{5.49}$$

The variational equation may again be reduced to the variables $\boldsymbol{\delta}'=(\xi',\eta')$ by eliminating the azimuthal coordinate, viz

$$\frac{d^2\boldsymbol{\delta}'}{dt^2} + \mathbf{M}\,\boldsymbol{\delta}' = \mathbf{K}\,\delta\alpha \quad , \quad \mathbf{M} = \begin{bmatrix} M^*_{11} & \vdots & M_{13} \\ \cdots & \cdots & \cdots \\ M_{31} & \vdots & M_{33} \end{bmatrix} \quad , \quad \mathbf{K} = \begin{bmatrix} \frac{\partial a_\rho}{\partial \alpha} \\ \frac{\partial a_z}{\partial \alpha} \end{bmatrix} \tag{5.50}$$

where the general first order attitude change $\delta\mathbf{n}$ now becomes the change in Sun-sail pitch angle $\delta\alpha$. The solar radiation pressure acceleration partial derivatives $\mathbf{K}=(K_1,K_2)^T$ are given by

$$K_1 = \beta\cos^3\alpha\,(1 - 2\tan^2\alpha) \tag{5.51a}$$

$$K_2 = -3\beta\cos^2\alpha\,\sin\alpha \tag{5.51b}$$

Writing the system in standard state variable form $\mathbf{x}=(\boldsymbol{\delta}',d\boldsymbol{\delta}'/dt)$ a set of four first order equations are obtained

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}^*\mathbf{x} + \mathbf{K}^*\delta\alpha \quad , \quad \mathbf{M}^* = \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{I} \\ \cdots & \cdots & \cdots \\ -\mathbf{M} & \vdots & \mathbf{0} \end{bmatrix} \quad , \quad \mathbf{K}^* = \begin{bmatrix} \mathbf{0} \\ \\ \mathbf{K} \end{bmatrix} \tag{5.52}$$

The controllability matrix $\mathbf{C}=(\mathbf{K}^*,\mathbf{M}^*\mathbf{K}^*,\mathbf{M}^{*2}\mathbf{K}^*,\mathbf{M}^{*3}\mathbf{K}^*)$ formed from the system and input distribution matrices then becomes

$$\mathbf{C} = \begin{Bmatrix} 0 & K_1 & 0 & -M^*_{11}K_1 - M_{33}K_2 \\ 0 & K_2 & 0 & -M_{31}K_1 - M_{33}K_2 \\ K_1 & 0 & -M^*_{11}K_1 - M_{33}K_2 & 0 \\ K_2 & 0 & -M_{31}K_1 - M_{33}K_2 & 0 \end{Bmatrix} \tag{5.53}$$

For $\mathbf{C}$ to have full rank it is required that $\det\mathbf{C}\neq 0$. The matrix determinant is then given by

$$\det\mathbf{C} = K_1K_2(M^*_{11} - M_{33}) - (K_1^2 - K_2^2)M_{13} \tag{5.54}$$

For the optimal halo mode equation (5.54) reduces to $\det\mathbf{C}=3\beta^2\rho z/r^5$. Therefore the statite type solutions with $\rho=0$ are uncontrollable using the sail pitch alone. However, all halo orbit modes are in principle controllable using a feedback control to the Sun-sail pitch angle.

5.8.2 Control by Variable Sail Pitch

Since it has now been demonstrated that the geocentric halo orbit modes are all controllable using variable sail pitch a closed loop feedback control will now be investigated. A general expression for the feedback will be used, viz

$$\delta\alpha = \sum_{j=1}^{4} g_j x_j \quad , \quad \mathbf{x} = (x_j) \quad (j=1,4) \tag{5.55}$$

where the gains (g_j), $(j=1,4)$ are chosen to ensure that all four of the system eigenvalues are in the left hand complex plane so that the system has asymptotic stability. Constraints on the gains are found

using the Routh-Hurwitz criterion.

The closed loop response of an unstable 15.51 day optimal halo orbit with a rotating Sun-line is shown in Figure 5.8. It can be seen that large variations in sail pitch angle are required to maintain the orbit and that there is a large amplitude response. Due to the variable transformations used in equations (5.25) the control scheme attempts to damp the system to $\xi'=\eta'=0$. Therefore there are constant residual displacements, as with the heliocentric case. These errors can however be removed through a simple open loop manoeuvre.

5.8.3 Open Loop Control

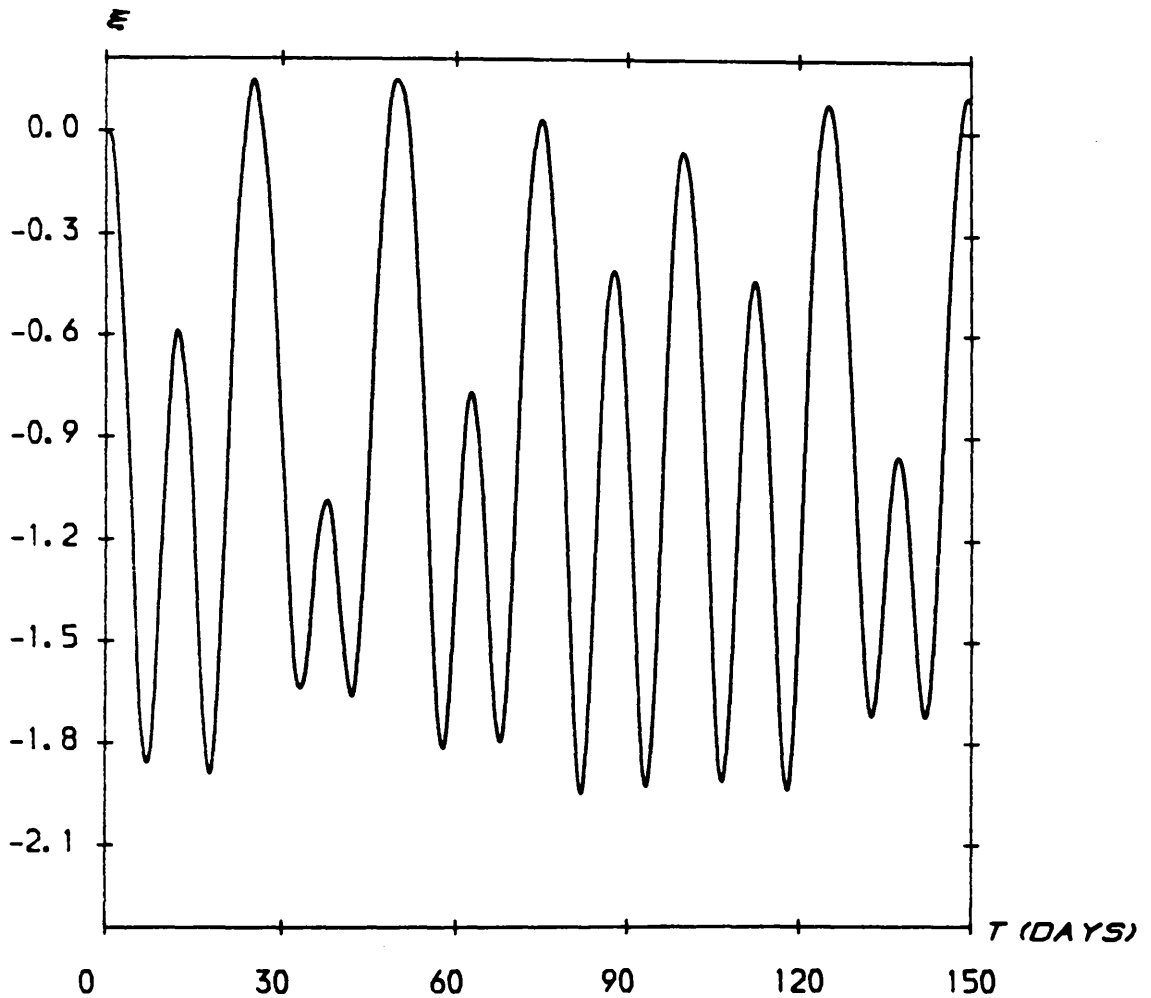
It was found in section 5.8.2 that with the pitch angle control the injection errors damped out to non-zero values and that there was an azimuthal drift in the sail position. Physically, this is due to the sail having excess azimuthal angular momentum at the injection point. Since the sail has pitch control only there is no means of removing this excess momentum. It will be shown now that the off-set and drift in sail position can in fact be removed through a simple open loop manoeuvre using a rotation about the sail yaw axis.

Considering the optimal halo orbit family with zero sail pitch angle and allowing first order changes in the sail yaw angle $\delta\chi$, the azimuthal component of the solar radiation pressure acceleration is obtained from the relation $(\partial a_\theta/\partial\chi)=\beta$. Therefore, the set of variational equations become

$$\frac{d^2\xi}{dt^2} - 2\Omega\rho_0\frac{d\Psi}{dt} + M_{11}\xi + M_{13}\eta = 0 \qquad (5.56a)$$

$$\frac{d^2\Psi}{dt^2} + \frac{2\Omega}{\rho_0}\frac{d\xi}{dt} = \frac{\beta}{\rho_0}\delta\chi \qquad (5.56b)$$

Figure 5.8(a)



Closed loop response for an unstable 15.51 day optimal halo orbit with ρ=10, z=40 and with a variable Sun-sail pitch angle control. The feedback gains are g_1=-0.6, g_2=-0.3, g_3=-0.6 and g_4=-0.3; (a) ξ response (b) η response (c) Ψ response (d) sail pitch angle control.

Figure 5.8(b)

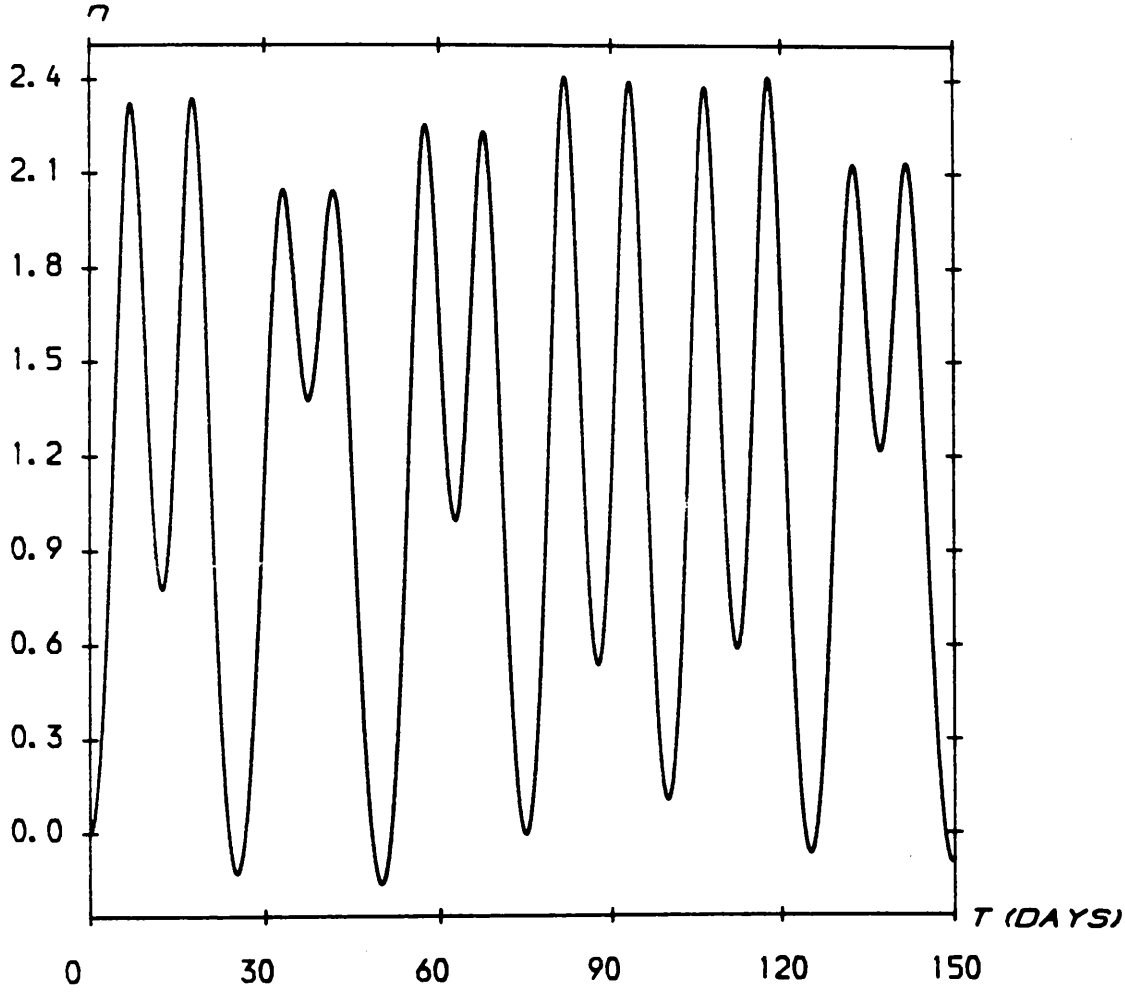


Figure 5.8(c)

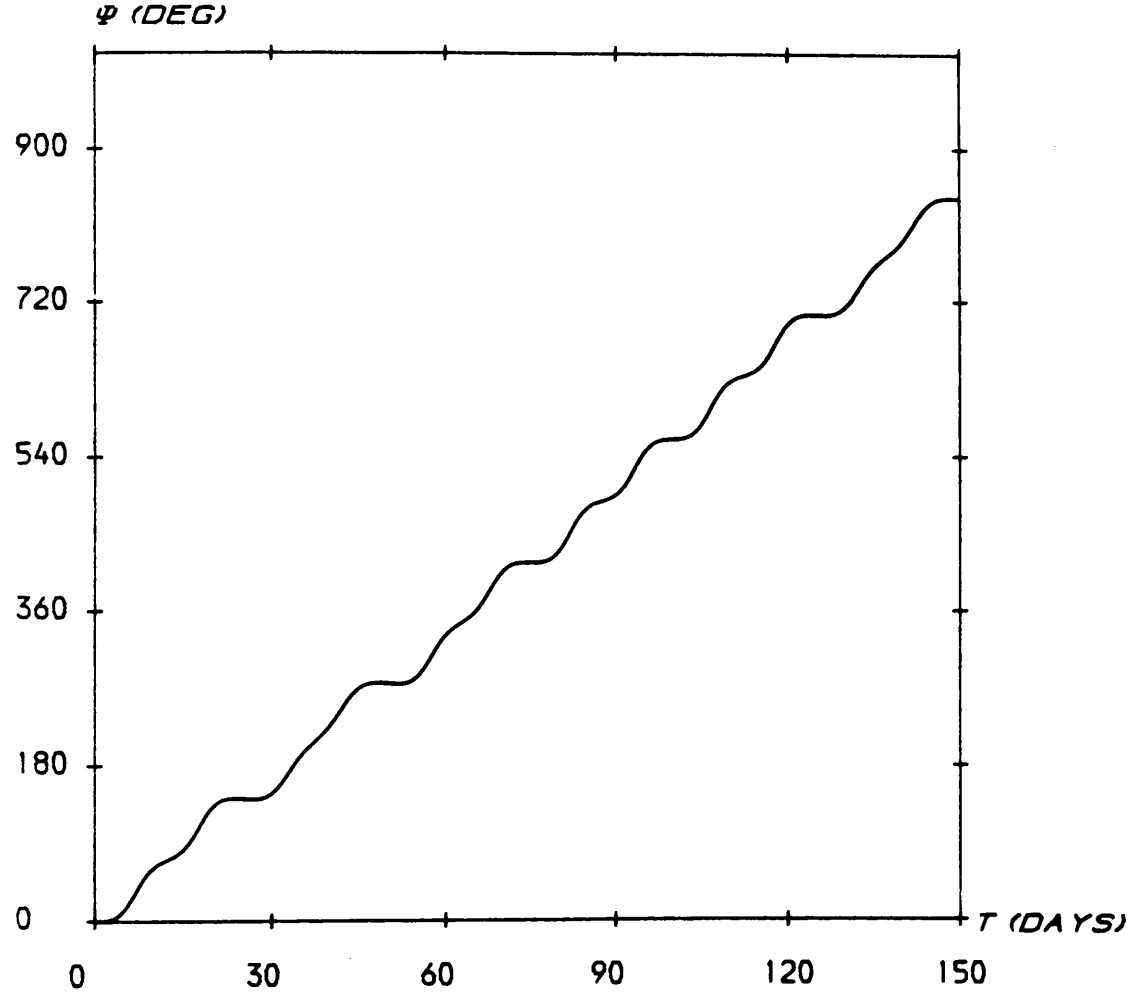
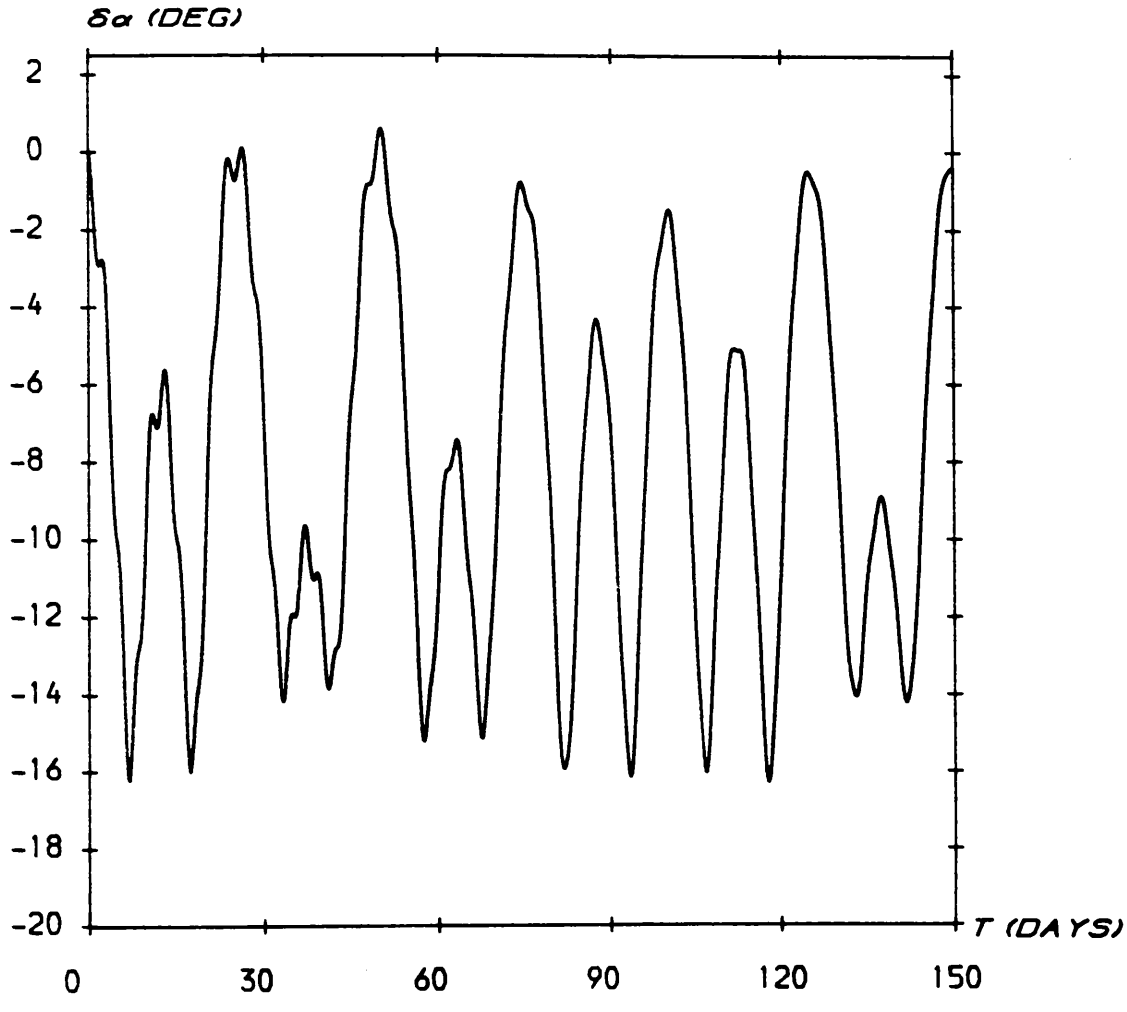


Figure 5.8(d)



$$\frac{d^2\eta}{dt^2} + M_{13}\xi + M_{33}\eta = 0 \qquad (5.56c)$$

Integrating equation (5.56b) the azimuthal drift rate of the sail may be written as

$$\frac{d\Psi}{dt} = \left\{\frac{d\Psi}{dt}\right\}_o - \left\{\frac{2\Omega}{\rho_o}\right\}(\xi(t) - \xi_o) + \frac{\beta t}{\rho_o}\varepsilon_X \qquad (5.57)$$

Therefore, $\xi(t)$ is required to obtain the full azimuthal time response. Substituting equation (5.57) in equation (5.56a) to eliminate the azimuthal term a set of variational equations in $\boldsymbol{\varepsilon}=(\xi,\eta)$ is obtained

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} + \mathbf{M}\,\boldsymbol{\varepsilon} = \boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 t \qquad (5.58)$$

where the vectors $\boldsymbol{\Gamma}_j=(\Gamma_{j1},\Gamma_{j2})$ $(j=1,2)$ are given by

$$\boldsymbol{\Gamma}_1 = \left\{2\Omega\rho_o\left\{\left\{\frac{d\Psi}{dt}\right\}_o + \left\{\frac{2\Omega}{\rho_o}\right\}\xi_o\right\},\ 0\right\}^T, \quad \boldsymbol{\Gamma}_2 = (2\Omega\beta\,\varepsilon_X,\ 0)^T \qquad (5.59)$$

Using the variable transformation $\boldsymbol{\varepsilon}'=\boldsymbol{\varepsilon}-\mathbf{M}^{-1}\boldsymbol{\Gamma}_1$ equation (5.58) may be reduced to

$$\frac{d^2\boldsymbol{\varepsilon}'}{dt^2} + \mathbf{M}\,\boldsymbol{\varepsilon}' = \boldsymbol{\Gamma}_2 t \qquad (5.60)$$

This equation may now be solved by standard Laplace transform methods. Defining the transform variable as s such that $L[f(t)]=F(s)$ the solution of equation (5.60) may be written as

$$\boldsymbol{\varepsilon}(t)' = L^{-1}\left\{\mathbf{D}(s)\left\{s\boldsymbol{\varepsilon}_o' + \left\{\frac{d\boldsymbol{\varepsilon}'}{dt}\right\}_o\right\}\right\} + L^{-1}\left\{\mathbf{D}(s)\,\frac{1}{s^2}\,\boldsymbol{\Gamma}_2\right\} \qquad (5.61)$$

where $\mathbf{D}(s)=(s^2\mathbf{I}+\mathbf{M})^{-1}$. Noting that $\det(s^2\mathbf{I}+\mathbf{M})$ is the characteristic polynomial of the system, the matrix $\mathbf{D}(s)$ may be written as

$$\mathbf{D}(s) = \text{Adj}(s^2 + \mathbf{I}) \left\{ \prod_{j=1}^{4} (s - \omega_j) \right\}^{-1} \tag{5.62}$$

where ω_j (j=1,4) are the four eigenvalues of the system defined by equation (5.33). The first term of equation (5.61) represents the open loop response of the system and the second term represents the response due to the sail yaw. For the stable, optimal halo orbit family the open loop response is of the form of periodic oscillations. The general solution for the response may be written in the physical variables (ξ,η) as

$$\xi(t) = \Gamma_{11} \frac{M_{33}}{M_{11}*M_{33}-M_{13}^2} + \Gamma_{21} \frac{M_{33}}{\omega_1^2\omega_2^2} t + \sum_{j=1}^{6} G_j P_j(\omega_1,\omega_2;t) \tag{5.63a}$$

$$\eta(t) = \Gamma_{11} \frac{M_{13}}{M_{11}*M_{33}-M_{13}^2} - \Gamma_{21} \frac{M_{13}}{\omega_1^2\omega_2^2} t + \sum_{j=1}^{6} H_j Q_j(\omega_1,\omega_2;t) \tag{5.63b}$$

where G_j, H_j (j=1,6) are constants and P_j, Q_j (j=1,6) are the short and long term periodic functions corresponding to the open loop response. The yaw response may now be used to remove the residual displacement terms in equations (5.63).

If the first two terms in equations (5.63) are now set to zero at time T then the sail yaw angle required is found to be

$$\delta\chi^* = - \frac{\rho_0}{\beta T} \left\{ \left\{ \frac{d\Psi}{dt} \right\}_0 + \left\{ \frac{2\Omega}{\rho_0} \right\} \xi_0 \right\} \tag{5.64}$$

Therefore, applying a yaw control angle $\delta\chi^*$ for time T will place the sail on the nominal halo orbit with periodic motion about the orbit. Substituting for $\xi(t)$ in equation (5.57) the azimuthal drift rate becomes

$$\frac{d\Psi}{dt} = \frac{\Gamma_{11}+\Gamma_{21}}{2\Omega\rho_o} + \frac{\Gamma_{11}M_{33}}{M_{11}*M_{33}-M_{13}^2} + \frac{\Gamma_{21}M_{33}}{\omega_1^2\omega_2^2}\,t + \sum_{j=1}^{6} G_j P_j(\omega_1,\omega_2;t) \quad (5.65)$$

As before, the sail yaw angle may be used to remove the azimuthal drift. If the first three terms in equation (5.65) are now set to zero at time T, it is found that the same yaw angle δx^* of equation (5.64) is required. Therefore, a single open loop control manoeuvre may be used to remove the off-sets along the ρ and z axes and the azimuthal drift.

The fact that the orbit off-sets and azimuthal drift are removed simultaneously with the one manoeuvre is related to the sail azimuthal angular momentum $h_\theta=\rho^2(d\theta/dt)$. The first order difference in azimuthal angular momentum between the nominal and perturbed halo orbit is given by

$$\delta h_\theta = \rho_o^2 \left\{\frac{d\Psi}{dt}\right\} + 2\Omega\rho_o\xi \quad (5.66)$$

Furthermore, the azimuthal acceleration $\beta\delta x$ due to the yaw manoeuvre generates a torque on the orbit $(dh_\theta/dt)=\rho_o\beta\delta x$. Therefore, the total azimuthal angular momentum change at time T is given by

$$\delta H_\theta = \delta h_\theta + \rho_o\beta\delta x T \quad (5.67)$$

To remove the excess angular momentum it is required that $\delta H_\theta=0$. Therefore, solving for the required sail yaw angle the expression δx^* given by equation (5.64) is obtained.

It has been shown then that the excess azimuthal orbital angular momentum due to the injection errors may be removed with a simple open loop yaw manoeuvre. This manoeuvre may be used initially to bring the sail to the nominal halo orbit before the pitch control

scheme is used for orbit maintenance. An extension of the scheme may also be used for the controlled, unstable halo orbit families.

5.8.4 Control by Variable Sail Attitude and Loading

As with the heliocentric halo orbits it has been found that pitch only control leads to long damping timescales. Therefore, a variable sail attitude and loading control is now developed. The variations in both the sail attitude and loading allows arbitrary control accelerations to be generated. Therefore, short damping timescales may be achieved.

Allowing for an arbitrary control acceleration Δ the full three-dimensional variational equation may be written as

$$\frac{d^2\delta}{dt^2} + M_1 \frac{d\delta}{dt} + M\,\delta = \Delta \quad , \quad \Delta = \Lambda_1\delta + \Lambda_2 \frac{d\delta}{dt} \tag{5.68}$$

where the gain matrices $\Lambda_{1,2}$ are chosen to ensure asymptotic stability. The variations in the sail attitude and loading required to generate the control acceleration may be obtained from equation (5.2), viz

$$(\beta+\delta\beta)\,(\mathbf{l}.(\mathbf{n}+\delta\mathbf{n}))^2\,(\mathbf{n}+\delta\mathbf{n}) = \beta\,(\mathbf{l}.\mathbf{n})^2\,\mathbf{n} + \Delta \tag{5.69}$$

Taking vector and scalar products of equation (5.69) and using the normalisation condition $|\mathbf{n}+\delta\mathbf{n}|=1$ the required sail attitude control is obtained as

$$\delta\mathbf{n} = \frac{\mathbf{a} + \Delta}{|\mathbf{a} + \Delta|} - \mathbf{n} \tag{5.70}$$

where the nominal radiation pressure acceleration $\mathbf{a}$ is given by equation (5.2). The attitude control $\delta\mathbf{n}$ can then be related to the

scalar pitch angle $\delta\alpha$ and yaw angle $\delta\chi$ controls. The sail loading control can also be obtained from equation (5.69) as

$$\delta\beta = \frac{|\mathbf{a} + \boldsymbol{\Delta}|}{(\mathbf{l}.(\mathbf{a}+\boldsymbol{\Delta}))^2} - \beta \qquad (5.71)$$

so that expressions for the sail attitude and loading control are obtained as a function of the control acceleration.

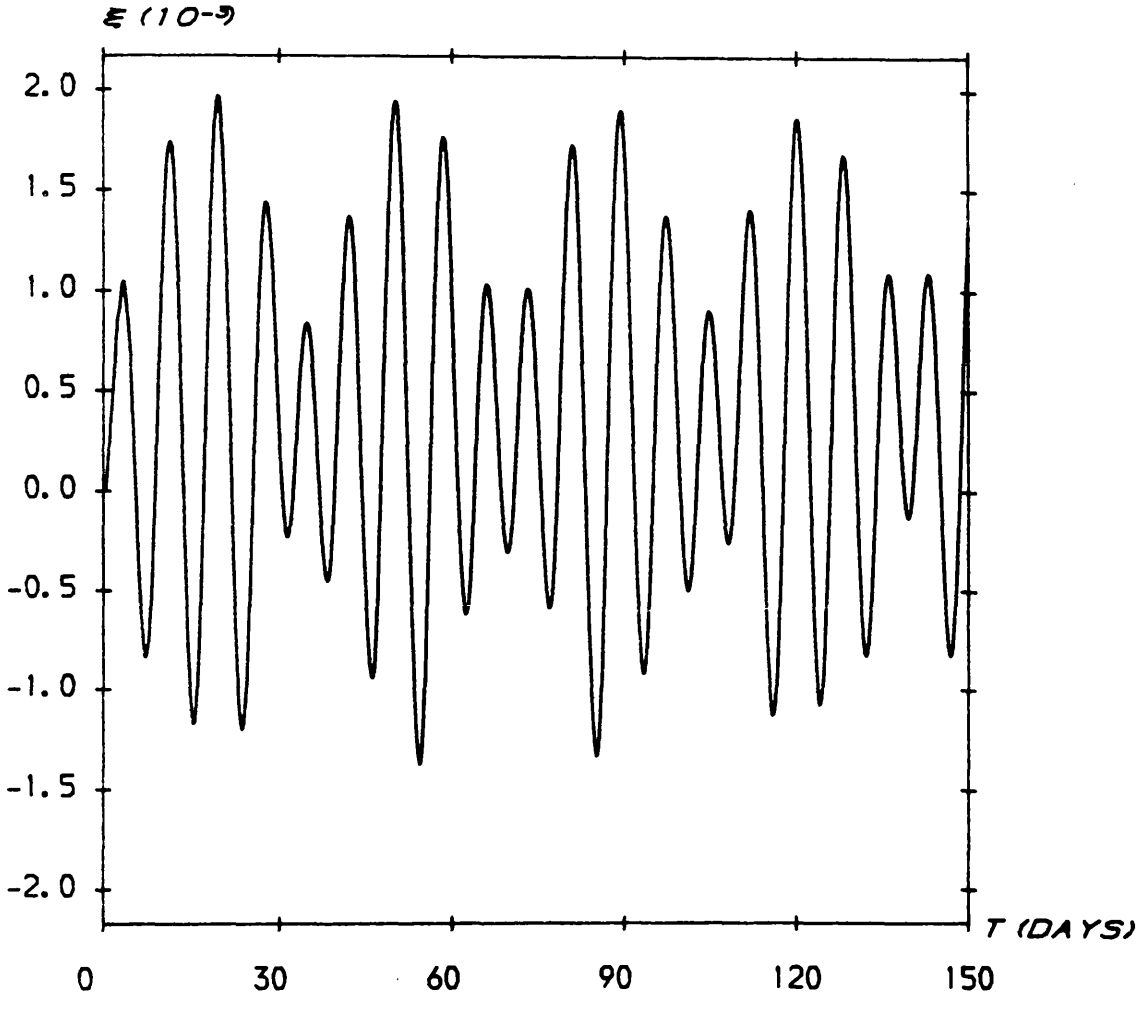
Using diagonal gain matricies $\Lambda_j=(\delta_{ji}g_i)$ (j=1,2, i=1,3) a suitably damped response may be obtained, as shown in Figure 5.9. It can be seen that with a maximum variation in the sail loading of 3×10^{-2} the amplitude of the response to the perturbations due to the rotating Sun-line are extremely small. A smaller amplitude response is possible but at the expense of larger variations in the sail loading.

5.9 Solution by the Hamilton-Jacobi Method

The dynamics of the optimal halo orbit problem will now be investigated through the use of Hamilton-Jacobi theory, Goldstein (1980). It will be demonstrated that a general closed analytic solution exists for the optimal case with **n=l**, or indeed the off-axis case with **n=k**. The dynamics are exactly equivalent to the Stark effect on the hydrogen atom when a uniform electric field is superimposed upon the Coulomb potential. The fact that the Hamilton-Jacobi equation of this system is separable can be demonstrated using the Staeckel conditions on the Hamiltonian when using a set of parabolic coordinates, Epstein (1916). The solution of the Stark problem is considered in detail in Born (1927).

Through the use of Hamilton-Jacobi theory constraints on the sail motion will be obtained by deriving the canonical momentum in terms

Figure 5.9(a)



Closed loop response for an unstable 15.51 day optimal halo orbit with ρ=10, z=40 and with sail pitch, yaw and loading control. The feedback gains are $g_{11}=g_{13}=-4\times10^{-4}$, $g_{21}=g_{23}=-3\times10^{-4}$, $g_{12}=-3\times10^{-3}$ and $g_{22}=-2\times10^{-3}$; (a) ξ ($\times10^{-3}$) response (b) η ($\times10^{-3}$) response (c) Ψ response (d) sail pitch angle control (e) sail yaw angle control (f) sail loading control.

Figure 5.9(b)

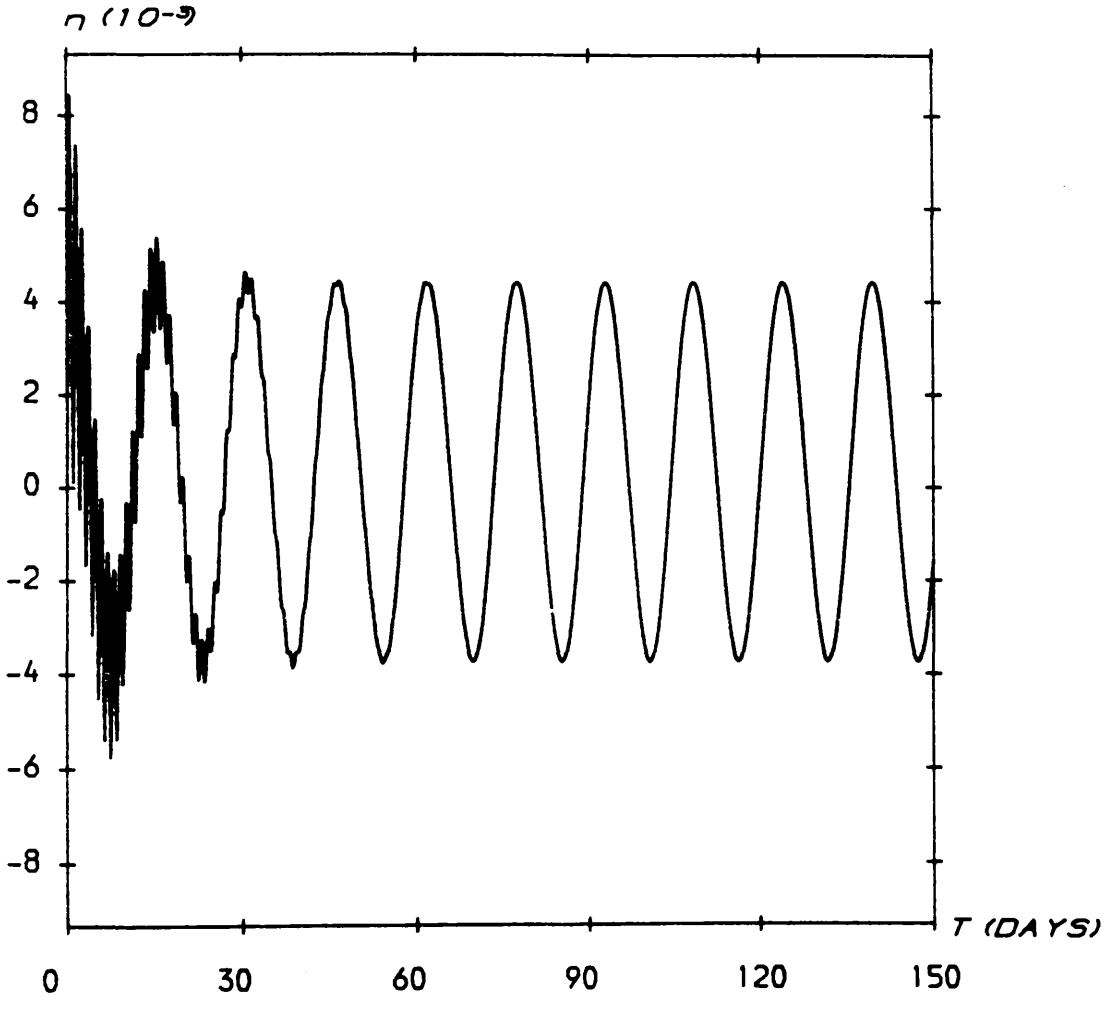


Figure 5.9(c)

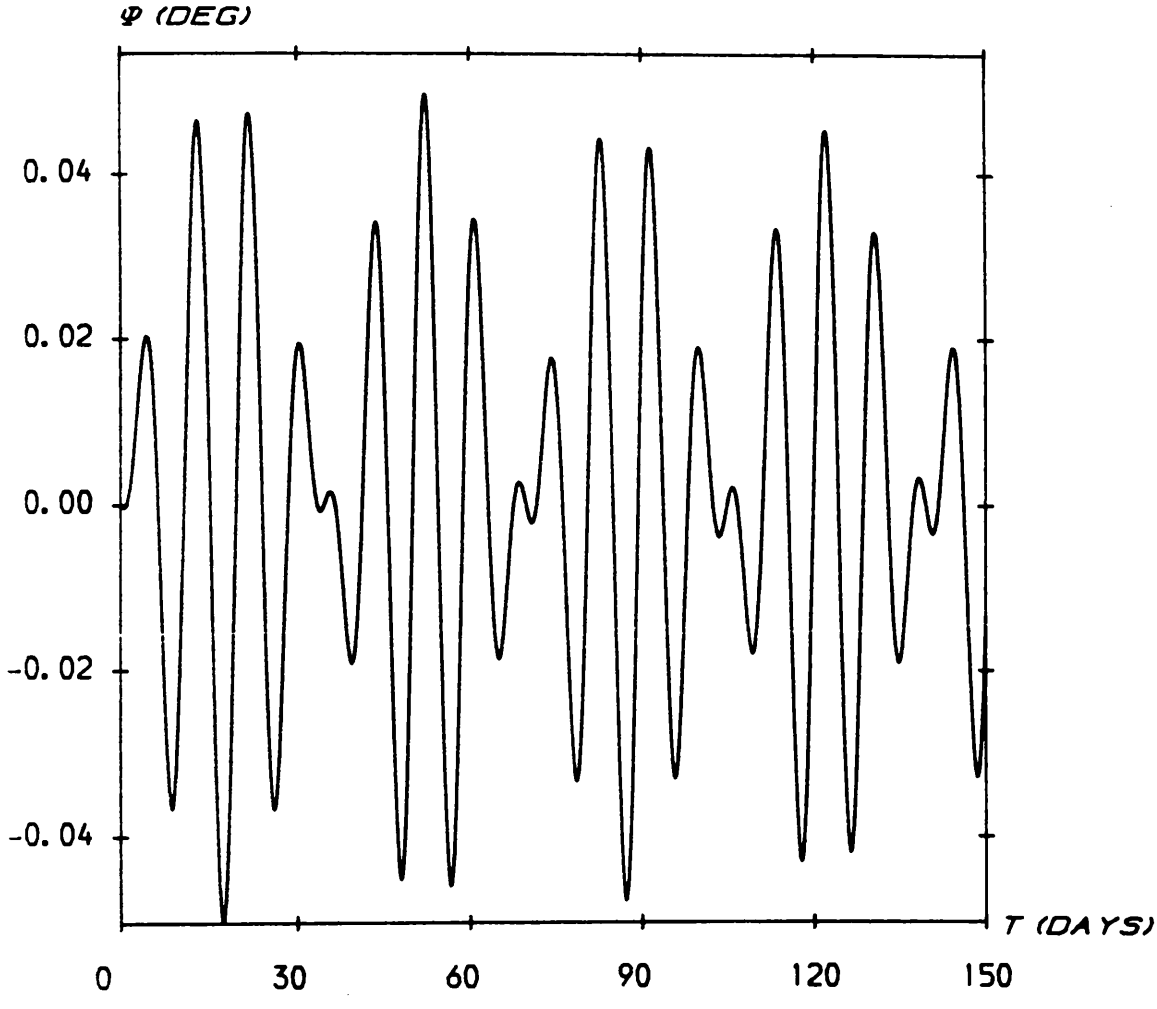


Figure 5.9(d)

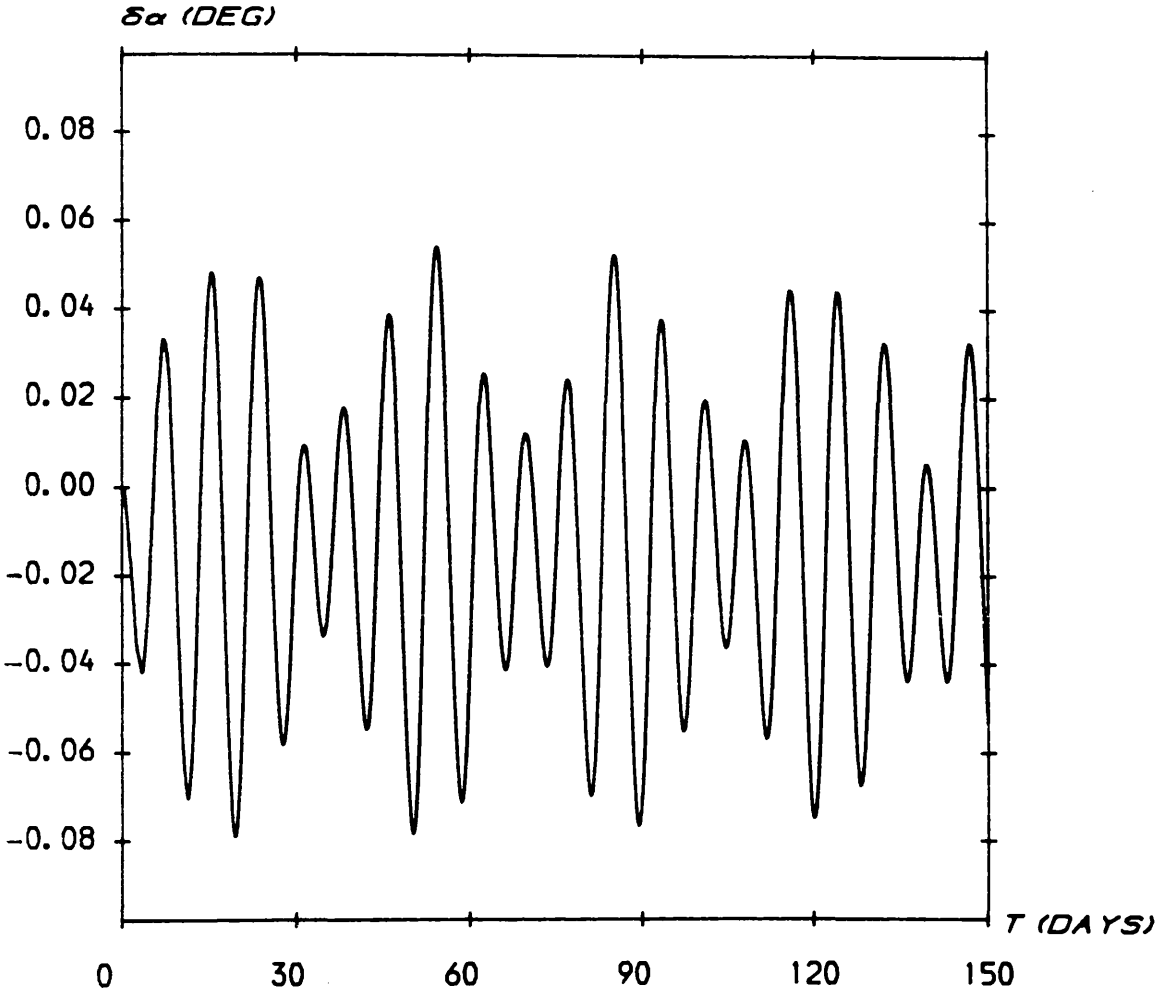

δα (DEG)
0.08
0.06
0.04
0.02
0.00
-0.02
-0.04
-0.06
-0.08
0
30
60
90
120
150
T (DAYS)

Figure 5.9(e)

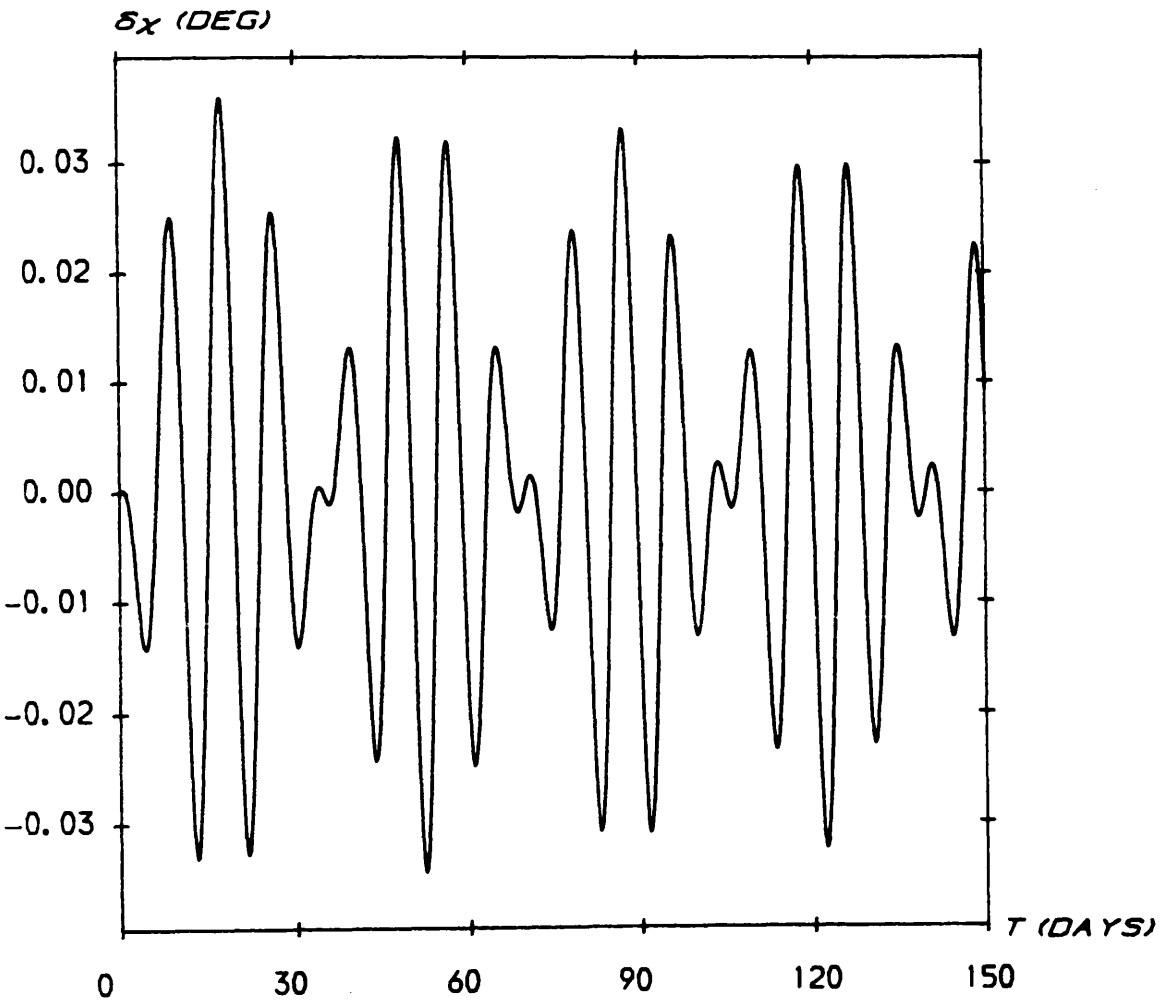
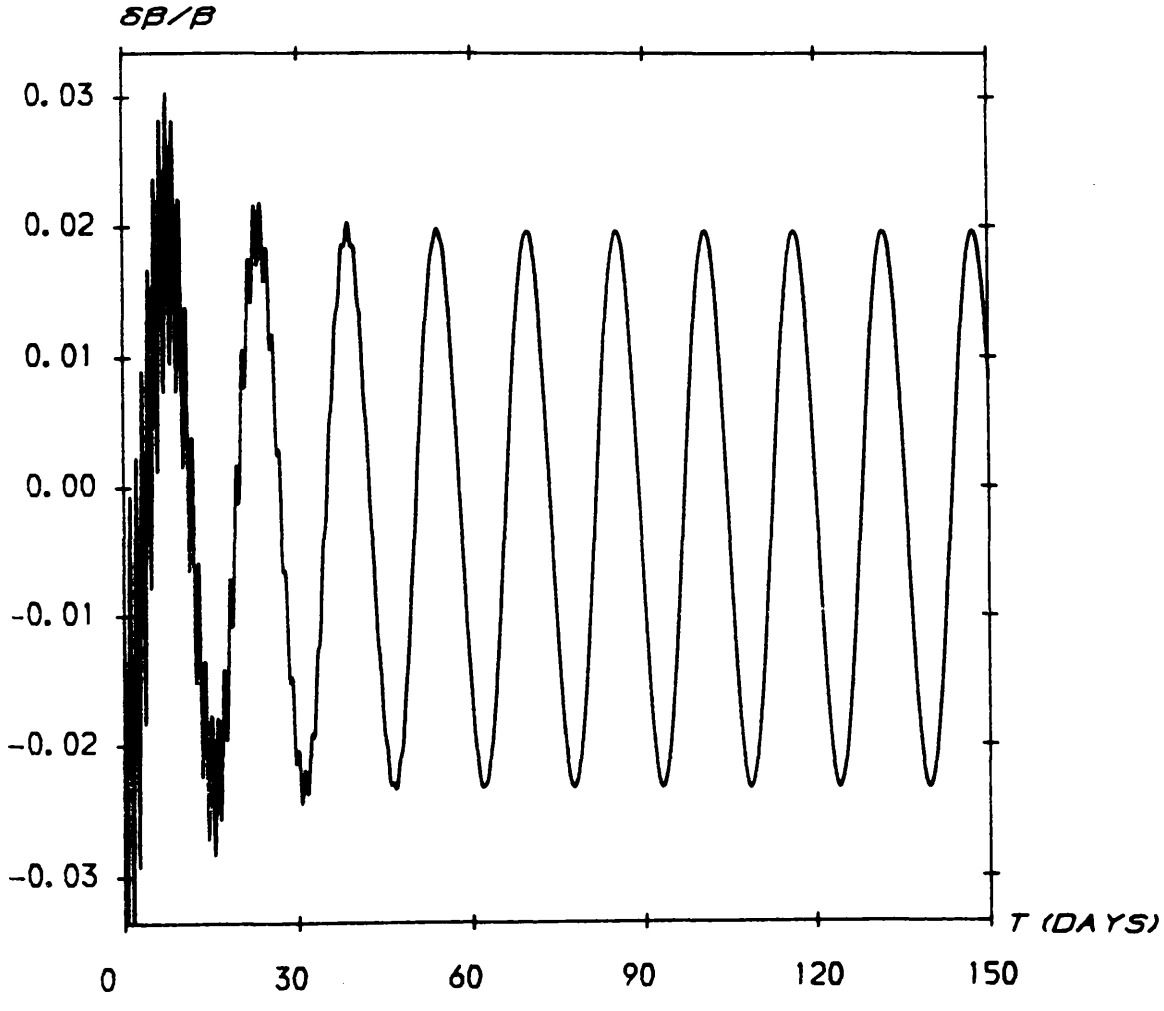


Figure 5.9(f)



of parabolic coordinates. These constraints can then be related to the existence of geocentric halo orbits. A set of parabolic coordinates (ξ,η,θ) defined as

$$x = \xi\eta\cos\theta \quad , \quad y = \xi\eta\sin\theta \quad , \quad z = \frac{1}{2}(\xi^2 - \eta^2) \tag{5.72}$$

will now be used where constant ξ and η coordinates define diametrically oriented parabolae of revolution about the z-axis, with θ the azimuthal angle. Using these new coordinates the kinetic energy of the system may be written as

$$T = \frac{1}{2}\left\{ \left\{\left\{\frac{d\xi}{dt}\right\}^2 + \left\{\frac{d\eta}{dt}\right\}^2\right\}\left\{\xi^2 + \eta^2\right\} + \left\{\frac{d\theta}{dt}\right\}^2\xi^2\eta^2 \right\} \tag{5.73}$$

so that the new conjugate momemta $p_j=\partial T/\partial x_j$ $(j=1,3)$ may be obtained as

$$P_\xi = \left\{\frac{d\xi}{dt}\right\}\left\{\xi^2 + \eta^2\right\} \quad , \quad P_\eta = \left\{\frac{d\eta}{dt}\right\}\left\{\xi^2 + \eta^2\right\} \quad , \quad P_\theta = \left\{\frac{d\theta}{dt}\right\}\xi^2\eta^2 \tag{5.74}$$

where P_θ is an integral of the system since the potential $\Gamma(\mathbf{r},\mathbf{n})+\Phi_2(|\mathbf{r}|)$ is independent of θ.

Equations (5.73) and (5.74) may now be used to construct the Hamiltonian of the system $H=T+\Gamma+\Phi_2$. In terms of the parabolic coordinates and conjugate momenta it is found that

$$H = \frac{1}{2}(\xi^2 + \eta^2)^{-1}\left\{P_\xi^2 + P_\eta^2 + \left\{\frac{1}{\xi^2} + \frac{1}{\eta^2}\right\}P_\theta^2 - \beta(\xi^4 - \eta^4) - 4\right\} \tag{5.75}$$

Defining the Hamilton-Jacobi function as S, the Hamilton-Jacobi equation may now be written as

$$\frac{\partial S}{\partial t} + \frac{1}{2}(\xi^2+\eta^2)^{-1}\left\{\left[\frac{\partial S}{\partial \xi}\right]^2 + \left[\frac{\partial S}{\partial \eta}\right]^2 + \left[\frac{1}{\xi^2} + \frac{1}{\eta^2}\right]\left[\frac{\partial S}{\partial \theta}\right]^2 - \beta(\xi^4-\eta^4) - 4\right\} = 0 \quad (5.76)$$

where $p_\xi=\partial S/\partial \xi$, $p_\eta=\partial S/\partial \eta$ and $p_\theta=\partial S/\partial \theta$. Since the Hamiltonian is independent of θ and t the Hamilton-Jacobi function becomes

$$S = \alpha_1 t + \alpha_3 \theta + S_*(\xi,\eta) \quad , \quad (5.77)$$

where α_1 may be identified with the total energy of the system and α_3 with the azimuthal momentum p_θ. If the separability property is now used and $S_*(\xi,\eta)=S_1(\xi)+S_2(\eta)$ then the Hamilton-Jacobi equation may be separated into

$$\left[\frac{\partial S_1}{\partial \xi}\right]^2 + 2\alpha_1\xi^2 + \alpha_3\xi^{-2} - \beta\xi^4 = -\alpha_2{}^2 \quad (5.78a)$$

$$\left[\frac{\partial S_2}{\partial \eta}\right]^2 + 2\alpha_1\eta^2 + \alpha_3\eta^{-2} + \beta\eta^4 - 4 = \alpha_2{}^2 \quad (5.78b)$$

where α_2 is the separation constant.

The momenta may now be written in terms of the parabolic coordinates and the constants of the motion, viz

$$P_\xi = F_1(\xi)^{1/2} \quad , \quad P_\eta = F_2(\eta)^{1/2} \quad (5.79)$$

where the functions F_1 and F_2 are defined as

$$F_1 = \frac{\beta}{\xi}\left\{\xi^6 - \frac{2\alpha_1}{\beta}\xi^4 - \frac{\alpha_2{}^2}{\beta}\xi^2 - \frac{\alpha_3{}^2}{\beta}\right\}^{1/2} \quad (5.80a)$$

and

$$F_2 = \frac{\beta}{\eta}\left\{-\eta^6 - \frac{2\alpha_1}{\beta}\eta^4 + \frac{\alpha_2{}^2+4\mu}{\beta}\eta^2 - \frac{\alpha_3{}^2}{\beta}\right\}^{1/2} \quad (5.80b)$$

The roots of these bicubic polynomials $\pm\xi_j$ (j=1,3) and $\pm\eta_j$ (j=1,3) then

define the region of space accessible to the sail.

For $\xi_1 \neq \xi_2$ and $\eta_1 \neq \eta_2$ the motion is periodic and is constrained to the interior of an annular ring defined by the intersections of the four paraboloids of revolution. Similarly for $\xi_1 \neq \xi_2$ and $\eta_1 = \eta_2$ the sail trajectory lies on the surface of a paraboloid of revolution defined by $\eta_1 = \eta_2$ and is constrained by the two paraboloids of revolution ξ_1 and ξ_2. Finally, the halo orbit case corresponds $\xi_1 = \xi_2$ and $\eta_1 = \eta_2$ so that the motion is constrained to a circle displaced in the anti-Sun direction. Therefore, the minimal loading halo orbits are in fact defined by the intersection of two paraboloids of revolution.

The problem can be solved fully by obtaining the parabolic coordinates through the inversion of elliptic quadratures, Isayev and Kunitsyn (1972). Defining a new time variable τ such that, $dt=(\xi+\eta)d\tau$, it may be shown that a closed solution is obtained in terms of Jacobian elliptic functions sn, cn as

$$\xi(\tau) = X_{11} + X_{12}\,sn^2\left\{X_{13},\ X_{14}(\tau-\tau_0)\right\} \tag{5.81a}$$

$$\eta(\tau) = X_{21} + X_{22}\,cn^2\left\{X_{23},\ X_{24}(\tau-\tau_0)\right\} \tag{5.81b}$$

where $X_{i,j}(i=1,2,\ j=1,4)$ and τ_0 are constants of the system defined by the canonical constants α_j $(j=1,3)$. For a small solar radiation pressure acceleration the elliptic functions may be expanded in rapidly convergent power series. However, this technique is not available for the geocentric halo orbits as the solar radiation pressure acceleration is of the same order as the local gravitational acceleration.

5.10 Patched Geocentric Halo Orbits

Now that the stability and control of the fundamental modes of

operation of geocentric halo orbits have been established the patching of halo orbits to form complex and elaborate new trajectories will be investigated. The patching process will be carried out by transferring from one halo orbit to another through a simple, assumed instantaneous, switching operation on the sail attitude. At each of these switching points a number of boundary conditions must be satisfied to make the transfer possible. The conditions to be met are

(i) $r_1=r_2$; Intersection of the halo orbits.

(ii) $\mathbf{v}_1=\mathbf{v}_2$; No velocity impulse required at the switching point.

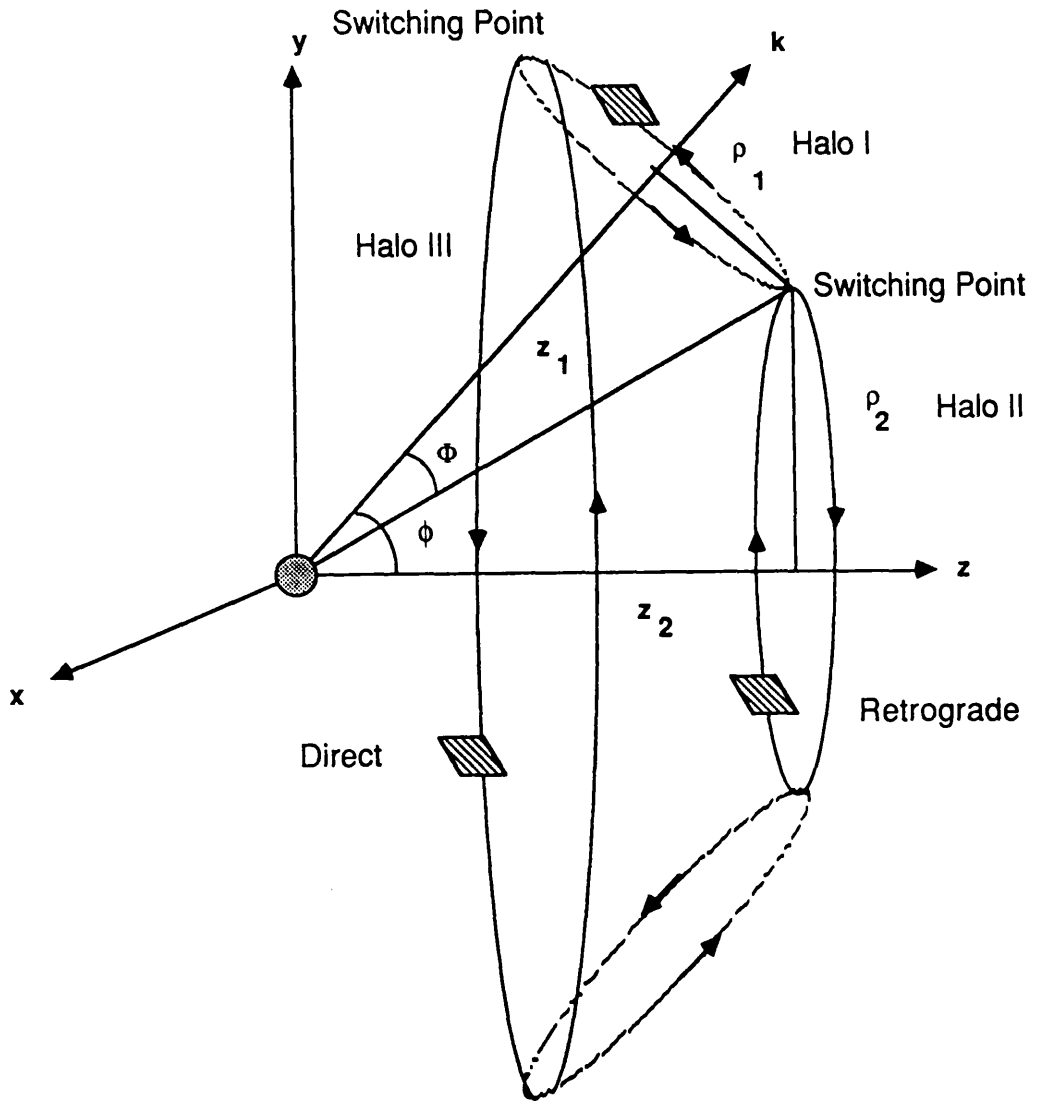
(iii) $E_1=E_2$; Sail energy is continuous across the switching operation.

(iv) $\beta_1=\beta_2$; Sail loading is continuous across the switching operation.

Condition (i) ensures that the transfer is possible while conditions (ii), (iii), (iv) ensure that no other operation other than the switching of the sail attitude is required. As a consequence of condition (ii) the halo orbits must intersect tangentially. It is clear then that the spacecraft can only transfer to or from an off-axis halo orbit (I) to an on-axis halo orbit (II), or between two off-axis halo orbits. The transfer may occur at an upper or lower point on the off-axis halo orbit, Figure 5.10.

Since the switching operation is assumed to take place instantaneously there is no change in the spacecraft gravitational potential across the switching operation. Furthermore, since the radiation field potential is non-conservative with respect to rotations of the sail (ie. no work is done against the field by re-orienting the sail), condition (iii) reduces to ensuring that $|\mathbf{v}_1|=|\mathbf{v}_2|$, (ie. $\rho_1\Omega_1=\rho_2\Omega_2$). This condition therefore determines the period of halo

Figure 5.10



Patched trajectories formed from off-axis and on-axis halo orbits. The switching of the sail attitude occurs at the intersection points.

orbit II as $\Omega_2=\Omega_1(\rho_1/\rho_2)$.

5.10.1 Upper Transfer

Firstly the upper switching point will be considered with a transfer from the off-axis to on-axis halo orbit. Owing to the dynamic reversibility of the problem the transfer can of course take place from the on-axis to off-axis halo orbit. The sail loadings required for each halo orbit are given by equations (5.13b) and (5.10b) as, using conditions (i-iii)

$$\beta_1 = \frac{z_1}{r^3}\cos^{-2}\Phi \quad , \quad \Omega_1 = \Omega_* \quad , \quad \alpha = \Phi \tag{5.82a}$$

$$\beta_2 = \frac{z_2}{r^3}\left\{1 + \left\{\frac{\rho_2}{z_2}\right\}\left\{1 - \left\{\frac{\Omega_2}{\Omega_*}\right\}^2\right\}^2\right\}^{3/2}, \quad \Omega_2 = \Omega_*\left\{\frac{\rho_1}{\rho_2}\right\} \tag{5.82b}$$

with the sail pitch angle on halo orbit II given by equation (5.10a). The sail loadings can be equated using condition (iv) to obtain a function $\beta_1-\beta_2=0$. This function can be reduced to two variables by eliminating (ρ_2,z_2), which are related to (ρ_1,z_1) through a rotation

$$\begin{Bmatrix} \rho_2 \\ z_2 \end{Bmatrix} = \begin{bmatrix} \cos\Phi & \sin\Phi \\ -\sin\Phi & \cos\Phi \end{bmatrix} \begin{Bmatrix} \rho_1 \\ z_1 \end{Bmatrix} \tag{5.83}$$

The resulting condition for a patched, equal loading trajectory then becomes $F_1(\Phi,\Phi)=0$, where

$$F_1(\Phi,\Phi) = \cos^2\Phi\left\{1 + \tan^2(\Phi+\Phi)\left\{1 - \cos^{-2}\Phi\left\{1 + \frac{\tan\Phi}{\tan\Phi}\right\}^{-2}\right\}^2\right\} \tag{5.84}$$

$$- \left\{1 - \tan\Phi\tan\Phi\right\}^{-2/3} \quad , \quad \tan\Phi = \frac{\rho_1}{z_1}$$

The numerical solution to this equation is shown in Figure 5.11 as C_1 in the (Φ,Φ)-plane. It can be seen that $\Phi+\Phi \leqslant (\pi/2)$ so that halo orbit II always lies on the +z (night-side) of the planet, as is required for the solution to be physical.

It should be noted that the system is rotationally symmetric about the z-axis so that the sail may transfer from halo orbit II to halo orbit I at any azimuthal position.

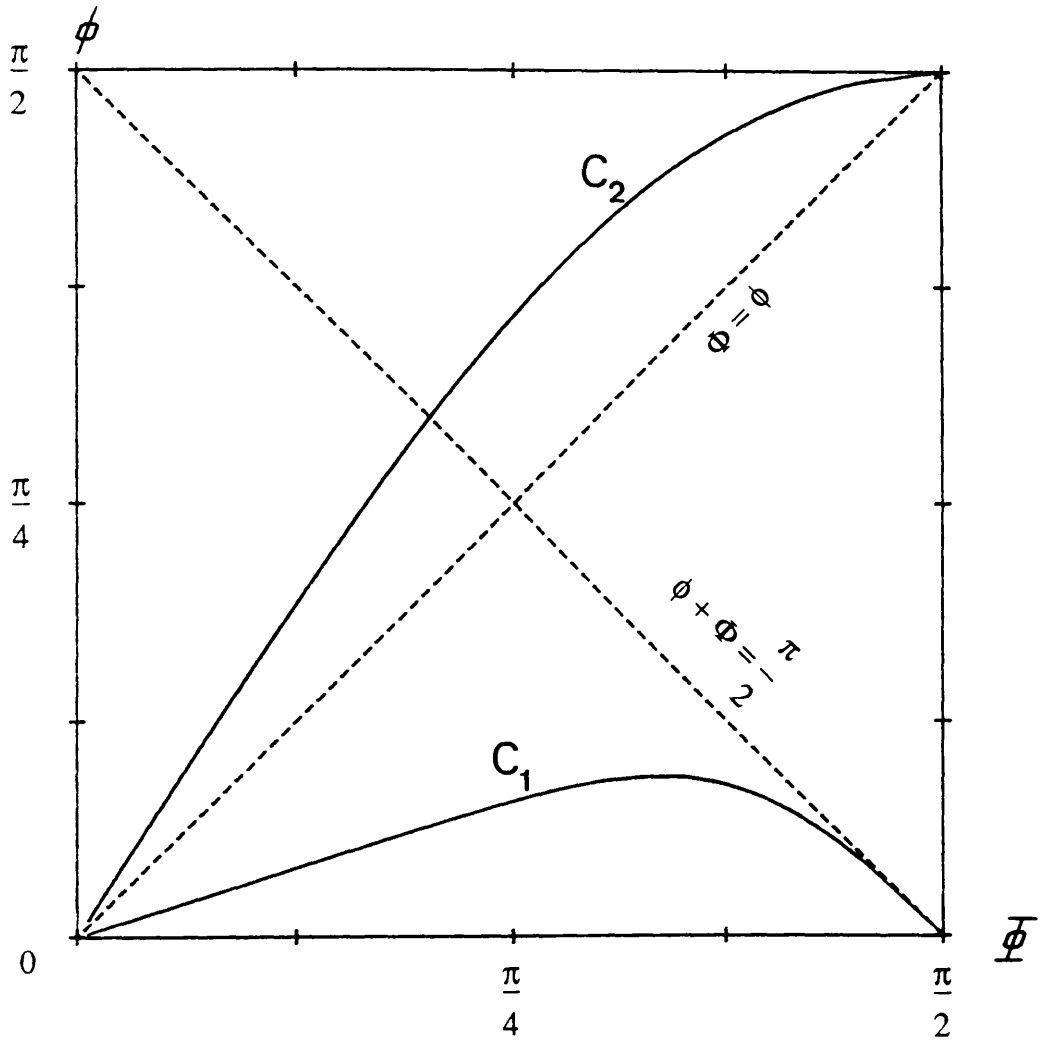
5.10.2 Lower Transfer

The analysis for the lower switching point is similar to that of the upper point with the resulting condition $F_2(\Phi,\Phi)=F_1(\Phi,-\Phi)=0$. The numerical solution is shown as C_2 in Figure 5.11. The sail may transfer at will to and from halo orbit I as it moves along halo orbit III, with halo orbit III being retrograde with respect to halo orbit II. It can be seen that there is no intersection of the curves in Figure 5.11 so that it is not possible, with the imposed boundary conditions, to transfer at both the upper and lower switching points from the same initial off-axis halo orbit I.

5.10.3 Azimuthal Transfer

The possibility of patching off-axis halo orbits together will now be investigated. Consideration of the geometry shows that for condition (iv) to hold the off-axis angles Φ must be equal. Then, for conditions (i-iii) to hold, it is clear that the sail may only transfer onto neighbouring off-axis halo orbits of the same amplitude and displacement. Therefore, by patching together these azimuthally neighbouring halo orbits with on-axis halo orbits, extremely complex patched halo trajectories may be constructed.

Figure 5.11



Conditions for patching off-axis to on-axis halo orbits. The curve C_1 represents the conditions for a transfer from the upper point of an off-axis halo orbit and C_2 represents the condition for the lower transfer.

5.10.4 Keplerian Transfer

The sail attitude may also be switched into a null orientation with $\alpha=\pi/2$, so that the sail is transferred onto a Keplerian orbit. This Keplerian orbit may be circular or elliptical depending on the initial halo orbit.

Off-Axis Transfer Firstly a transfer to or from an off-axis halo orbit I to a Keplerian ellipse will be considered (it may be shown that condition (iii) cannot be satisfied for transfer to a circular orbit). The switching point may be at either the apogee or perigee of the ellipse and at any point on the halo orbit, Figure 5.12. At these switching points the Keplerian perigee and apogee velocities are given by

$$v_p = \left\{ \frac{1}{a} \left\{ \frac{1+e}{1-e} \right\} \right\}^{1/2}, \quad v_a = \left\{ \frac{1}{a} \left\{ \frac{1-e}{1+e} \right\} \right\}^{1/2} \tag{5.85}$$

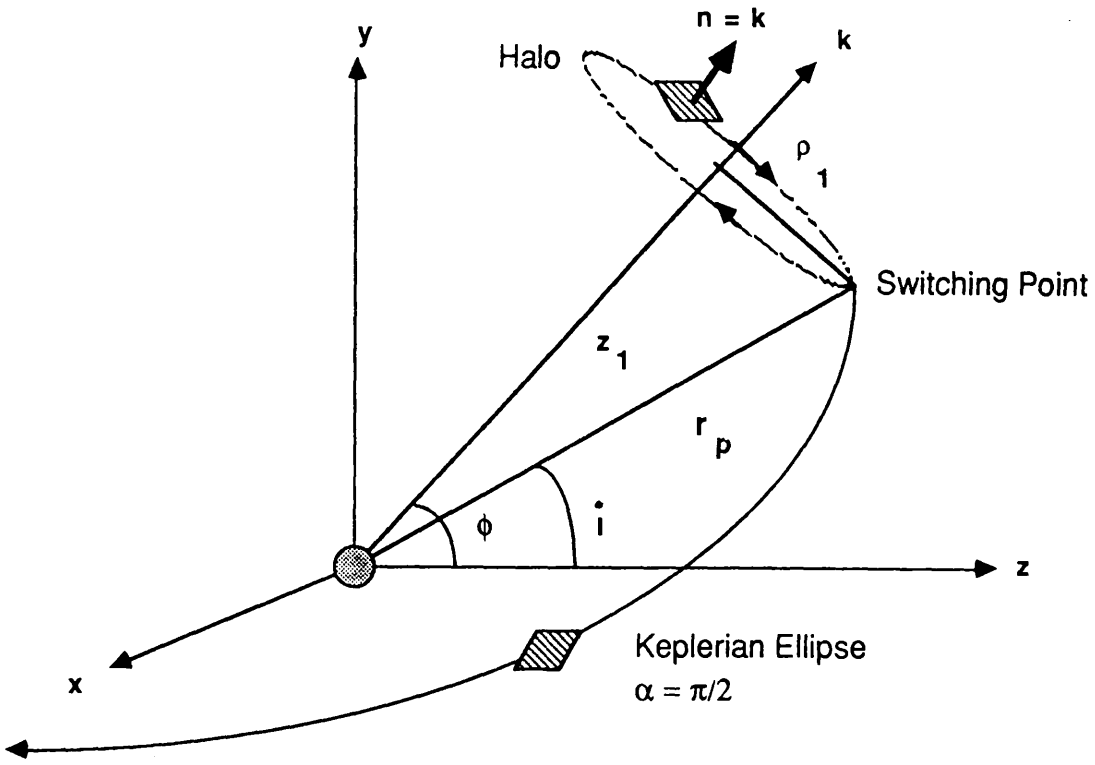
where a is the semi-major axis and e is the eccentricity of the ellipse. If a transfer to or from the perigee is considered then for condition (i) to be satisfied the sail orbital angular velocity on the off-axis halo orbit must be $\Omega_*=r_p^{-3/2}$, where $r_p=a(1-e)$ is the perigee distance. Therefore, equating v_p to the sail orbital velocity $\Omega_*\rho_1$ to satisfy condition (iii) the required halo amplitude is found to be

$$\rho_1 = a(1-e)(1+e)^{1/2} \tag{5.86}$$

Furthermore, since $\sin(\Phi-i)=\rho_1/r_p$ the required orientation of the off-axis halo orbit is obtained as

$$\Phi = i + \sin^{-1}(1+e)^{1/2} \tag{5.87}$$

Figure 5.12



Patching of an off-axis halo orbit to a Keplerian ellipse. The transfer may occur at either the perigee or apogee point and to or from an off-axis or on-axis halo orbit.

where i is the ecliptic inclination of the Keplerian ellipse. In particular the sail may be transferred from an off-axis halo orbit onto a Keplerian polar orbit. A similar set of expressions are obtained for an apogee transfer (ie. $\rho_1=\rho_1(-e)$, $\Phi=\Phi(-e)$).

On-Axis Transfer With this type of patched trajectory the spacecraft is transferred from an on-axis halo orbit I to a Keplerian orbit, which may in this case be circular. The circular transfer is possible since the sail angular velocity is now a free parameter, unlike the off-axis case, and can be chosen to satisfy the boundary conditions.

It can be seen from Figure 5.12 that for a perigee transfer the halo amplitude and displacement distances are given in terms of the orbital elements of the ellipse by $\rho_1=a(1-e)\sin(i)$, $z_1=a(1-e)\cos(i)$. Therefore to satisfy condition (iii) the required sail orbital angular velocity is given by

$$\Omega_1 = a^{-3/2} \sin^{-1}(i) \frac{(1+e)^{1/2}}{(1-e)^{3/2}} \qquad (5.88)$$

For an apogee transfer the required angular velocity is given by $\Omega_1(a,e,i)=\Omega_1(a,-e,i)$ and with $\Omega_1(a,0,i)$ the angular velocity for a transfer onto a circular orbit is obtained.

5.11 Conclusions

It has been shown in this chapter that geocentric halo type orbits are possible with solar sail spacecraft. As with the heliocentric halo orbits three distinct families of halo orbit exist with differing stability characteristics. The least demanding mode of operation is the optimal halo orbit with the sail orbital period chosen to be that of a

circular Keplerian orbit with a geocentric distance equal to that of the sail. With this orbital period the sail loading is minimized. Furthermore, with the optimal halo orbit family off-axis halo orbits are possible with the sail normal directed along the new halo axis.

Since it was found that unstable halo orbit families exist simple feedback control schemes have been developed. Although a feedback control to the sail pitch gives a damped response, the associated damping timescale is rather long. Therefore, a feedback to the sail loading has been included to achieve a well damped response.

Finally, it has been shown that individual halo orbits may be patched to form complex new trajectories. The patching may be between individual halo orbits or, by switching the sail attitude into a null orientation, a halo orbit and a Keplerian ellipse.

6. SOLAR SAIL PARKING IN RESTRICTED THREE-BODY SYSTEMS

6.1 Introduction

In this chapter stationary solutions for solar sails in the Earth-Sun and Earth-Moon restricted three-body dynamical systems are investigated. The classical circular restricted three-body problem has five well known stationary solutions L_j (j=1,5) where an infinitesimal mass will remain at rest with respect to the two primary masses of the system. It is found that in general the collinear points L_j (j=1,3) are unstable while the triangular points L_j (j=4,5) are stable in the Lyapunov sense, (see for example Roy (1982)). The classical restricted problem has been extended to include a radiation pressure force from either or both of the primary masses, exerted on the infinitesimal mass. This formulation generates four new additional stationary solutions with interesting stability characteristics. The radiation pressure force vector is however constrained to lie along the Sun-mass line.

For the Earth-Sun-sail three-body system the sail attitude may be freely oriented so that the solar radiation pressure force vector is not constrained to lie along the Sun-sail line. Furthermore, the magnitude of the solar radiation pressure force may be chosen through the spacecraft mass per unit area. Therefore, since certain parameters of the system can be arbitrarily chosen it is clear that rich new possibilites for artificial stationary solutions will arise. In fact it will be demonstrated that there is a continuum of new stationary solutions parameterised by the sail attitude and loading. The dynamical stability of these new stationary solutions is investigated and their instability established. Therefore, a control

scheme is developed whereby a feedback to the sail attitude is used to ensure the asymptotic stabilty of the solutions.

For the Earth-Moon three-body system the dynamics are no longer autonomous as the Sun-line rotates once per synodic month with respect to the Earth-Moon co-rotating reference frame. However, stationary solutions are again possible by utilising small trims in the sail loading parameter to compensate for the rotation of the Sun-line. Using this technique the spacecraft may be parked for a short duration in the Earth-Moon system. Furthermore, by linearising the dynamical equations it will be shown that an out-of-plane halo type trajectory is possible about the lunar L_2 point. These new three-body stationary solutions have potential applications for space science missions and solar sail parking, as will be discussed in chapter 7.

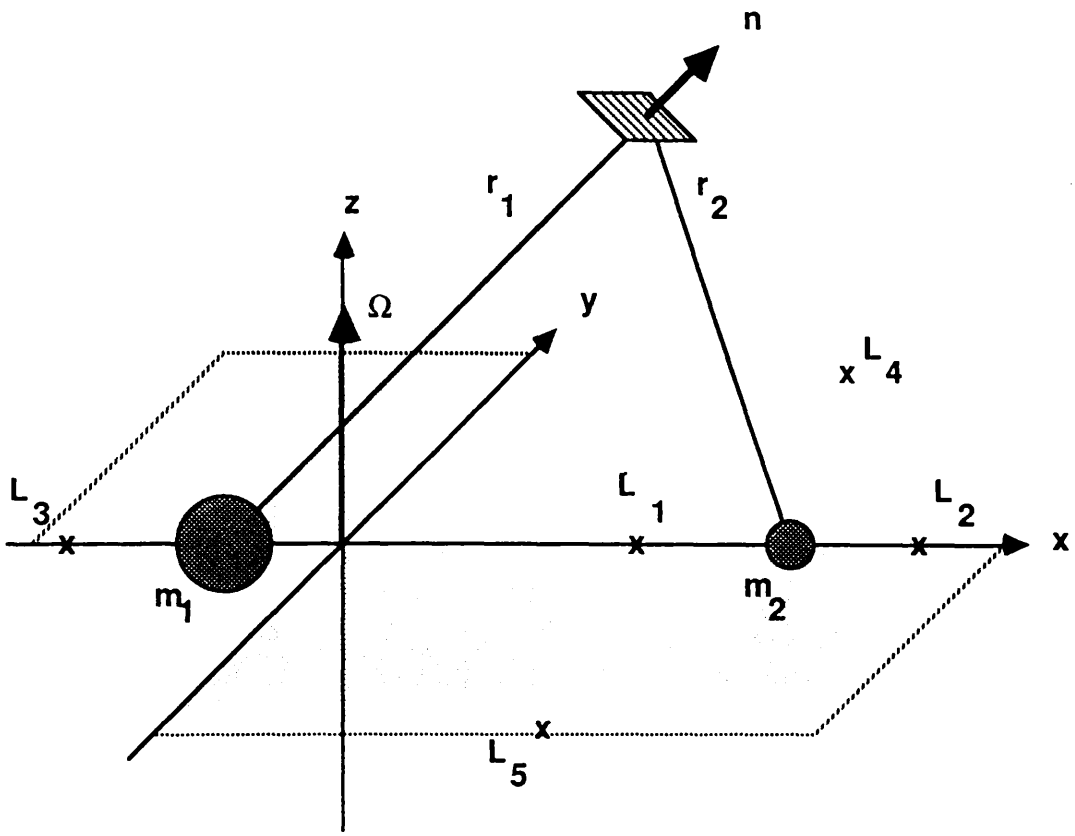
6.2 The Classical Restricted Three-Body Problem

The classical Lagrange stationary solutions to the restricted three-body problem L_j (j=1,5), have been studied in great depth since their discovery in 1772. More recently much attention has been given to the modified photogravitational problem, where one or both of the primary masses is a source of radiation pressure, the most complete work to date being that of Simmons et. al (1985). It has been shown that the five in-plane stationary solutions are modified by the existence of radiation pressure and that there exists four new out-of-plane solutions, L_j (j=6,9), with L_8 and L_9 existing only when both of the primary masses are luminous. Other investigations, such as Schuerman (1980), have included the relativistic Poynting-Robertson effect in the dynamical equations, which has bearing on the stability of the solutions.

The geometry of the circular restricted three-body problem with the five Lagrange points is shown schematically in Figure 6.1 for the Earth-Sun system. The mass ratio of the system is defined as $\mu=m_2/(m_1+m_2)$ and the Sun-Sail and Earth-Sail distances given by r_1 and r_2 respectively. The motion of m_1 and m_2 about their common centre of mass defines a co-rotating cartesian reference frame (x,y,z) with angular velocity Ω in which the dynamical equations are formulated. Dimensionless units are used so that the m_1-m_2 distance is unity. The classical stationary solutions are found at the equilateral points L_j (j=4,5) which are located at vertices of equilateral triangles defined by $r_1=r_2=1$ and the collinear points L_j (j=1,3) are given by the solution of a quintic polynomial. It may be shown that the equilateral points are Lyapunov stable provided $\mu<0.0385$ (Routh's value) and that the collinear points are always unstable. It is around the collinear L_1 and L_2 points that the new stationary solutions for solar sails are found to exist. Away from the Earth the solar sail solutions are similar to the Earth synchronous heliocentric halo orbits of chapter 4.

With the introduction to the classical case, of radiation pressure from m_1 the gravitational potential of m_1 is modified by the ratio of the radiation pressure force to the gravitational force β. This modification results in the classical Lagrange points moving to new positions as β increases, the triangular points $L_{4,5}$ now being defined by the constraints $r_1=(1-\beta)^{1/3}$, $r_2=1$. It can be seen that as $\beta\to1$ the triangular points will coalesce on m_1. It is also found that $L_{1,3}$ coalesce at m_1 with L_2 moving onto m_2. Furthermore, two new stationary solutions $L_{6,7}$ appear out of the ecliptic plane, moving asymptotically towards the z-axis with $z\to\infty$ as $\beta\to1$.

Figure 6.1



Schematic geometry of the Earth-Sun solar sail restricted three-body system with the sail positioned at $\mathbf{r}=(x,y,z)$. The sail attitude is defined by the unit vector $\mathbf{n}$ fixed in the co-rotating frame and the frame rotates with an angular velocity $\boldsymbol{\Omega}$. The centre of mass of m_1 and m_2 is located at the origin with m_1 located at $(-\mu,0,0)$ and m_2 located at $(1-\mu,0,0)$ where $\mu=m_2/(m_1+m_2)$.

6.3 Dynamical Equations for the Earth-Sun System

Consider now an idealised perfectly reflecting, planar solar sail in a co-rotating cartesian reference frame of angular velocity $\boldsymbol{\Omega}$ with two point primary masses m_1 and m_2, as shown in Figure 6.1. The sail attitude is defined by a unit vector $\mathbf{n}$, fixed in the co-rotating frame and the ratio of the solar radiation pressure force to the solar gravitational force exerted on the sail is again given by the sail loading parameter $\beta=\sigma^*/\sigma$. Since the sail attitude is to be fixed in the co-rotating frame the sail must rotate about the normal to the plane of the system once with respect to a fixed inertial frame in time $2\pi/|\boldsymbol{\Omega}|$. The units of the system will be chosen such that the gravitational constant, the distance between the two primary masses, the sum of the primary masses and so the angular velocity of co-rotation are all taken to be unity.

The vector dynamical equation for a solar sail in the co-rotating frame may be written in the usual form as

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) = \mathbf{a} - \nabla\Phi_3(\mathbf{r}) \qquad (6.1)$$

where the three-body gravitational potential $\Phi_3(\mathbf{r})$ and the non-conservative solar radiation pressure acceleration $\mathbf{a}$ are given by

$$\Phi_3(\mathbf{r}) = -\left\{\frac{1-\mu}{|\mathbf{r}_1|} + \frac{\mu}{|\mathbf{r}_2|}\right\} \quad , \quad \mathbf{a} = \beta \frac{1-\mu}{|\mathbf{r}_1|^4} (\mathbf{r}_1.\mathbf{n})^2 \mathbf{n} \qquad (6.2)$$

Although the system is non-conservative it is however autonomous owing to the co-rotation of the reference frame. The radiation solar pressure force vector can never be directed sunward so that the sail attitude is constrained such that $\mathbf{r}_1.\mathbf{n} \geqslant 0$. The sail position vectors are given with respect to the cartesian frame by

$$\mathbf{r}_1 = (x+\mu,\ y,\ z) \quad , \quad \mathbf{r}_2 = (x-(1-\mu),\ y,\ z) \tag{6.3}$$

where $\mu=m_2/(m_1+m_2)$ is the mass ratio of the system.

Since the centrifugal term in equation (6.1) is conservative it may again be written as a scalar potential $\Psi(\mathbf{r})$ such that

$$\nabla\Psi(\mathbf{r}) = \boldsymbol{\Omega} \times(\boldsymbol{\Omega} \times \mathbf{r}) \quad , \quad \Psi(\mathbf{r}) = -\frac{1}{2}\,|\boldsymbol{\Omega} \times \mathbf{r}|^2 \tag{6.4}$$

Defining a new potential, $U(\mathbf{r})=\Phi_3(\mathbf{r})+\Psi(\mathbf{r})$, a reduced dynamical equation is obtained, viz

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \nabla U(\mathbf{r}) = \mathbf{a} \tag{6.5}$$

In the co-rotating frame stationary solutions are required so that the first two terms of equation (6.5) vanish. The five classical stationary solutions $\mathbf{r}^{L}_{j}$ (j=1,5) are then given by the solutions to the equation $\nabla U(\mathbf{r})=0$. However, for the solar sail three-body system there exists an additional acceleration term $\mathbf{a}$ which is a function of the sail loading parameter β and attitude $\mathbf{n}$ so that new artificial stationary solutions may be generated.

Since the vector $\mathbf{a}$ is oriented in direction $\mathbf{n}$, taking the vector product of $\mathbf{n}$ with equation (6.5) it follows that

$$\nabla U(\mathbf{r}) \times \mathbf{n} = 0 \ \Rightarrow\ \mathbf{n} = \lambda \nabla U(\mathbf{r}) \tag{6.6}$$

where λ is an arbitrary scalar multiplier. Using the normalisation condition $|\mathbf{n}|=1$, λ is identified as $|\nabla U(\mathbf{r})|^{-1}$ so that the sail attitude required for a stationary solution is given by

$$\mathbf{n} = \frac{\nabla U(\mathbf{r})}{|\nabla U(\mathbf{r})|} \tag{6.7}$$

The sail attitude may then be expressed in terms of two angles (α,χ), defined with respect to the coordinate triad ($\hat{\mathbf{r}}_1$,$\hat{\mathbf{r}}_1 \times \boldsymbol{\Omega}$,$(\hat{\mathbf{r}}_1 \times \boldsymbol{\Omega}) \times \hat{\mathbf{r}}_1$) centred on the sail, where $\hat{\mathbf{r}}_1 = \mathbf{r}_1/|\mathbf{r}_1|$. The pitch angle α is defined as the angle of $\mathbf{n}$ with respect to $\hat{\mathbf{r}}_1$ and the clock angle χ is defined as the angle of the projection of $\mathbf{n}$ in the plane normal to $\hat{\mathbf{r}}_1$ with respect to $(\hat{\mathbf{r}}_1 \times \boldsymbol{\Omega})$. Therefore, taking vector and scalar products of equation (6.7) with $\hat{\mathbf{r}}_1$ these angles may be written as

$$\tan \alpha(\mathbf{r}) = \frac{|\mathbf{r}_1 \times \nabla U(\mathbf{r})|}{\mathbf{r}_1 \cdot \nabla U(\mathbf{r})} \tag{6.8a}$$

$$\tan \chi(\mathbf{r}) = \frac{|(\mathbf{r}_1 \times \boldsymbol{\Omega}) \times (\mathbf{r}_1 \times \nabla U(\mathbf{r}))|}{(\mathbf{r}_1 \times \boldsymbol{\Omega}) \cdot (\mathbf{r}_1 \times \nabla U(\mathbf{r}))} \tag{6.8b}$$

The sail loading required may also be obtained by taking a scalar product of equation (6.5) with $\mathbf{n}$. Requiring a stationary solution it is found that

$$\beta(\mathbf{r}) = (1-\mu)^{-1} \, |\mathbf{r}_1|^4 \, \frac{\nabla U(\mathbf{r}) \cdot \mathbf{n}}{(\mathbf{r}_1 \cdot \mathbf{n})^2} \tag{6.9}$$

Therefore, general vector valued functions for the sail attitude and loading required for stationary solutions have been obtained in terms of the co-rotating three-body potential $U(\mathbf{r})$. Since the sail loading and attitude may be chosen at will the set of five classical stationary solutions will be replaced by an infinite set of artificially generated stationary solutions. The classical solutions then correspond to the subset β=0. This infinite set of solutions is parametrised into level surfaces by the sail loading. A particular stationary solution on a level surface is then defined by the two

attitude angles (α,χ).

6.4 Existence of Stationary Solutions

Now that the existence of new stationary solutions has been established the regions in which these solutions may exist must be investigated. These regions are defined by the constraint

$$\mathbf{r}_1 . \nabla U(\mathbf{r}) \geqslant 0 \tag{6.10}$$

with the boundary surface defined by an equality in equation (6.10). This constraint may be understood physically since the solar radiation pressure acceleration vector **a**, and so the sail attitude vector **n**, can never be directed sunwards. The sail pitch angle is therefore constrained such that $|\alpha| < \pi/2$.

In scalar form the three-body potential $U(\mathbf{r})=\Phi_3(\mathbf{r})+\Psi(\mathbf{r})$ may be written as

$$U(\mathbf{r}) = - \left\{ \frac{1}{2} (x^2 + y^2) + \frac{1-\mu}{r_1} + \frac{\mu}{r_2} \right\} \tag{6.11}$$

Therefore, evaluating the gradient of $U(\mathbf{r})$ in equation (6.10) a function $S(\mathbf{r})=0$ is obtained, viz

$$S(\mathbf{r}) = x(x+\mu) + y^2 - \frac{1-\mu}{r_1} - \frac{\mathbf{r}_1 . \mathbf{r}_2}{r_2^3} \tag{6.12}$$

The function $S(\mathbf{r})$ has two topologically disconnected boundary surfaces S_1 and S_2 which define the boundary to the region of existence of stationary solutions, Figures 6.2 and 6.3. The region of existence of the stationary solutions lies between these two surfaces and is defined by the region $S_1 \cap S_2'$.

Figure 6.2(a)

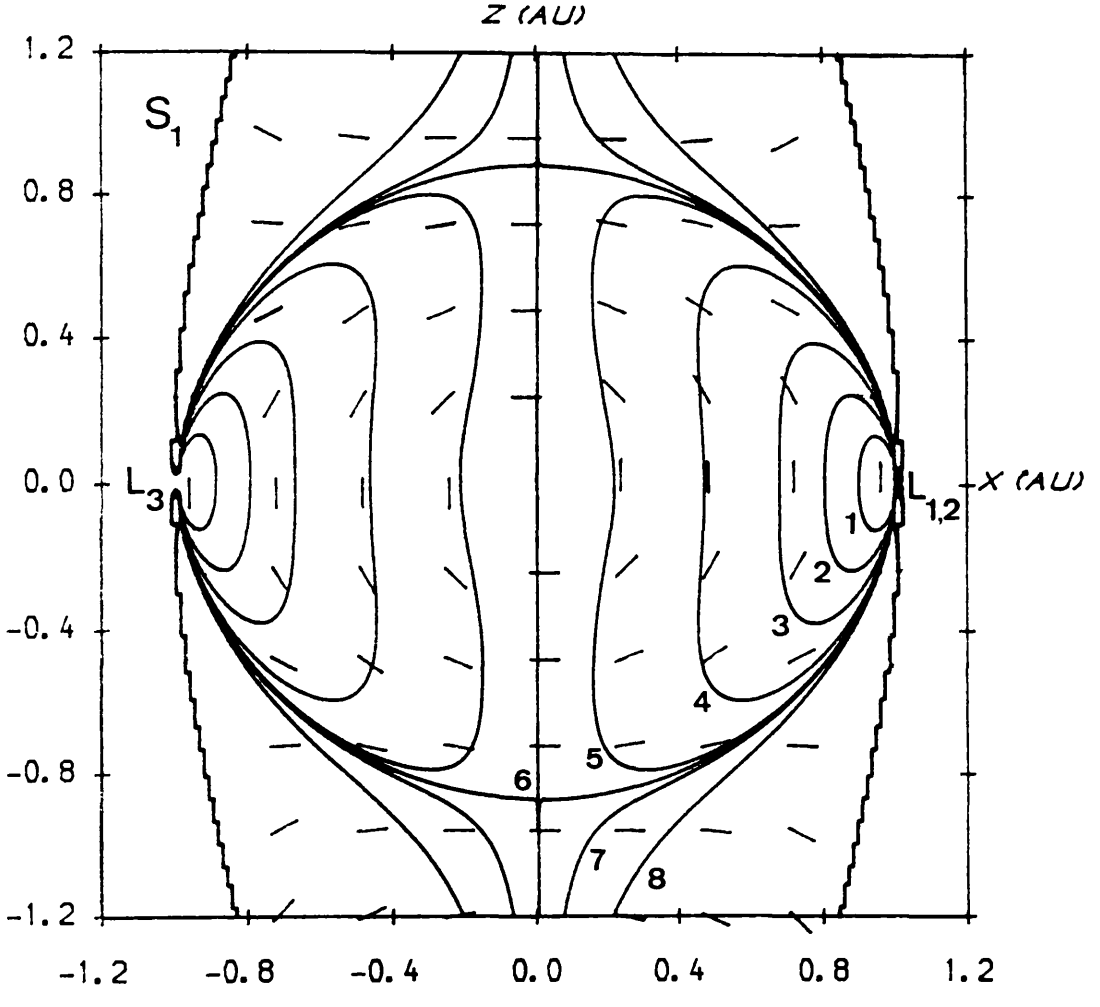


Figure 6.2 Section of the level surfaces in the Earth-Sun system (a) normal to the plane of the system and (b) in the plane of the system. The required sail loadings are given by; (1) 0.3 (2) 0.5 (3) 0.7 (4) 0.9 (5) 1.0 (6) 1.01 (7) 1.1. The contour S_1 represents the outer boundary surface.

Figure 6.2(b)

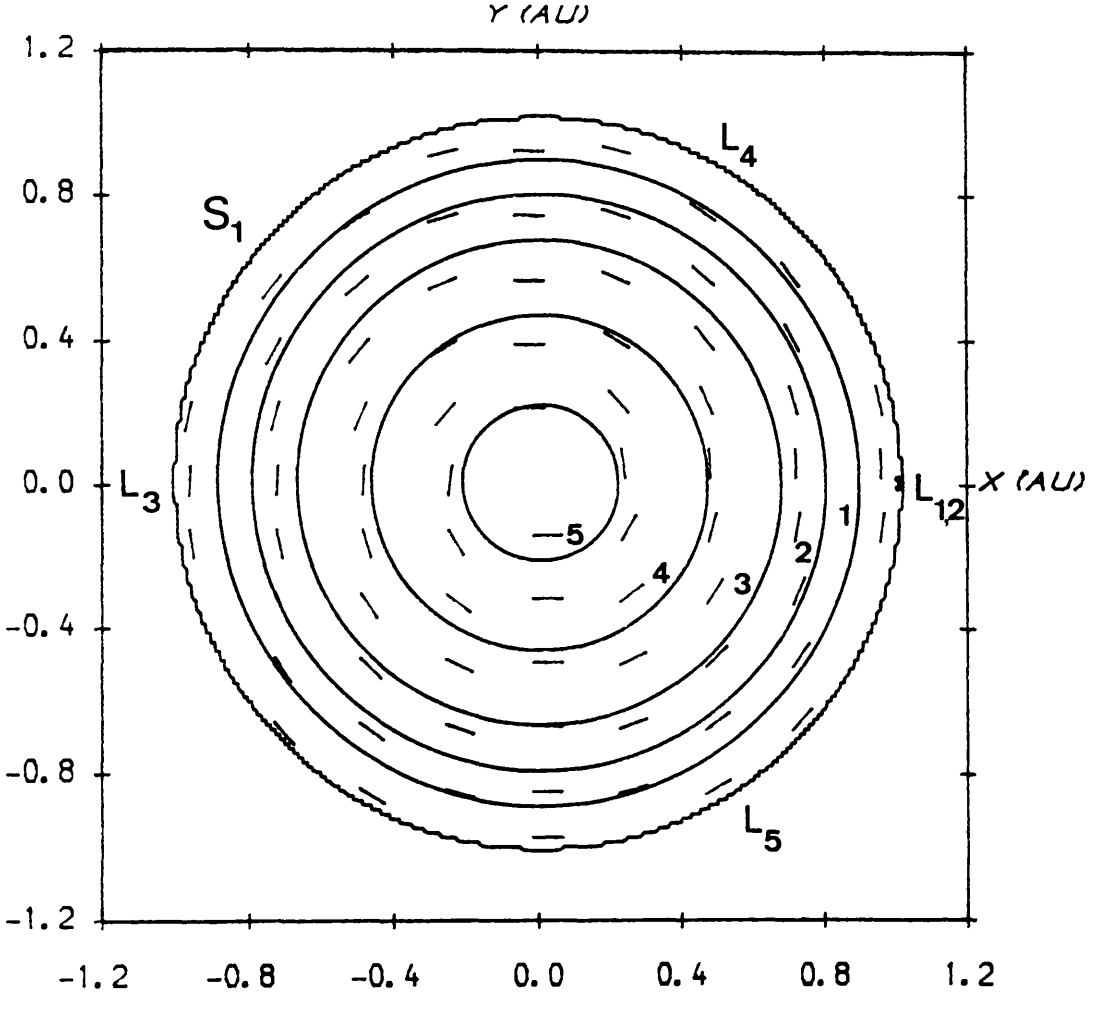
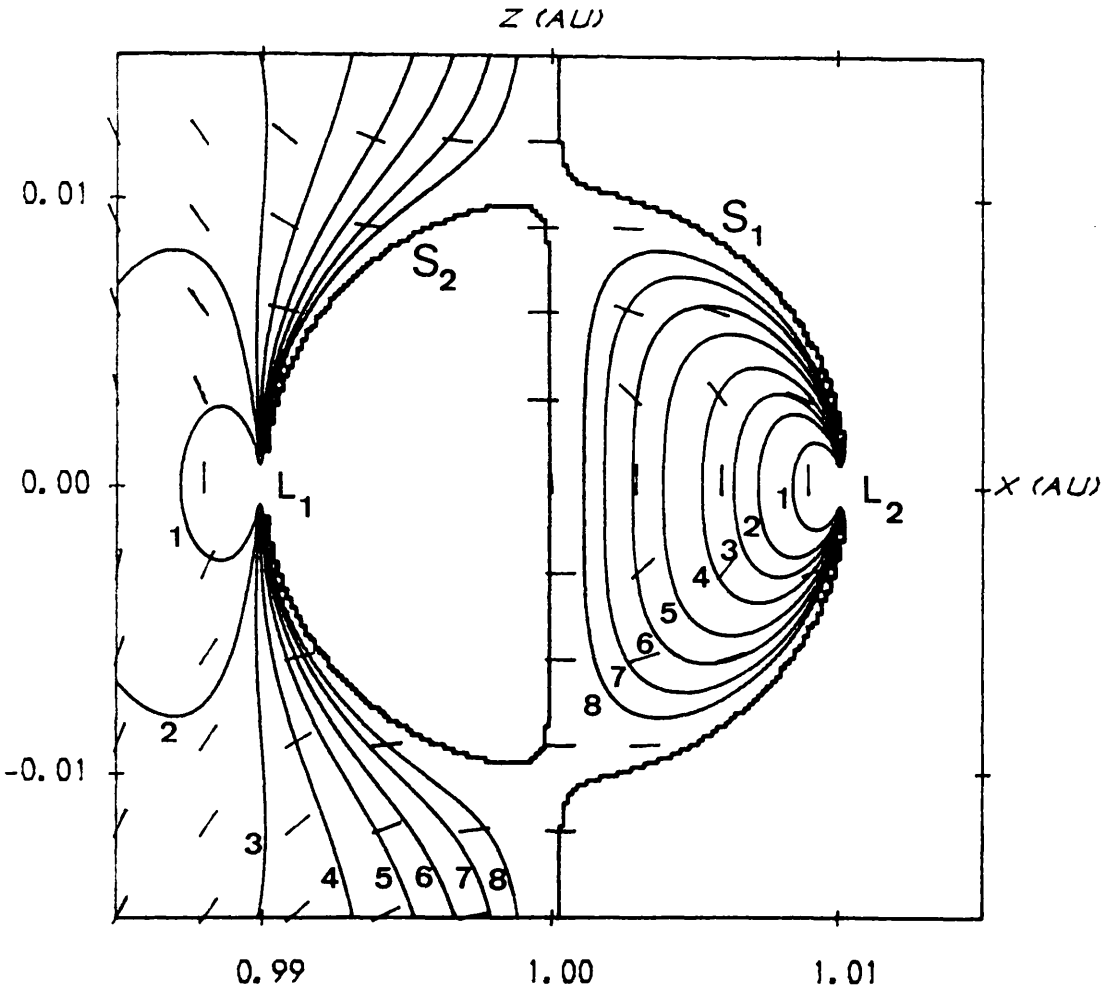
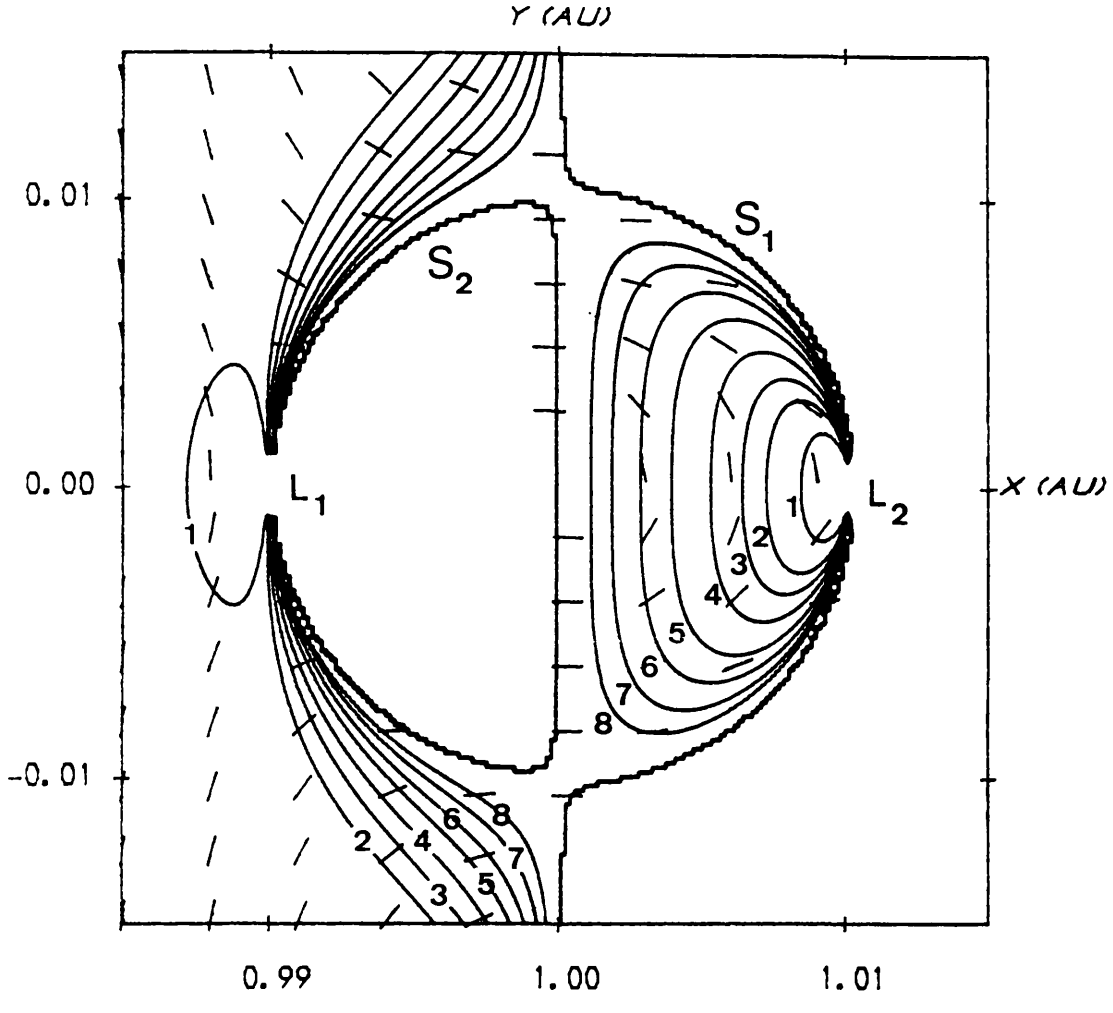

Y (AU)
X (AU)
1.2
0.8
0.4
0.0
-0.4
-0.8
-1.2
-1.2
-0.8
-0.4
0.0
0.4
0.8
1.2
L_4
S_1
L_3
$L_{1,2}$
L_5
1
2
3
4
5

Figure 6.3(a)



Section of the level surfaces in the near Earth region (a) normal to the plane of the system and (b) in the plane of the system. The required sail loadings are given by; (1) 0.02 (2) 0.04 (3) 0.06 (4) 0.1 (5) 0.2 (6) 0.4 (7) 1.0 (8) 3.0. The contour S_2 represents the inner boundary surface.

Figure 6.3(b)



The outer surface S_1 possesses a cylindrical topology and excludes solutions along the x-axis from $-\infty<x<xL_3$ and $xL_2<x<+\infty$ while the inner surface S_2 excludes solutions along $xL_1<x<1-\mu$. All of the five classical stationary solutions lie on $S_1 \cup S_2$ since they are the solutions $\nabla U(\mathbf{r})=0$ of equation (6.10). In general the sail loading level surfaces approach the boundary surface asymptotically with $\beta\to\infty$ as then $\mathbf{r}_1.\nabla U(\mathbf{r})\to 0$ in equation (6.9).

6.5 Stationary Solutions in the Earth-Sun System

Level surfaces of constant sail loading may now be generated from equation (6.9) for the Earth-Sun system ($\mu=3.036\times10^{-6}$). For ease of illustration sections of the surfaces through the x-y and x-z planes are used. Then, only the pitch angle α will be required to completely describe the sail attitude required for a stationary solution. In general though the two angles (α,χ) are required to describe the sail attitude for a stationary solution at some arbitrary position.

The sections of the level surfaces generated by equation (6.9) are shown in Figures 6.2 and 6.3. The sections define families of one parameter level curves representing subsets of the continuum of new artificial stationary solutions with equal sail loading. The sail attitude required is also shown. From Figure 6.2 it can be seen that in the far Earth region the level surfaces are a family of topologically nested tori with the inner radius of the torus vanishing as $\beta\to1$. In the plane of the system the level curves are near circular with the sail pitch angle $\alpha\approx0$. These curves represent the sail loading required for a circular heliocentric orbit with an orbital period of one year. Out of the plane of the system it can be seen that the solutions are essentially Earth synchronous heliocentric halo orbits, as discussed in

chapter 4. Along the z-axis of the system there are stationary solutions above the poles of the Sun with $\beta=1$.

A detailed plot of the sections of the level surfaces near the Earth is shown in Figure 6.3. It can be seen that the level surfaces around L_1 and L_2 accessible to the solar sail expand with increased sail loading, but always contain the classical Lagrange point they are associated with, as this corresponds to the solution $\beta\neq0$, $\mathbf{r}_1.\mathbf{n}=0$.

For $\beta\gg1$ there are no solutions in the plane of the system, except at the five classical Lagrange points where $\mathbf{r}_1.\mathbf{n}=0$. It can be seen in Figure 6.2 away from the Earth that for $\beta>1$ the level surfaces have undergone a topological change to become a family of nested cylinders. There are also no intersections with, and so no solutions in, the x-y plane except at the $L_{4,5}$ points where $\mathbf{r}_1.\mathbf{n}=0$ and the sail loading is undefined. There are however out-of-plane solutions corresponding to halo type orbits at greater distances above the plane of the system. In the near Earth region, Figure 6.3, it is seen that for an increased sail loading the level surfaces continue to expand, asymptotically approaching the boundary surfaces S_1 and S_2.

6.6 Stability and Control

Now that the existence of the artificial stationary solutions has been established it is necessary to examine their stability. The general vector dynamical equation is given by equation (6.5) as

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \nabla U(\mathbf{r}) = \mathbf{a} \qquad (6.13)$$

It will be assumed that the sail is stationary on some level surface at a point $\mathbf{r}_0$. Then the dynamical equation in a local neighbourhood of

$\mathbf{r}_o$ is obtained in the usual manner using an arbitrary linear perturbation $\boldsymbol{\varepsilon}$, such that $\mathbf{r}_o \rightarrow \mathbf{r}_o + \boldsymbol{\varepsilon}$. Since $\mathbf{r}_o$ is a stationary solution a variational equation is obtained, viz

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\boldsymbol{\varepsilon}}{dt} + \nabla U(\mathbf{r}_o+\boldsymbol{\varepsilon}) - \mathbf{a}(\mathbf{r}_o+\boldsymbol{\varepsilon}) = 0 \qquad (6.14)$$

The potential gradient and the radiation pressure acceleration may be expanded in trivariate Taylor series about the stationary solution to first order as

$$\nabla U(\mathbf{r}_o+\boldsymbol{\varepsilon}) = \nabla U(\mathbf{r}_o) + \frac{\partial}{\partial \mathbf{r}} \nabla U(\mathbf{r}) \Big|_{\mathbf{r}=\mathbf{r}_o} \boldsymbol{\varepsilon} + 0\ (|\boldsymbol{\varepsilon}|^2) \qquad (6.15a)$$

$$\mathbf{a}(\mathbf{r}_o+\boldsymbol{\varepsilon}) = \mathbf{a}(\mathbf{r}_o) + \frac{\partial}{\partial \mathbf{r}} \mathbf{a}(\mathbf{r}) \Big|_{\mathbf{r}=\mathbf{r}_o,\ \mathbf{n}=\mathbf{n}_o} \boldsymbol{\varepsilon} + 0\ (|\boldsymbol{\varepsilon}|^2) \qquad (6.15b)$$

Then, since $\nabla U(\mathbf{r}_o)=\mathbf{a}(\mathbf{r}_o)$ for a stationary solution, a linear variational equation is obtained, viz

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} + \mathbf{M}_1 \frac{d\boldsymbol{\varepsilon}}{dt} + (\mathbf{M}_2 - \mathbf{M}_3)\, \boldsymbol{\varepsilon} = \mathbf{0} \qquad (6.16)$$

where $\mathbf{M}_{2,3}$ the gravity and radiation gradient tensors and the skew symmetric gyroscopic matrix $\mathbf{M}_1$ are given by

$$\mathbf{M}_1 = \begin{Bmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{Bmatrix}, \quad \mathbf{M}_2 = \langle U_{ij} \rangle_{(i,j)\in(x,y,z)}, \quad \mathbf{M}_3 = \langle a_{ij} \rangle \qquad (6.17)$$

where U_{ij} is the (i,j) partial derivative of the potential with respect to the cartesian axes and a_{ij} is the j^{th} derivative of the i^{th} component of the solar radiation pressure acceleration. The stability of the

system may be investigated in the usual manner by examining the system eigenvalues resulting from the characteristic polynomial. This may be carried out by substituting an exponential solution of the form

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_0 e^{st} \quad , \quad s = \sigma + i\omega \quad , \quad i=\sqrt{-1} \tag{6.18}$$

Substituting this solution into equation (6.16) yields a matrix equation of the form

$$(s^2\mathbf{I} + s\mathbf{M}_1 + \mathbf{M}^*)\boldsymbol{\varepsilon}_0 = 0 \tag{6.19}$$

where $\mathbf{M}^*=\mathbf{M}_2-\mathbf{M}_1$. For non-trivial solutions a vanishing secular determinant is required, which then gives the characteristic polynomial of the system P(s)=0, viz

$$P(s) = \sum_{j=0}^{6} a_{6-j}\, s^j \tag{6.20}$$

where, owing to the fundamental theorem of algebra, P(s) has roots $s_j=\sigma_j+i\omega_j$ (j=1,6). For asymptotic stability it is required that all of the system eigenvalues are in the left hand complex plane so that $\sigma_j<0$ (j=1,6). However, for stability in the Lyapunov sense the weaker condition, that all the roots of P(s) are at least purely imaginary, is required. This constrains the motion to a local neighbourhood of the nominal stationary solution.

The coefficients of the polynomial P(s) are given by

$$a_0 = 1 \tag{6.21a}$$

$$a_1 = 0 \tag{6.21b}$$

$$a_2 = M^*_{11} + M^*_{22} + M^*_{33} + 4 \tag{6.21c}$$

$$a_3 = 2(M^*_{21} - M^*_{12}) \tag{6.21d}$$

$$a_4 = M^*_{11}M^*_{22} + M^*_{11}M^*_{33} + M^*_{22}M^*_{33} - M^*_{23}M^*_{32} - M^*_{13}M^*_{31} - M^*_{12}M^*_{21} + 4M^*_{33} \tag{6.21e}$$

$$a_5 = 2M^*_{33}(M^*_{21} - M^*_{12}) + 2(M^*_{32}M^*_{13} - M^*_{23}M^*_{31}) \tag{6.21f}$$

$$a_6 = M^*_{11}M^*_{22}M^*_{33} - M^*_{11}M^*_{23}M^*_{32} - M^*_{33}M^*_{12}M^*_{21} - M^*_{22}M^*_{31}M^*_{13} + M^*_{21}M^*_{32}M^*_{13} + M^*_{12}M^*_{23}M^*_{31} \tag{6.21g}$$

Since $a_1=0$, an application of the Routh-Hurwitz criterion implies that at least one eigenvalue will not lie in the left hand complex plane, (ie. there is at least one eigenvalue with $\sigma_j \geqslant 0$). Therefore the system does not naturally possess asymptotic stability. Given this fact the condition for Lyapunov type stability with purely imaginary eigenvalues, $\sigma_j=0$ (j=1,6), will be established. Substituting for $s=j\omega$ P(s) becomes

$$P(j\omega) = -\omega^6 + a_2\omega^4 - ja_3\omega^3 - a_4\omega^2 + ja_5\omega + a_6 \tag{6.22}$$

For the condition P(s)=0 to hold it is required that both the real and purely imaginary parts of the polynomial are identically zero, viz

$$-\omega^6 + a_2\omega^4 - a_4\omega^2 + a_6 = 0 \tag{6.23a}$$

$$j\omega(a_5 - \omega^2 a_3) = 0 \tag{6.23b}$$

Six consistent solutions of equations (6.23) with $\omega_j^2>0$ (j=1,6) are now required. From equation (6.23b) it is seen that $\omega_1=0$, $\omega_{2,3}=\pm\sqrt{(a_5/a_3)}$. However, the solution $\omega_1=0$ is obviously inconsistent with equation (6.23a). The remaining solutions $\omega_{2,3}$ are also not generally consistent with equation (6.23a). However, equation (6.23b) is satisfied if $a_3=a_5=0$.

The eigenvalues of the system are then determined as conjugate pairs from equation (6.23a), which may or may not have real solutions.

Therefore a necessary, but not sufficient, condition for Lyapunov stability is then $a_3=0 \Rightarrow (M^*_{21}-M^*_{12})=0$. However, since the potential is conservative $U_{yx}-U_{xy}=0$, so that $a_3=0 \Rightarrow (a_{yx}-a_{xy})=0$. Similarly the condition $a_5=0$ requires that $a_{zx}-a_{xz}=0$ and $a_{yz}-a_{zy}=0$. Taken together these conditions imply that $\beta=0$ or

$$\nabla \mathbf{x}\, \mathbf{a} = 0 \qquad (6.24)$$

That is, that the solar radiation pressure acceleration must be conservative and so must be derivable from some scalar potential. Therefore, the required conditions for Lyapunov stability are $\alpha=0$ (modified photogravitational system), or $\beta=0$ (classical restricted system). In practice the solutions away from the Earth will behave as Earth synchronous heliocentric halo orbits with their associated regions of Poincaré stability and instability.

It has been shown then that the set of new stationary solutions do not possess a natural asymptotic stability and that Lyapunov stability is only possible for the particular solutions when the sail is oriented along the Sun-line. In general therefore a control scheme is required to ensure asymptotic stability. A simple control scheme using a combination of proportional and derivative feedback to the sail attitude will now be developed.

Including first order variations in the sail attitude $\mathbf{n}_o \rightarrow \mathbf{n}_o + \delta\mathbf{n}$, the open loop variational system becomes

$$\frac{d^2\delta}{dt^2} + M_1 \frac{d\delta}{dt} + M^* \delta = N\, \delta n \quad , \quad N = \left.\frac{\partial \mathbf{a}}{\partial \mathbf{n}}\right|_{r=r_o,\ n=n_o} \qquad (6.25)$$

In the full six-dimensional phase space $\mathbf{x}=(\delta,d\delta/dt)$ the system then becomes

$$\frac{d\mathbf{x}}{dt} = \mathbf{P}\mathbf{x} + \mathbf{Q}\,\delta\mathbf{n} \quad , \quad \mathbf{P} = \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{I} \\ \cdots & \cdots & \cdots \\ -\mathbf{M}^* & \vdots & -\mathbf{M}_1 \end{bmatrix} \quad , \quad \mathbf{Q} = \begin{bmatrix} \mathbf{0} \\ \\ \mathbf{N} \end{bmatrix} \tag{6.26}$$

where $\mathbf{P}$ is the system matrix and $\mathbf{Q}$ is the input distribution matrix.

In order to proceed further it is necessary to determine if the system is fully controllable. Therefore, the 6x6 controllability matrix $\mathbf{C}=\langle \mathbf{Q}, \mathbf{PQ}, \mathbf{P}^2\mathbf{Q}, \mathbf{P}^3\mathbf{Q} \rangle$ must have full rank, viz

$$\mathbf{C} = \begin{bmatrix} \mathbf{0} & \vdots & \mathbf{N} & \vdots & -\mathbf{M}_1\mathbf{N} & \vdots & -\mathbf{M}^*\mathbf{N} + \mathbf{M}_1{}^2\mathbf{N} \\ & \vdots & & \vdots & & \vdots & \\ \mathbf{N} & \vdots & -\mathbf{M}_1\mathbf{N} & \vdots & -\mathbf{M}^*\mathbf{N} + \mathbf{M}_1{}^2\mathbf{N} & \vdots & 2\mathbf{M}^*\mathbf{M}_1\mathbf{N} - \mathbf{M}_1{}^3\mathbf{N} \end{bmatrix} \tag{6.27}$$

Since $\mathbf{N}\neq 0$ if $\mathbf{r}_1.\mathbf{n}\neq 0$, in general all the rows of $\mathbf{C}$ are linearly independent so that $r(\mathbf{C})=6$ and the system is fully controllable. The control will then be defined as

$$\delta\mathbf{n} = \Lambda_1.\delta + \Lambda_2.\left\{\frac{d\delta}{dt}\right\} \tag{6.28}$$

so that the sail attitude trim is given as function of the sail position and velocity relative to the nominal stationary solution. The closed loop system is then given by

$$\frac{d^2\delta}{dt^2} + (\mathbf{M}_1-\mathbf{N}\Lambda_2)\,\frac{d\delta}{dt} + (\mathbf{M}^*-\mathbf{N}\Lambda_1)\,\delta = 0 \tag{6.29}$$

Therefore, the feedback control now allows $\sigma_j<0$ $(j=1,6)$ with a suitable choice of gain matrices $\Lambda_{1,2}$ so that the stationary solutions may have asymptotic stability. In general the choice of gains will depend to a

large extent on operational requirements. However, to demonstrate the existence of asymptotic stability the gain matrices will be chosen as

$$\Lambda_1 = N^{-1}(M^* - \lambda_2 I) \quad , \quad \Lambda_2 = N^{-1}(M_1 - \lambda_1 I) \tag{6.30}$$

Substituting these gains into the variational system a damped harmonic equation is obtained with the damping proportional to the gain constants $\lambda_{1,2}$, viz

$$\left\{\frac{d^2\delta}{dt^2}\right\} + \lambda_1\left\{\frac{d\delta}{dt}\right\} + \lambda_2\,\delta = 0 \tag{6.31}$$

which has a characteristic polynomial

$$s^2 + \lambda_1 s + \lambda_2 = 0 \tag{6.32}$$

Since the gain constants $\lambda_{1,2}$ may be arbitrarily chosen the eigenvalues may be chosen to be in the left hand complex plane ensuring asymptotic stability. A solution of equation (6.31) is then

$$\delta = \delta_0 e^{st} \quad , \quad s = -\left\{\frac{\lambda_1}{2}\right\} \pm \left\{\frac{\lambda_1^2}{4} - \lambda_2\right\}^{1/2} \tag{6.33}$$

so that for asymptotic stability it is required that $\lambda_1>0$, $\lambda_2>\lambda_1^2/4$. It has been shown then that, in principle, the stationary solutions are controllable using a feedback to the sail attitude and that asymptotic stability can therefore be achieved.

6.7 Dynamical Equations for the Earth-Moon System

An idealised perfectly reflecting solar sail will now be considered in a co-rotating reference frame of constant angular velocity Ω with a point mass Earth m_1 and Moon m_2. The dynamics of the Earth-Moon

restricted three-body system are quite different from the Earth-Sun system in that the Sun-line **S** is not fixed in the co-rotating frame, but rotates once per synodic lunar month. It will be assumed that the solar radiation pressure is constant in magnitude over the scale of the problem. In the units of the system the Earth-Moon distance is taken to be unity. Therefore, the sail loading parameter is now defined as the solar radiation pressure acceleration made dimensionless with respect to the Earth's gravitational acceleration at the lunar distance. The spacecraft mass per unit area is then related to the sail loading parameter by the relation $\sigma=3.385\beta^{-1}$ gm^{-2}.

The vector dynamical equation for a solar sail in this co-rotating frame may be written as

$$\frac{d^2\mathbf{r}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\mathbf{r}}{dt} + \nabla U(\mathbf{r}) = \mathbf{a} \qquad (6.34)$$

where the co-rotating three-body potential $U(\mathbf{r})$ and the solar radiation pressure acceleration **a** are given by

$$U(\mathbf{r}) = -\left\{\frac{1}{2}\,|\boldsymbol{\Omega} \times \mathbf{r}|^2 + \frac{1-\mu}{|\mathbf{r}_1|} + \frac{\mu}{|\mathbf{r}_2|}\right\}, \quad \mathbf{a} = \beta\,(\mathbf{S}.\mathbf{n})^2\,\mathbf{n} \qquad (6.35)$$

where $\mu=(m_2/m_1+m_2)=0.01215$ is the mass ratio of the Earth-Moon system. The sail attitude is constrained such that $\mathbf{S}.\mathbf{n}\geqslant 0$ and the direction of the Sun-line is given by

$$\mathbf{S} = (\cos(\Omega_* t),\ -\sin(\Omega_* t),\ 0) \qquad (6.36)$$

where $\Omega_*=0.9252$ is the angular rate of the Sun-line in the co-rotating frame in dimensionless units. The small annual changes ($\pm 5^{\circ}$) in the inclination of the Sun-line with respect to the plane of the system are

ignored.

By again requiring stationary solutions in the co-rotating frame and taking vector products equation (6.34) may be solved to obtain the required sail attitude as

$$\mathbf{n} = \frac{\nabla U(\mathbf{r})}{|\nabla U(\mathbf{r})|} \tag{6.37}$$

which is time independant. The required sail loading may also be obtained, but is however time dependent due to the rotating Sun-line, viz

$$\beta(t) = \frac{\nabla U(\mathbf{r}).\mathbf{n}}{(\mathbf{S}(t).\mathbf{n})^2} \tag{6.38}$$

As before the region of existence of stationary solutions is bounded, the boundary being defined by the time dependent condition $\mathbf{S}.\mathbf{n}\geqslant 0$. This condition yields a function $T(\mathbf{r};t)=0$ defining the time dependent boundary surface to the regions of existence of solutions. This function is given by

$$T(\mathbf{r};t) = \frac{\partial U(\mathbf{r})}{\partial x}\cos(\Omega_* t) - \frac{\partial U(\mathbf{r})}{\partial y}\sin(\Omega_* t) \tag{6.39}$$

so that on this surface the sail attitude is normal to the Sun-line. Again there are two topologically disconnected regions T_1 and T_2.

For a fixed sail loading the conditions for stationary solutions derived above are only valid instantaneously at some time t_o. For a short, finite duration stay a small open loop control acceleration is required to compensate for the moving Sun-line. This acceleration will require a small variation in the sail loading since the sail attitude required is time independent. Although equation (6.38) gives the

required change in sail loading as a function of the Sun-line position, an approximate expression can be obtained by expanding the solar radiation pressure acceleration about the stationary solution at position $\mathbf{r}_o$ and time t_o. At a time Δt later the condition for continued stationarity is given by

$$\nabla U(\mathbf{r}_o) = \mathbf{a}(t_o) + \frac{1}{j!}\sum_{j=1}^{\infty}\frac{d^j\mathbf{a}}{dt^j} \tag{6.40}$$

Therefore, for a short duration stay ($\Delta t \ll 1$) the first order trim in the sail loading is given by

$$\frac{\Delta\beta}{\beta} = -\,2\mathbf{S}.\mathbf{n}\left\{\mathbf{n}.\frac{d\mathbf{S}}{dt}\right\}\Delta t \tag{6.41}$$

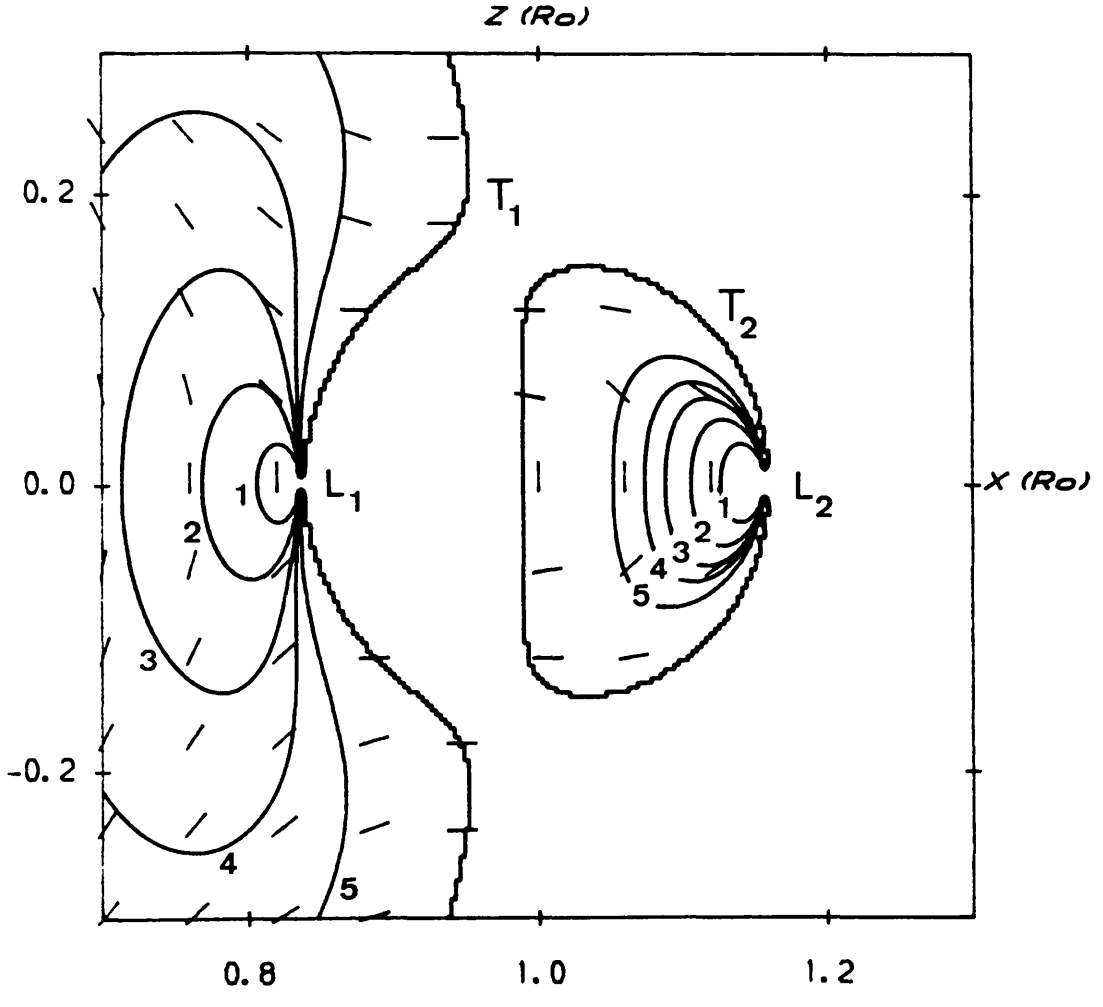
If $\Delta\beta/\beta$ is remains by limiting Δt to 5×10^{-3} a stay of 3.3 hours is possible. For sails with a large trim capability much longer duration stays would however be possible. Furthermore, if $\mathbf{n}=\mathbf{S}$ then $\mathbf{n}.(d\mathbf{S}/dt)=0$ so that, to first order, no variation in the sail loading is required.

6.8 Stationary Solutions in the Earth-Moon System

Using equation (6.38) level surfaces of constant sail loading in the Earth-Moon sytem may be generated. Again, sections of the surfaces through the x-y and x-z planes are used. Then, only the pitch angle α is required to describe the sail attitude.

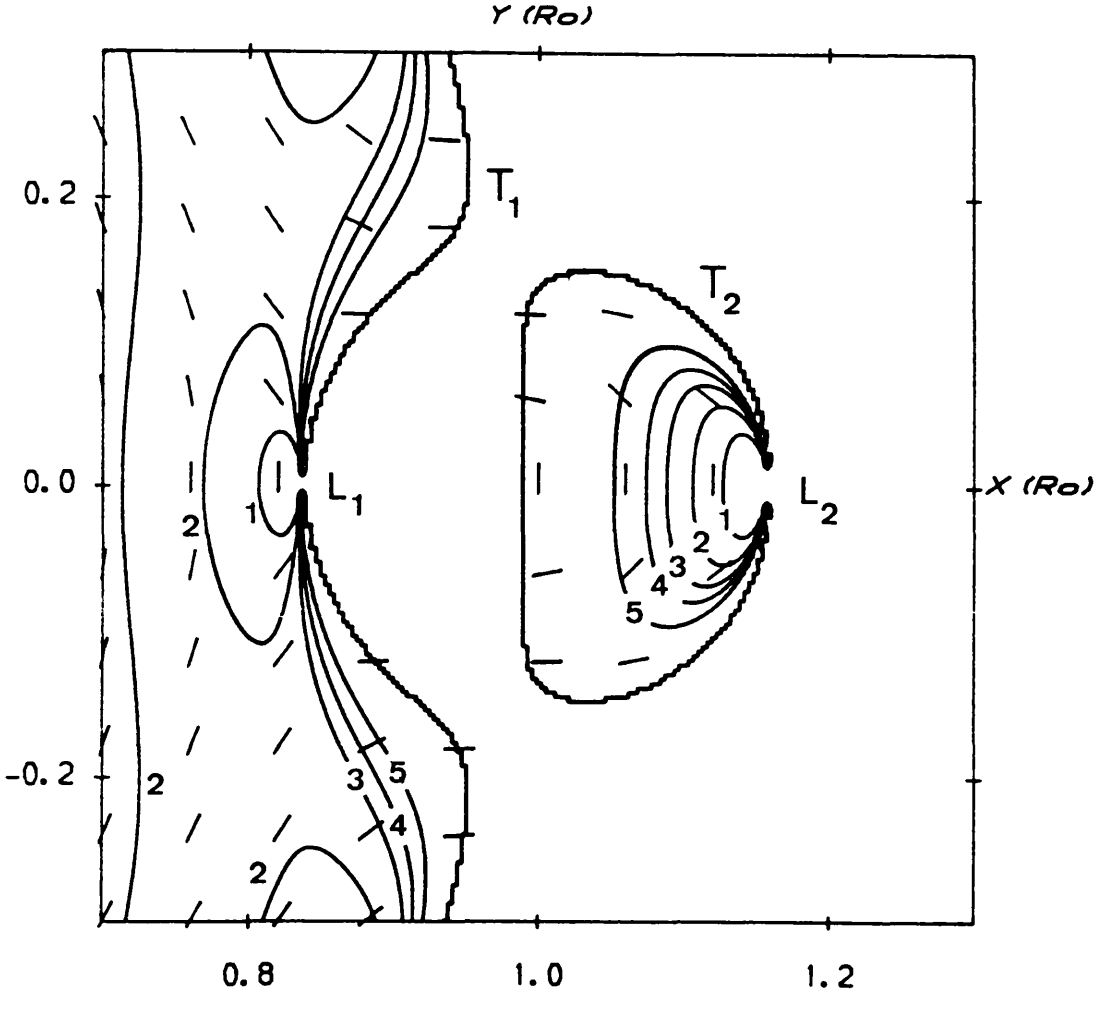
At time $\Omega_* t = 0^o$ the Sun-line is directed along the Earth-Moon line. Sections of the level surfaces of constant sail loading at this time are shown in Figure 6.4. It can be seen that the surfaces expand with increasing sail loading in a similar manner to the Earth-Sun system.

Figure 6.4(a)



Section of the level surfaces in the near lunar region (a) normal to the plane of the system and (b) in the plane of the system at time $\Omega_* t = 0^o$. The sail loading values are given by; (1) 0.3 (2) 0.6 (3) 1.0 (4) 1.5 (5) 3.0. The contours T_1 and T_2 represent the time dependent boundary surfaces.

Figure 6.4(b)



This is to be expected due to the configuration of the Earth-Moon-Sun system at this time. Sections of the boundary surfaces T_1 and T_2 are also shown.

Some time later in the synodic month when $\Omega_* t = 45^\circ$ the topology of the surfaces radically transforms, Figure 6.5. Normal to the plane of the system the surfaces are still symmetric. However, in the plane of the system the surfaces about the L_1 and L_2 points are asymmetric and connect at a loading value of approximately 1.2. Similarly, the boundary surfaces T_1 and T_2 are now connected. It can however be seen that the required sail attitude is time independent. Finally, at time $\Omega_* t = 180^\circ$ the regions of existance of solutions have reversed with respect to $\Omega_* t = 0^\circ$ and so solutions are now forbidden within surface T_2 and are allowed outwith T_1, Figure 6.6.

6.9 Lunar Lagrange Point Halo-Type Orbits

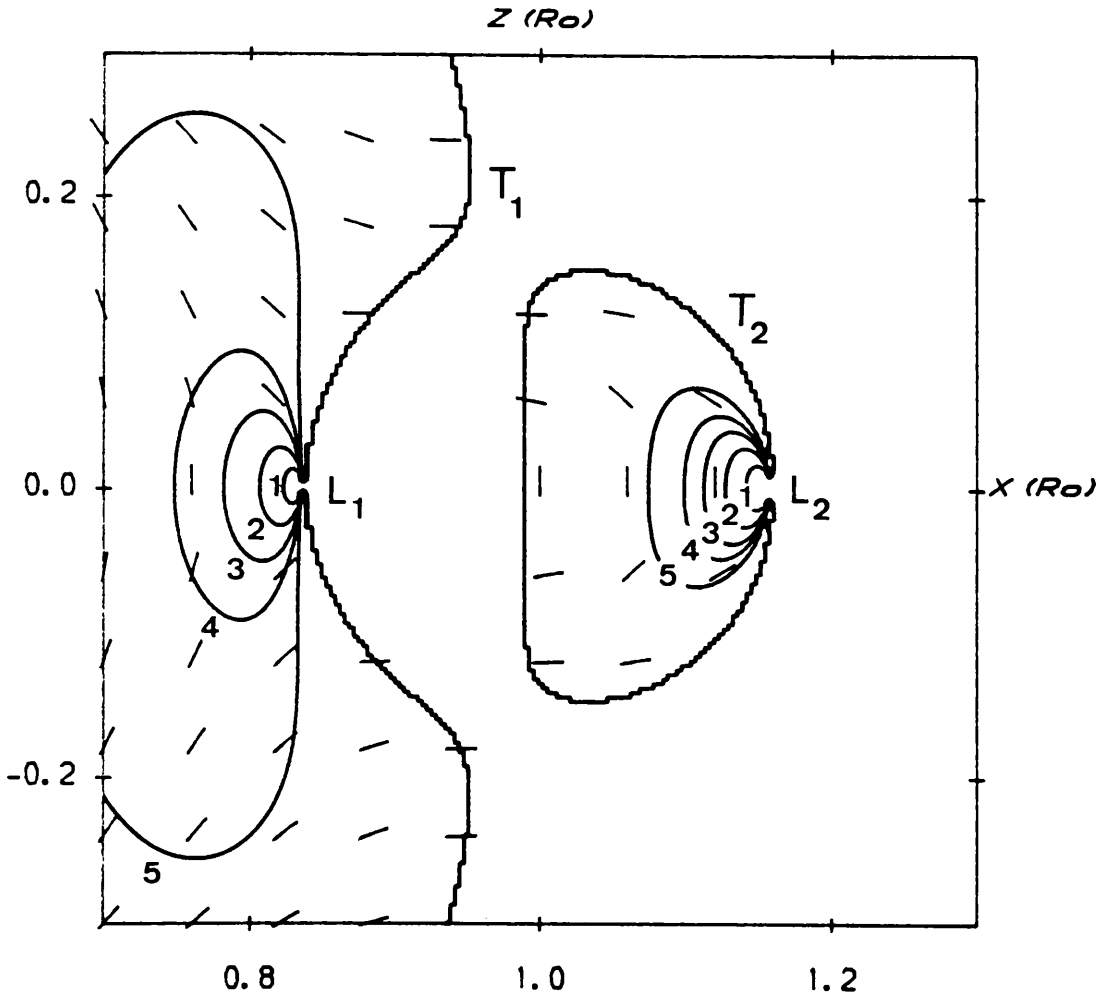
The dynamics of a solar sail in the neighbourhood of the lunar L_2 point at $\mathbf{r}_L$ will now be investigated and it will be demonstrated that a periodic out-of-plane trajectory exists. Perturbing the vector dynamical equation such that $\mathbf{r}_L \rightarrow \mathbf{r}_L + \boldsymbol{\varepsilon}$ it is found that

$$\frac{d^2\boldsymbol{\varepsilon}}{dt^2} + 2\boldsymbol{\Omega} \times \frac{d\boldsymbol{\varepsilon}}{dt} + \nabla U(\mathbf{r}_L + \boldsymbol{\varepsilon}) = \mathbf{a}(\mathbf{r}_L + \boldsymbol{\varepsilon}) \tag{6.42}$$

where $\boldsymbol{\varepsilon} = (\xi, \eta, \zeta)$ represent small displacements from the L_2 point along the (x,y,z) directions. The potential gradient may be expanded in a trivariate Taylor seies about the L_2 point to first order as

$$\nabla U(\mathbf{r}_L + \boldsymbol{\varepsilon}) = \nabla U(\mathbf{r}_L) + \frac{\partial}{\partial \mathbf{r}} \nabla U(\mathbf{r}) \Big|_{\mathbf{r}=\mathbf{r}_L} \boldsymbol{\varepsilon} + O(|\boldsymbol{\varepsilon}|^2) \tag{6.43}$$

Figure 6.5(a)



Section of the level surfaces in the near lunar region (a) normal to the plane of the system and (b) in the plane of the system at time $\Omega_* t = 45^{\circ}$. The sail loading values are given by; (1) 0.3 (2) 0.6 (3) 1.0 (4) 1.5 (5) 3.0. The contours T_1 and T_2 represent the time dependent boundary surfaces.

Figure 6.5(b)

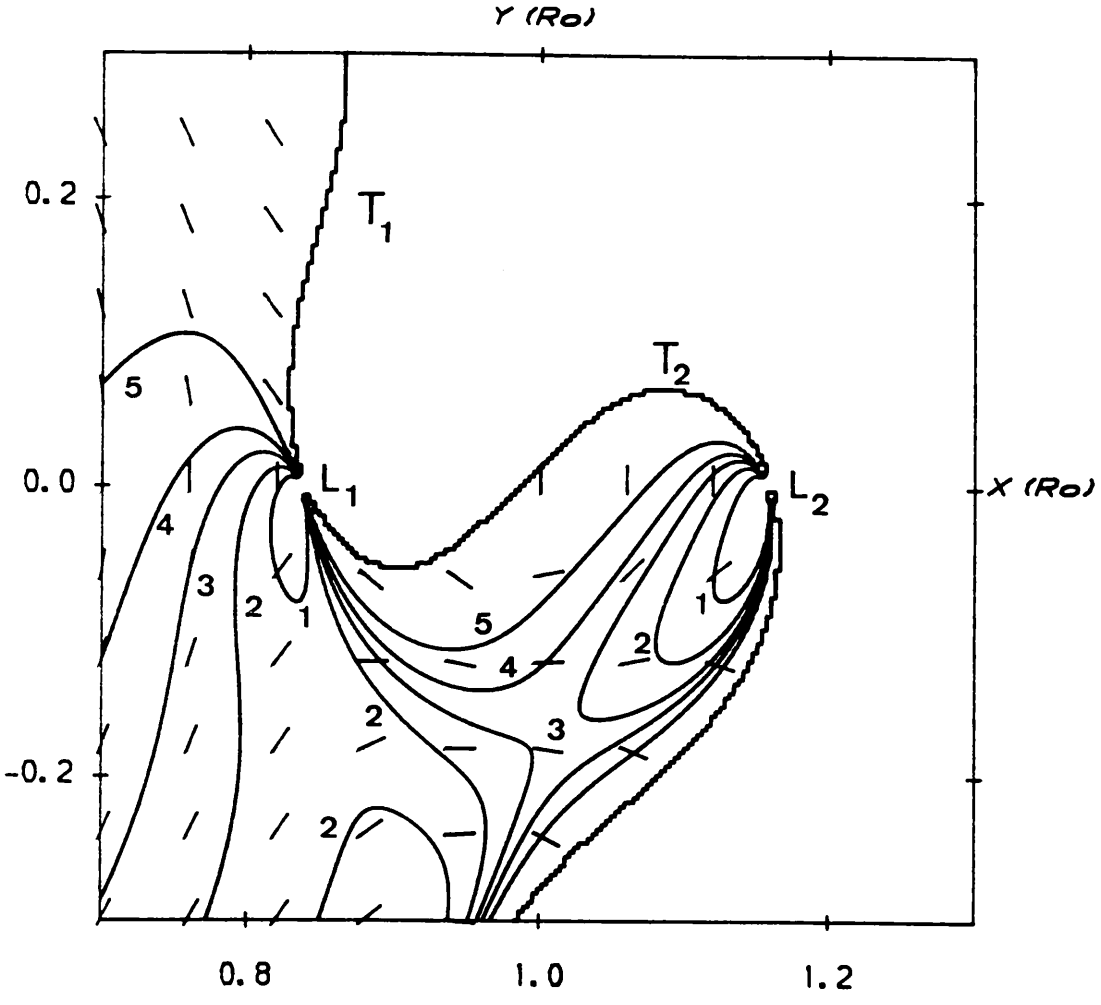
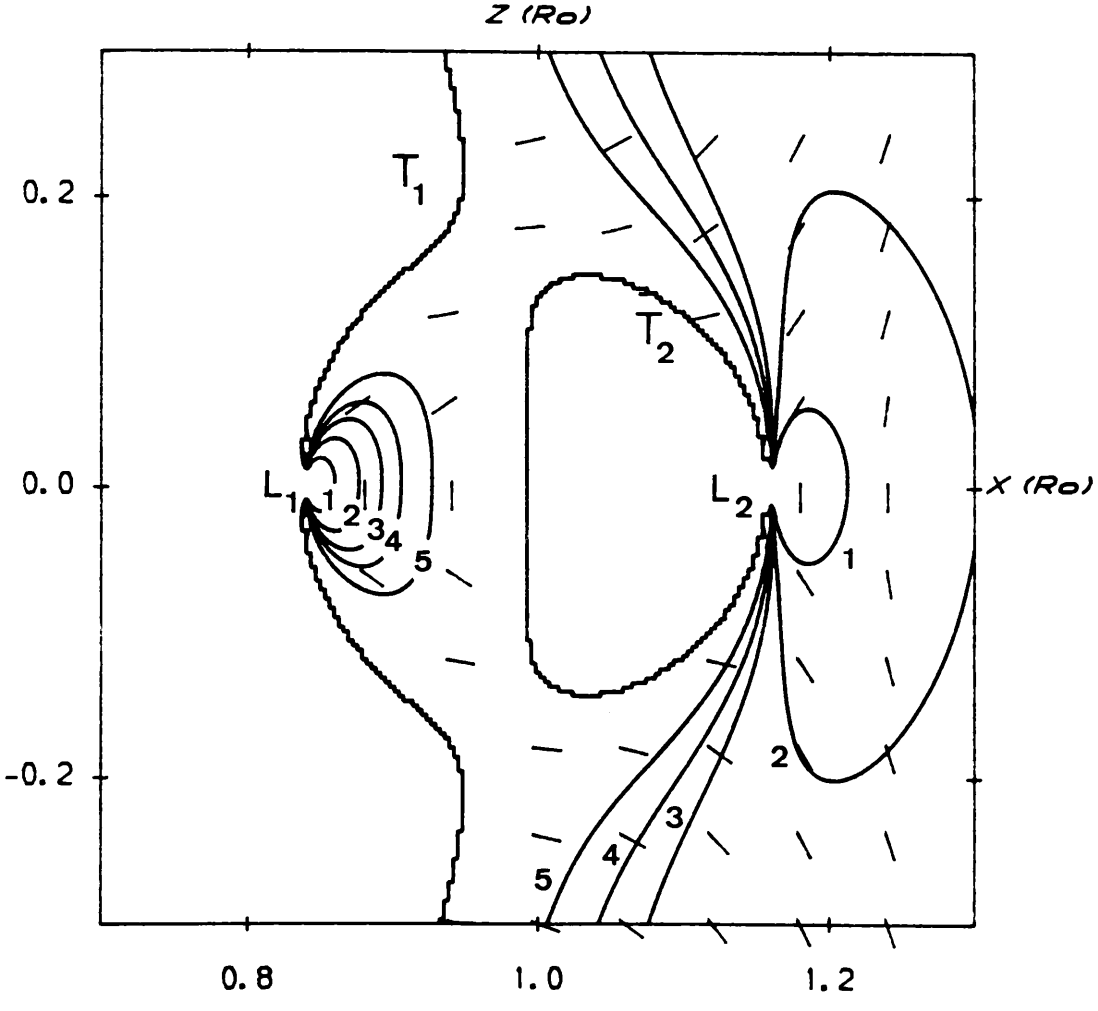
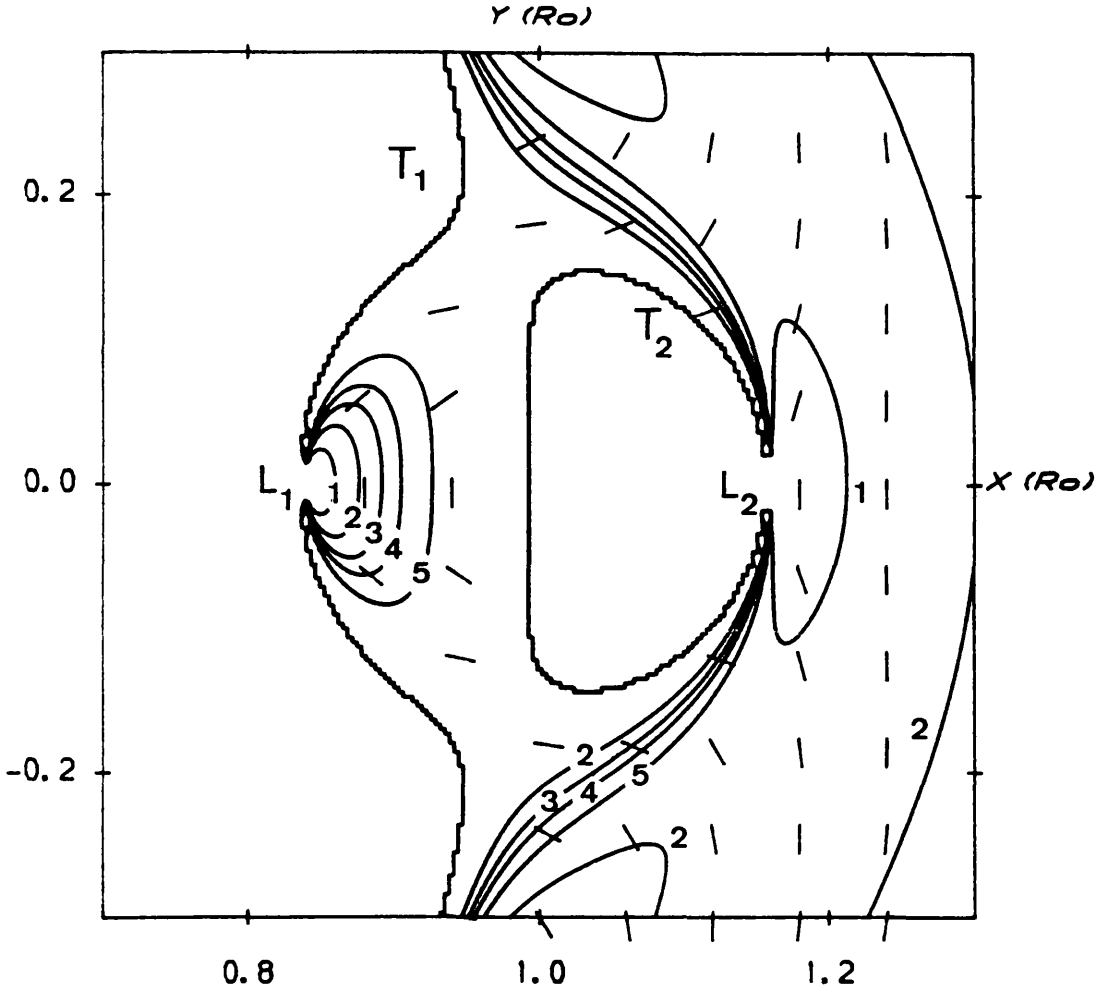


Figure 6.6(a)



Section of the level surfaces in the near lunar region (a) normal to the plane of the system and (b) in the plane of the system at time $\Omega_* t = 180^o$. The sail loading values are given by; (1) 0.3 (2) 0.6 (3) 1.0 (4) 1.5 (5) 3.0. The contours T_1 and T_2 represent the time dependent boundary surfaces.

Figure 6.6(b)



Then, since $\nabla U(\mathbf{r}_L)=0$ and $\partial \mathbf{a}/\partial \mathbf{r}=0$ (since the radiation field is assumed to be uniform) a linear variational system is obtained, viz

$$\frac{d^2\boldsymbol{\varsigma}}{dt^2} + \mathbf{M}_1 \frac{d\boldsymbol{\varsigma}}{dt} + \mathbf{M}_2 \boldsymbol{\varsigma} = \mathbf{a} \tag{6.44}$$

where $\mathbf{M}_2$, the gravity gradient tensor and the skew symmetric gyroscopic matrix $\mathbf{M}_1$ are given by

$$\mathbf{M}_1 = \begin{bmatrix} 0 & -2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{M}_2 = \underset{(i,j)\in(x,y,z)}{(U^o{}_{ij})} \tag{6.45}$$

where $U^o{}_{ij}$ is the (i,j) partial derivative of the potential with respect to the cartesian axes, evaluated at the L_2 point.

The sail attitude is now fixed such that the sail normal points along the Sun-line, but is pitched at an angle γ to the plane of the system, Figure 6.7. With this choice of sail attitude control equation (6.44) may be written in component form as

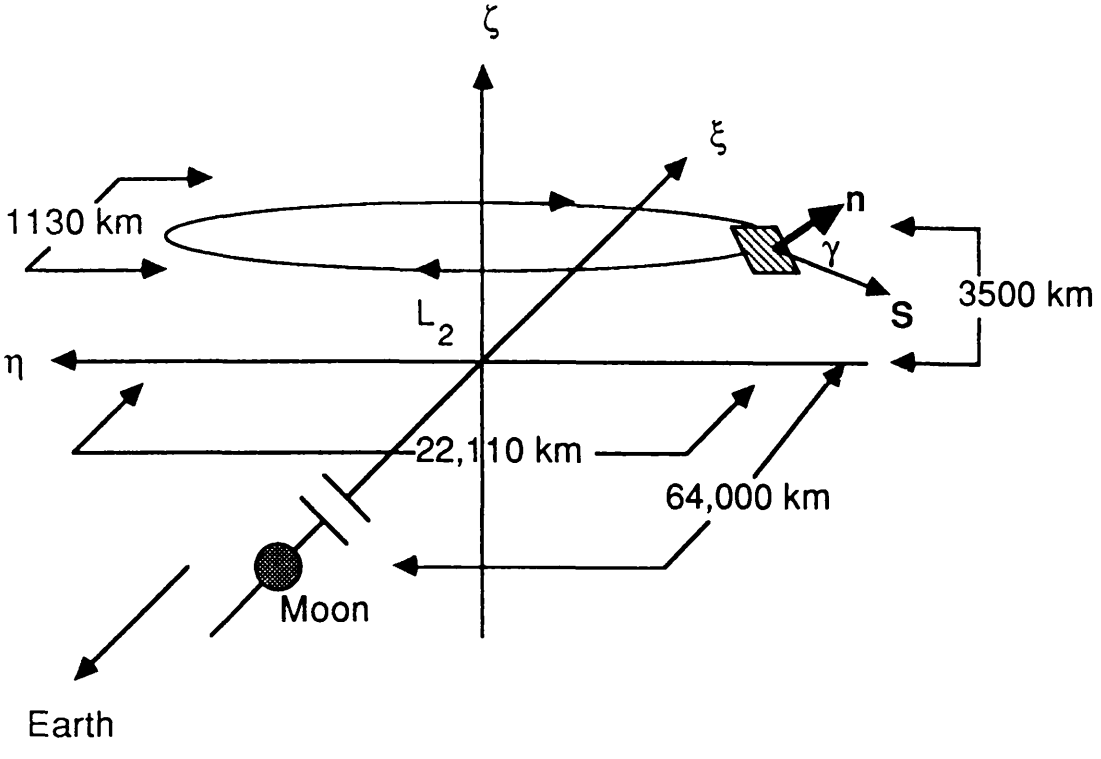
$$\frac{d^2\xi}{dt^2} - 2\frac{d\eta}{dt} + U^o{}_{xx}\xi = \beta\cos(\Omega_* t)\cos^3\gamma \tag{6.46a}$$

$$\frac{d^2\eta}{dt^2} + 2\frac{d\xi}{dt} + U^o{}_{yy}\eta = -\beta\sin(\Omega_* t)\cos^3\gamma \tag{6.46b}$$

$$\frac{d^2\zeta}{dt^2} + U^o{}_{zz}\zeta = \beta\cos^2\gamma\sin\gamma \tag{6.46c}$$

The complete solution to equations (6.46) will in general have divergent modes giving unbound motion. Therefore, since the instability timescale at the L_2 point is 12.7 days, active control using a feedback to the sail attitude is required to suppress these modes. A particular periodic in-plane solution will now be required, viz

Figure 6.7



A periodic out-of-plane solar sail trajectory at the lunar L_2 Lagrange point. The sail is pitched at an angle γ to the Sun-line **S**.

$$\xi(t) = \xi_0 \cos(\Omega_* t) \quad , \quad \eta(t) = \eta_0 \sin(\Omega_* t) \tag{6.47}$$

This periodic solution is now substituted in equations (6.46) to yield

$$\begin{Bmatrix} \xi_0 \\ \eta_0 \end{Bmatrix} = - \begin{Bmatrix} \dfrac{\Omega_*^2 + 2\Omega_*^2 - U_{yy}^0}{\Omega_*^2 + 2\Omega_*^2 - U_{xx}^0} \end{Bmatrix} \tag{6.48}$$

Therefore, since $U_{xx}^0 \neq U_{yy}^0$ the trajectory will be an ellipse centred on the L_2 point. The required sail loading may also be obtained as

$$\beta = \cos^{-3}\gamma \left\{ \frac{\Omega_*^4 - \Omega_*^2(U_{xx}^0 + U_{yy}^0 + 4) + U_{xx}^0 U_{yy}^0}{U_{yy}^0 - 2\Omega_* - \Omega_*^2} \right\} \xi_0 \tag{6.49}$$

The uncoupled out-of-plane motion defined by equation (6.46c) may now be solved by Laplace transforms to give a general solution of the form

$$\begin{aligned} \zeta(t) = {} & \zeta_0 \cos(\omega t) + \left\{ \frac{d\zeta}{dt} \right\}_0 (U_{zz}^0)^{-1/2} \sin(\omega t) \\ & + \beta \cos^2\gamma \sin\gamma \, (U_{zz}^0)^{-1} (U(t) - \cos(\omega t)) \end{aligned} \tag{6.50}$$

where $U(t)$ is the unit step function and $\omega = (U_{zz}^0)^{1/2}$. Choosing $(d\zeta/dt)_0 = 0$ the solution can be more conveniently expressed as

$$\begin{aligned} \zeta(t) = {} & U(t) \, \beta \cos^2\gamma \sin\gamma \, (U_{zz}^0)^{-1} \\ & + \cos(\omega t)(\zeta_0 - \beta \cos^2\gamma \sin\gamma \, (U_{zz}^0)^{-1}) \end{aligned} \tag{6.51}$$

It can be seen from this form of the solution that once the sail is pitched from $\gamma = 0$ at $t = 0$ the motion is of the form of periodic oscillations at an out-of-plane distance $\beta \cos^2\gamma \sin\gamma (U_{zz}^0)^{-1}$. However, by choosing the initial out-of-plane distance $\zeta_0 = \beta \cos^2\gamma \sin\gamma (U_{zz}^0)^{-1}$ the sail remains at this fixed distance. The ζ component of the solar radiation pressure acceleration, and so the out-of-plane distance, may be maximised by an optimal choice of sail pitch angle γ^*, viz

$$\frac{d}{d\gamma}\beta\cos^2\gamma\sin\gamma = 0 \Rightarrow \gamma^* = \tan^{-1}(2^{-1/2}) \qquad (6.52)$$

Using this optimal pitch angle (γ^*=35.264°) the sail may execute an out-of-plane elliptical trajectory centred at the L_2 point. The spacecraft may be placed on such a trajectory by inserting it into a suitable elliptical path about the L_2 point. Once the sail is pitched to an angle γ^* the out-of-plane oscillations may be damped in a time optimal manner.

At an out-of-plane distance of $3.5\text{x}10^3$ km both the lunar far-side and the equatorial regions of the Earth would be visible, requiring a very low sail acceleration of order 0.2 mms^{-2}. The applications of such a trajectory will be discussed in chapter 7. The trajectory itself would be a narrow ellipse with semi-major and minor axes of $1.105\text{x}10^4$ km (n_o=$2.876\text{x}10^{-2}$) and $5.655\text{x}10^2$ km (ξ_o=$1.471\text{x}10^{-3}$) and a period of 29.53 days (synodic lunar month), Figure 6.7. Since trajectories in the neighbourhood of the lunar L_2 point are naturally unstable active control is required to ensure damping of the divergent modes. Also, the neglected non-linear terms in equations (6.46) will perturb the spacecraft from its nominal elliptical trajectory.

6.10 Conclusions

It has been demonstrated that circular restricted three-body systems for solar sail spacecraft have an infinite set of new stationary solutions. For the Earth-Sun system these new solutions appear as level surfaces of constant sail loading around the classical Lagrange points. A linear stability analysis shows that the solutions are in general unstable, apart from the particular solutions with the sail oriented along the Sun-line. However, since it is found that the

system is controllable, asymptotic stability may be ensured through the use of a feedback control to the sail attitude.

For the Earth-Moon system the surfaces of constant sail loading become time dependent due to the synodic rotation of the Sun-line. Therefore, true stationary solutions are not in fact possible. However, since the required sail attitude is time independent small trims in the sail loading allow the sail to remain at a stationary point for a short duration. Lastly, by linearising the dynamical equations about the lunar L_2 point, periodic out-of-plane trajectories are obtained. These trajectories can be achieved with a relatively low sail loading, although active control is required due to the natural instability of trajectories about the collinear Lagrange points.

7. ADVANCED TRAJECTORY APPLICATIONS

7.1 Heliocentric Halo Orbit Applications

The dynamics, stability and control of heliocentric halo orbits were discussed in chapter 4 where it was shown that unique heliocentric solar sail trajectories are possible. The applications of these advanced trajectories are particularly interesting for solar system space science missions. Using heliocentric halo orbits unique observations of the Sun and interplanetary dust complex may be made.

7.1.1 Polar Stationed Solar Sail Missions

The simplest heliocentric solar sail mission utilises the solar radiation pressure force to balance the solar gravitational force so that the sail will remain stationary above the solar poles, or any other point on the solar surface. This case is a degenerate halo orbit with zero halo amplitude. The dynamics of this system have been investigated in detail in section 3.3 where it was shown that a total spacecraft mass per unit area of 1.53 gm^{-2} is required. For a small 10^2 kg payload and an advanced sail mass per unit area of 1 gm^{-2}, sail material and structure, a 434x434 m square sail is required. By being stationed above the solar poles continuous observations of the poles would be possible allowing the out-of-plane solar polar magnetic field and solar wind to be investigated. When combined with in-plane data obtained from spacecraft orbiting in the ecliptic plane this would allow the full three-dimensional structure of the solar magnetic field and wind to be mapped. This task is a primary objective of the Ulysses international solar polar mission which will spend a total of 230 days above a heliographic latitude of 70^{o} with a perihelion

distance of 1.28 AU, Wenzel et. al (1990). For a polar stationed solar sail however, truly continuous observations could be made from a close heliocentric distance.

Observations of the solar magnetic field near the poles involves less rotational twist of the field lines than near the equator and so yield a more direct indication of their connection to the solar surface. Similarly, the solar wind flow above the poles is less distorted. In the ecliptic plane the high and low speed solar wind flows mix due to their relative radial speeds and their azimuthal motion, induced by the solar rotation. The high speed flows overtake the low speed flows and are mixed due to the solar rotation leading to a complex mixed plasma flow. Over the solar poles however the solar wind flow is expected to be parallel to the near radial magnetic field lines giving a less complex flow which will be easier to understand. Furthermore, during the quiet part of the eleven year solar cycle coronal holes appear in the polar regions. These regions of low magnetic field are sources for energetic solar wind streams, with speeds of up to 10^3 kms^{-1}. Since the solar wind flows radially from coronal holes the flows can only be investigated from directly above the holes. Therefore, solar sail spacecraft positioned at high heliographic latitudes are advantageous.

Observations of the distribution of interplanetary dust from spacecraft orbiting in the ecliptic plane allow the in-plane density to be calculated. However, the out-of-plane distribution can only be obtained remotely from the observed brightness integrated along the line of sight and polarisation of scattered radiation. The dust density must then be obtained by data inversion techniques. By obtaining brightness and from polarisation observations out of the ecliptic plane the full three-dimensional structure of the dust complex may be easily

obtained, Dumont et. al (1980). Other unique observations possible from a polar stationed solar sail are uncorrupted measurements of cosmic rays. As cosmic rays pass through the solar system their trajectories are perturbed by the solar magnetic field. Similarly their flux at the Earth is modulated by the eleven year solar cycle. Along the polar axis of the Sun however, where the magnetic field lines are more radial, cosmic ray particles may penetrate more easily into the inner solar system, Wenzel et. al (1990).

7.1.2 Heliosynchronous Solar Sail Missions

At a heliocentric distance of 0.167 AU a spacecraft in the ecliptic plane will have a Keplerian orbital period of 25 days and so will follow the solar equatorial rotation, allowing the possibility of unique solar observations, Christensen-Dalsgaard et. al (1984). However, the heliosynchronism may be maintained at distances of less than 0.167 AU by decoupling the spacecraft orbital period from its heliocentric distance. This is achieved by using a solar sail spacecraft so that the solar radiation pressure force reduces the effective local gravitational force and so lengthens the orbital period. This is a degenerate halo orbit with zero out-of-plane displacement distance. The dynamics of these trajectories have been investigated in detail in section 3.4.

By following the solar rotation from close heliocentric distances, less than 0.167 AU, active regions on the solar surface may be followed and their evolution studied in detail. With terrestrial observations, or indeed with Earth orbiting or Lagrange point spacecraft, events may only be studied until they move across the solar disk and disappear behind the solar limb. By using a network of several heliosynchronous solar sail spacecraft, possibly in

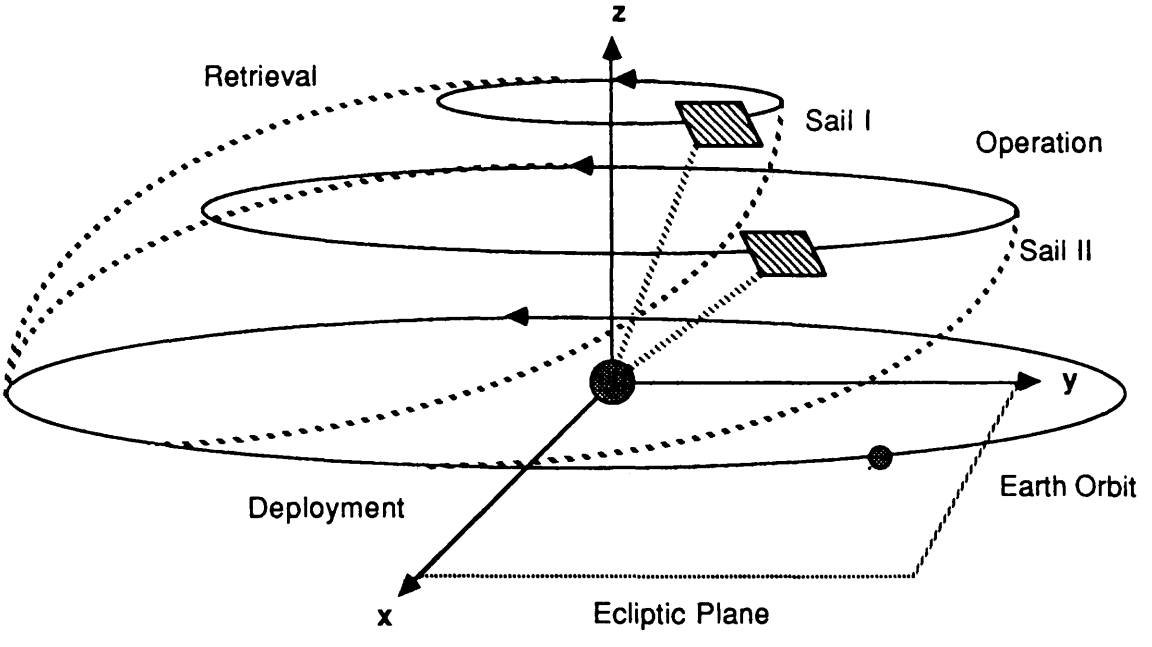
combination with polar stationed sails, global coverage and monitoring of the Sun would be possible. This would also allow stereoscopic views of the solar atmosphere and the multiple lines of sight would allow limb occultation studies of the differential height structure in the solar atmosphere, Kane (1982).

7.1.3 Halo Orbit Solar Sail Missions

As discussed in section 7.1.1 out-of-plane observations have many unique advantages both alone and when convolved with observations obtained from in-plane spacecraft. Stereoscopic imaging of the Sun at various wavelengths and stereo observations of the interplanetary dust complex may provide insights into many aspects solar system physics. By utilising solar sail spacecraft on heliocentric halo orbits continuous observations of the temporal and three-dimensional spatial structre of the heliosphere may be made.

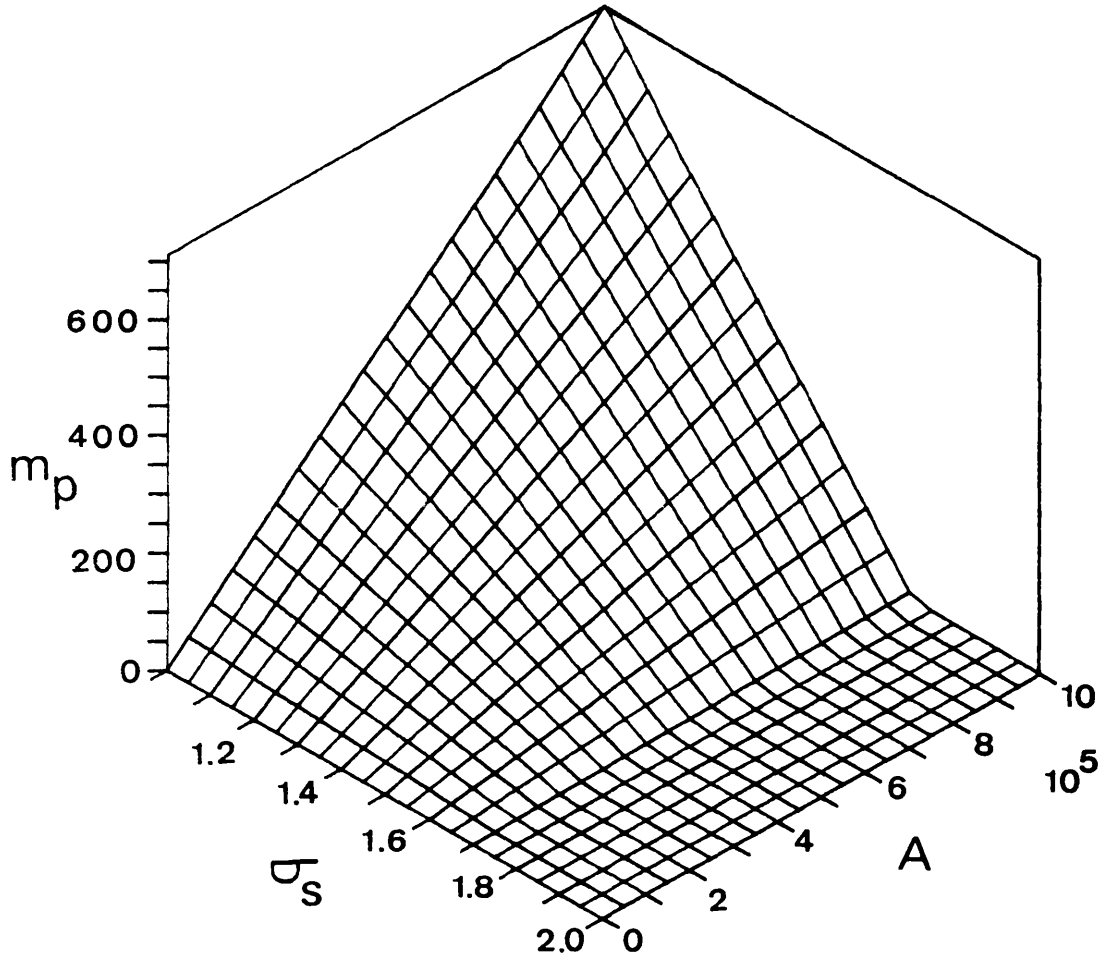
A mission with twin solar sails for such out-of-plane observations is shown schematically in Figure 7.1. The spacecraft are chosen to be on one year Earth synchronous halo orbits with ρ=0.56 AU, z=0.56 AU, β=0.9 (sail I) and ρ=0.88 AU, z=0.19 AU, β=0.5 (sail II). The loading required for sail I is extremely high, three times that of current solar sail designs, but could be achieved using the advanced fabrication techniques discussed in section 1.4.3. For a small payload mass of 10^2 kg and a sail mass per unit area, sail material and structure, of 1 gm^{-2} a 380x380 m square sail is required. A trade-off between payload mass, the sail total mass per unit area and sail area is shown in Figure 7.2, assuming an overall 98% sail efficiency. It can be seen that the sail area can be extremely sensitive to payload mass. However, payloads of up to 700 kg are possible with large 1x1 km

Figure 7.1



Schematic profile of a twin solar sail mission for three-dimensional solar observations. Both sail I and sail II are chosen to be on Earth synchronous halo orbits.

Figure 7.2



Surface of payload mass (kg) as a function of the sail area (m^2) and the sail total mass per unit area (gm^{-2}).

sails. The region with zero payload mass corresponds to the sail total mass per unit area, sail material and structure, being greater than the total spacecraft mass per unit area required to establish the halo orbit. Since the sail attiude is fixed, apart from a slow annual rotation to maintain a Sun pointing orientation, operational structural loads would be relatively low. Similarly, a small payload mass does not require a large sail structure to support it. In fact the payload may be distributed over the sail structure to reduce the relatively large structural loads experienced with a centralised payload.

Due to the large sail loading required for the two halo orbits direct spiralling out of the ecliptic plane would be possible for deployment of the two spacecraft, with transfer times of a few months only. Alternatively, direct injection with conventional upper stages would be possible so that the sails may be configured for halo orbit operation only. Once operational the twin spacecraft would have a relative viewing angle of the Sun of 33^{o} allowing stereoscopic observations. Furthermore, due to the one year orbital period there would be a constant path length, and so constant bit rate, for data returns. Apart from optical stereographic imaging of the solar atmosphere observations of anisotropic x-ray emissions from two different angles would allow anisotropic particle beams in the solar atmosphere to be investigated, Kane (1982). Similarly, the separation of the two spacecraft of nearly 0.5 AU allows a large baseline for the investigation of the three-dimensional structure of the solar magnetic field.

Since the upper halo orbit (sail I) has equal amplitude and displacement distances it would be possible for sail I to be patched to other halo orbits, as discussed in section 4.8.1. By patching to a halo

orbit perpendicular to the ecliptic plane sail I would fall below the ecliptic in six months. The sail could then be patched to another halo orbit below the ecliptic plane, but retrograde to the initial halo orbit. Alternatively, the sail could be patched to a Keplerian ellipse to allow closer in-situ solar observations at various distances above and below the ecliptic plane. For sail I this Keplerian ellipse would have a period of 147 days and a perigee distance of only 0.11 AU allowing close heliocentric observations. These patching maneouvres could be implemented at the end of the nominal twin solar sail mission above the ecliptic plane. Once the entire mission is complete the two spacecraft may be returned to Earth orbit for refurbishment and re-use.

Along with pure space science applications heliocentric halo orbits have applications as repeater stations for high bit rate interplanetary communications. The design and operation of such repeater satellites deployed on Keplerian orbits has been investigated by Mercader del Rio (1989). It was proposed that a heliocentric ring of satellites could be used to ensure communications with spacecraft on the opposite side of the Sun from the Earth. This would however be possible using just a single polar stationed solar sail.

7.2 Geocentric Halo Orbit Applications

Geocentric halo orbits were discussed in chapter 5 where it was demonstrated that solar sail spacecraft could achieve geocentric circular orbits normal to the ecliptic plane, but displaced in the anti-Sun direction. Although large families of unstable orbits were found to exist it was shown that they may be stabilised with a suitable control scheme. Geocentric halo orbits have many interesting

applications for near Earth space science missions, investigating the interaction of the solar wind with the geomagnetic field.

7.2.1 Static Equilibrium Solar Sail Missions

The simplest form of the geocentric halo orbit is the static equilibrium case, a degenerate halo orbit with zero orbital period. With the sail normal directed along the Earth-sail line the solar radiation pressure force balances the local gravitational force. The required sail loading then varies as z^{-2} and is independent of ρ, as discussed in section 5.2. For large geocentric distances however, the full three-body analysis of chapter 6 is required.

At a distance of 40 R_o along the Sun-line the required sail characteristic acceleration is 6.13 mms^{-2}, independent of the distance ρ. For a 10^2 kg payload and a sail mass per unit area, sail material and structure, of 1 gm^{-2} a 450x450 m square sail is required. A space plasma physics payload onboard such a sail may then be positioned in the geomagnetic tail and so may provide continuous observations of the tail. When combined with data obtained from near Earth spacecraft the propagation of disturbances along the geomagnetic tail may be observed. Since the sail is stationary, variations in magnetic field are purely temporal, whereas for a single orbiting satellite it is difficult to de-convolve temporal and spatial variations. Since the required sail loading is independent of the distance ρ, the spacecraft may be quasi-statically transferred from near the centre of the tail through the plasma sheet to the edge of the tail at the magnetopause boundary, thereby obtaining a cross-sectional profile of the tail. Furthermore, the 4^o aberrational tilt of the geomagnetic tail with respect to the Sun-line, due to the heliocentric motion of the Earth,

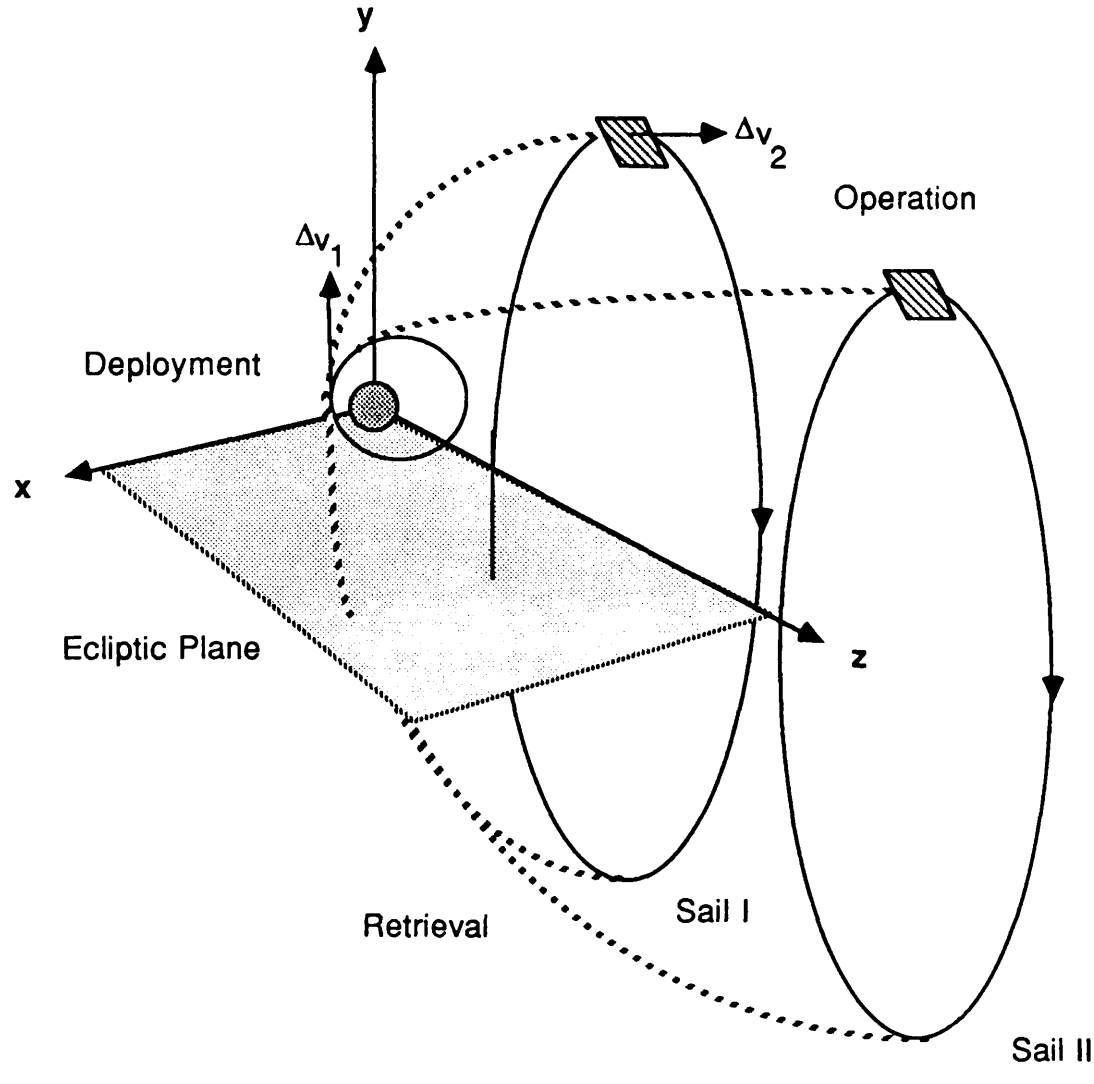
can be accommodated by a simple off-set of the stationary position.

7.2.2 Halo Orbit Solar Sail Missions

The investigation of the geomagnetic tail using solar sail spacecraft deployed on geocentric halo orbits allows the possibility of several unique missions. One such mission using twin solar sails is shown schematically in Figure 7.3. The amplitude ρ of the two halo orbits is chosen to be 20 R_o so that they are stationed on the magnetopause boundary. The required sail performance is then kept to a reasonable level by choosing optimal halo orbits of large displacement distances z of 40 R_o for sail I and 50 R_o for sail II. The separation of 10 R_o allows the possibility of observing the propagation of disturbances along the geomagnetic tail. Since optimal halo orbits are chosen the spacecraft will have differing orbital periods, with a period of 17.52 days for sail I and 23.15 days for sail II. The two spacecraft will therefore drift relative to each other in azimuthal position with a synodic period of 72.04 days. This drift may however be overcome by using polar synchronous halo orbits, but at the expense of increased sail loading requirements. For optimal halo orbits with 10^2 kg payloads and a sail total mass per unit area of $1 gm^{-2}$, sail I requires a 304x304 m square sail and sail II requires a 229x229 m square sail.

As with the heliocentric case the spacecraft may be transferred to their operational orbits by spiralling from low Earth orbit. However, the relatively long spiralling times from low Earth orbit, due to the high local gravitational acceleration, suggests that direct ballistic transfer should be considered. The transfer is made by applying an impulse Δv_1 with an upper stage to the undeployed sail

Figure 7.3



Schematic profile of a twin solar sail mission for three-dimensional observations of the geomagnetic tail. Both sail I and sail II are chosen to be on optimal halo orbits and are deployed using two impulses $\Delta v_{1,2}$.

on a low Earth parking orbit at geocentric distance r_o. The apogee of the resulting Keplerian ellipse is chosen to be equal to the halo orbit geocentric distance r, viz

$$\Delta v_1 = \left\{\frac{2\mu}{r_o}\left\{\frac{r}{r_o + r}\right\}\right\}^{1/2} - \left\{\frac{\mu}{r_o}\right\}^{1/2} \qquad (7.1)$$

Once the sail-upper stage stack is at the apogee point another small impulse is applied to attain the velocity $\rho\Omega_*$ required for the optimal halo orbit. The required impulse Δv_2 is then given by

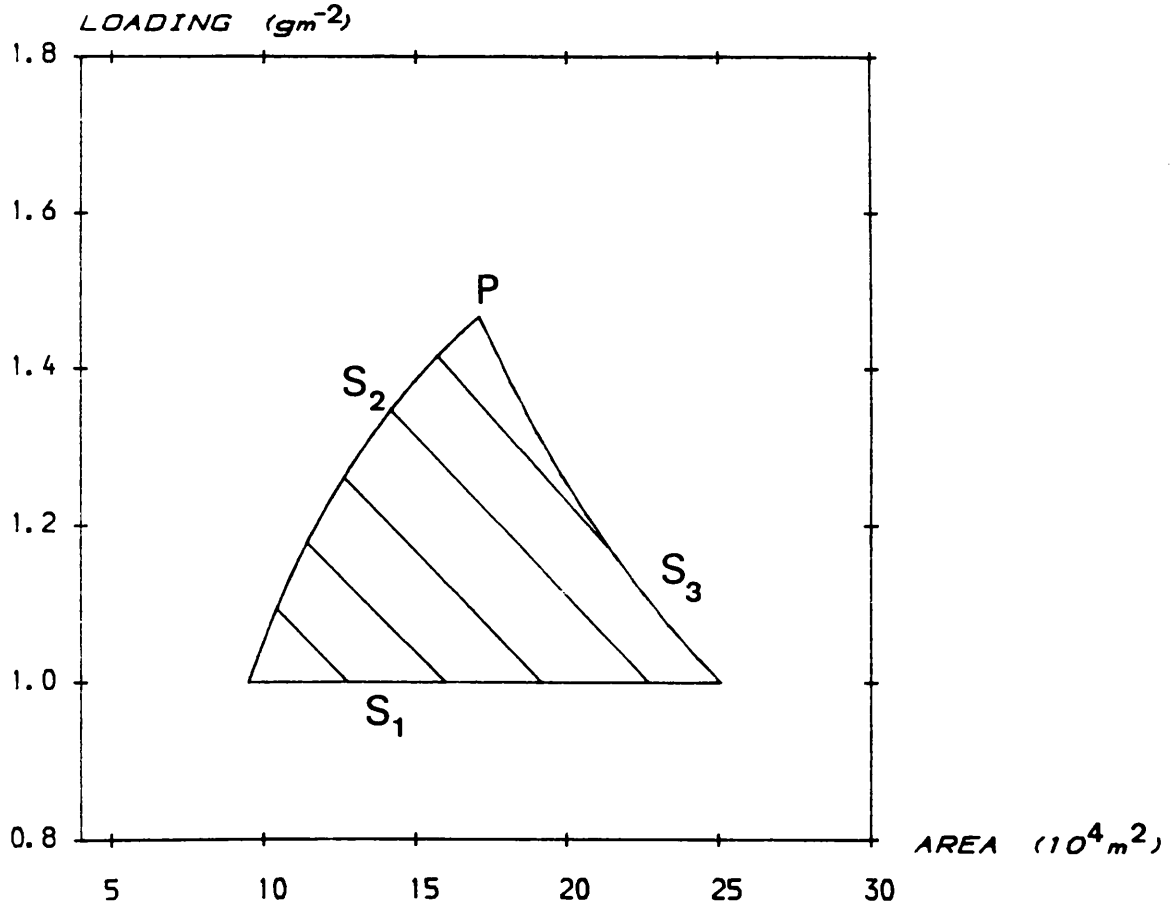
$$\Delta v_2 = \rho\left\{\frac{\mu}{r^3}\right\}^{1/2} - \left\{\frac{2\mu}{r}\left\{\frac{r_o}{r_o + r}\right\}\right\}^{1/2} \qquad (7.2)$$

When the final impulse is applied the sail is deployed and pitched to a Sun facing attitude. For sail I an initial impulse Δv_1 of 3.08 kms^{-1} and a final impulse Δv_2 of 0.28kms^{-1} is required for transfer from a 300 km parking orbit. The final mass m injected into the halo orbit is easily obtained from

$$m = m_o \exp\left\{\frac{-(\Delta v_1 + \Delta v_2)}{c}\right\} \qquad (7.3)$$

where m_o is the initial mass in low Earth orbit and c is the upper stage exhaust velocity. This final mass includes the total sail mass and upper stage booster dry mass. A trade-off between the sail total mass per unit area and the sail area is shown in Figure 7.4 for sail I using the above analysis. A maximum initial low Earth orbit mass of 2100 kg was assumed along with typical upper stage parameters, viz a 300 kg dry mass and 2.84 kms^{-1} exhaust velocity, Timnat (1987). The boundary S_3 then represents the maximum allowed mass in low Earth orbit. To the right of S_3 the total spacecraft and upper stage mass

Figure 7.4



Section of a surface of payload mass as a function of the sail area and the sail total mass per unit area. The curve S_2 is a contour of constant payload mass of 10^2 kg, with the point P the optimal design point on this contour. The contour S_3 represents the maximum low Earth orbit mass with S_1 representing the minimum sail total mass per unit area.

in low Earth orbit is greater than 2100 kg. Similarly, to the left of the payload mass contour S_2 the payload mass falls below a minimum of 10^2 kg. The boundary S_1 represents an imposed lower limit on the sail material and structural mass per unit area of 1 gm^{-2}. The allowed design point of the spacecraft then lies in the region defined by the intersection of these three curves. The optimal design point P lies on the intersection of S_2 and S_3. This point has a 10^2 kg payload mass but minimises the requirements on the sail material and structural design.

Once the twin solar sails have been deployed on their nominal halo orbits continuous observations of the geomagnetic tail may be made at fixed distances along the tail. For investigations using conventional ballistic trajectories long, highly elliptical orbits must be used in order to maximise the time spent in the tail, as is the case for the proposed Geotail mission, Dunham (1989). When combined with observations from near Earth orbiting spacecraft the propagation of disturbances along the geomagnetic tail may be observed and the full three-dimensional structure of the tail mapped. By using two spacecraft purely temporal variations in the observations may be de-convolved from spatial variations. The 4^o aberrational tilt of the tail may be compensated for by moving the halo orbits off-axis by the same angle. The increase in sail loading due to the oblique incidence of the incoming radiation is only 0.48 %.

At the end of the nominal twin spacecraft mission sail I (or sail II) may be patched to a Keplerian ellipse to investigate the inner regions of the geomagnetic tail, much in the same way as a conventional ballistic spacecraft. From the halo orbit of sail I the spacecraft may be patched to a 12.01 day Keplerian ellipse with a

perigee distance of 24.85 R_O, taking the payload through the bow shock region. However, unlike conventional ballistic trajectories the orientation of the ellipse may be chosen so that the polar or equatorial regions of the near Earth tail are traversed. That is, by switching the sail attitude into a null orientation as the spacecraft passes through the ecliptic plane the orbital plane of the resulting ellipse will be normal to the ecliptic. Similarly by switching the sail attitude at the upper arc of the halo orbit the line of nodes of the resulting ellipse will lie in the ecliptic plane. Therefore, observations of sections of the geomagnetic tail may be made. As with the heliocentric case, at the completion of the full mission the two spacecraft may be returned to low Earth orbit.

Not only can investigations of the geomagnetic tail be made using geocentric halo orbits but the magnetic tails of other planets can be investigated, with substantially lower sail loadings. Furthermore, by using a solar sail on an off-axis halo orbit in combination with a spacecraft orbiting on a low altitude polar ellipse real time stereographic imaging of a planetary surface may be made. Displacement distances of several thousand kilometers are required for a large relative viewing angle between the two spacecraft. Therefore, this application is extremely demanding for the geocentric case but is achievable for other inner solar system bodies.

7.3 Three-Body Stationary Solution Applications

The three-body dynamics of solar sail spacecraft in the Earth-Sun and Earth-Moon systems were discussed in chapter 6. It was shown that for the Earth-Sun system stationary solutions exist around the classical Lagrange points. For the Earth-Moon system no

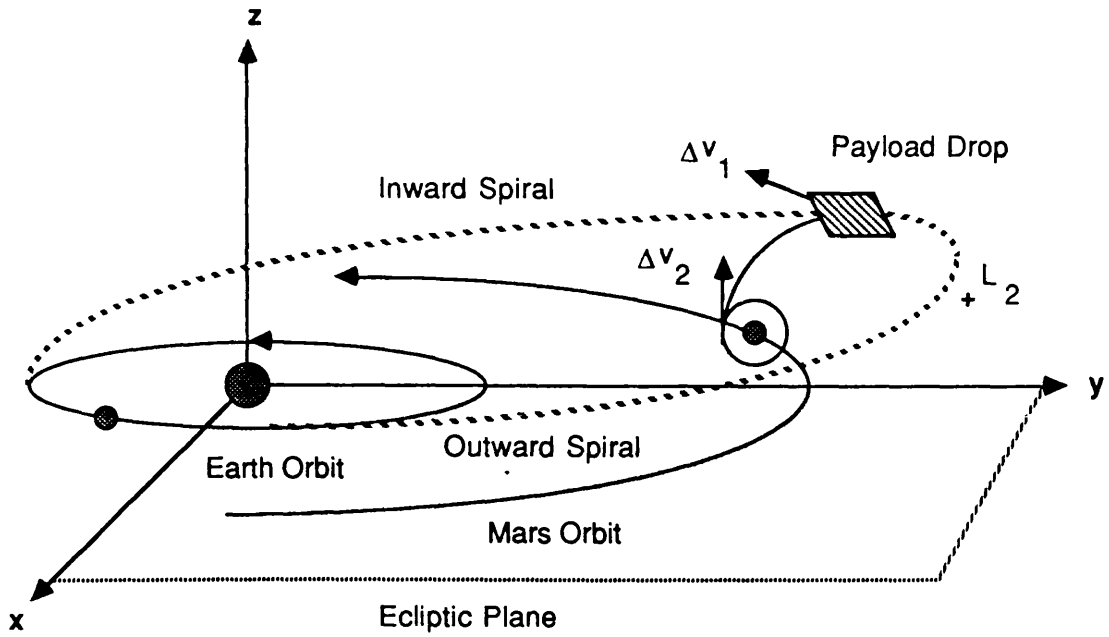
truly time independent stationary solutions exist. However, with small trims in the sail loading short duration stays were shown to be possible. These stationary solutions have potential applications for small space science missions and for the staging of large interplanetary missions.

7.3.1 Stationary Solutions in the Earth-Sun System

For interplanetary solar sail missions a large fraction of the total transfer time is contained in the initial and final planetocentric spirals to and from low planetary orbit and escape. This element of the total transfer time can however be eliminated by parking the solar sail near the planetary L_2 point on a level surface of constant sail loading equal to the total spacecraft loading. At the end of an interplanetary trajectory the spacecraft would transfer its payload to a low planetary orbit with an impulse from a chemical motor with space storable propellants, Figure 7.5. From an out-of-plane stationary point high planetary lattitudes may be easily reached for surface landers. For conventional schemes the payload is injected into a planetary parking orbit in the ecliptic plane. To reach the planetary polar regions then requires large plane changes and so a large Δv maneouvre. Furthermore, the window for transfer to the planetary surface is always open from a stationary point. For a parking orbit the window occurs only twice per planetary day when the parking orbit crosses the landing site.

The solar sail parking scheme also allows the sail to be optimally designed for interplanetary space. Planetocentric spiral trajectories require large attitude turning rates which puts demands on the sail structure and attitude control mechanisms. Similarly, for sample

Figure 7.5



Schematic profile of a Mars mission using artificial stationary point staging. The payload is transferred to low a planetary orbit or to the planetary surface from a level surface near the L_2 point using two impulses $\Delta v_{1,2}$.

returns, if the sample was transferred to the parked sail by chemical means the long spiral out of the planet's gravity well could be avoided. After the initial payload drop to the planetary surface the sail would have an increased loading and so would be stationed on a level surface closer to the planet to await the sample return from the planetary surface. Of course, there is a mass penalty imposed due to the need for propellant to be transported to the destination. However, the shortened total mission duration may more than compensate.

At the Earth-Sun L_1 point a number of spacecraft, such as ISEE-3, Farquhar (1980), have been, or will be, positioned for solar observations and for 'upwind' observations of the solar wind to be combined with data from near Earth orbiting spacecraft. However, spacecraft positioned directly at the L_1 point would appear in the centre of the solar radio disk when observed from Earth and so would be unable to make data returns. This problem is overcome however by forcing the spacecraft to execute a ballistic Lagrange point halo orbit. This highly unstable trajectory is a periodic orbit normal to the ecliptic plane so that, when viewed from Earth, the spacecraft appears to orbit around the solar radio disk allowing data returns. To maintain the Lagrange point orbit however requires regular station-keeping maneouvres so that the on-board propellant mass ultimately determines the mission lifetime.

Using a relatively small sail a payload could be positioned at the L_1 point, but simply displaced above the ecliptic plane to avoid the solar radio disk. At a distance of 1.2×10^5 km above the ecliptic plane and 7.2×10^4 km sunward of the L_1 point a minimised total spacecraft mass per unit area of 205 gm^{-2} is required. For a low performance sail of 5 gm^{-2} and a 10^3 kg payload only a small 70x70 m square sail

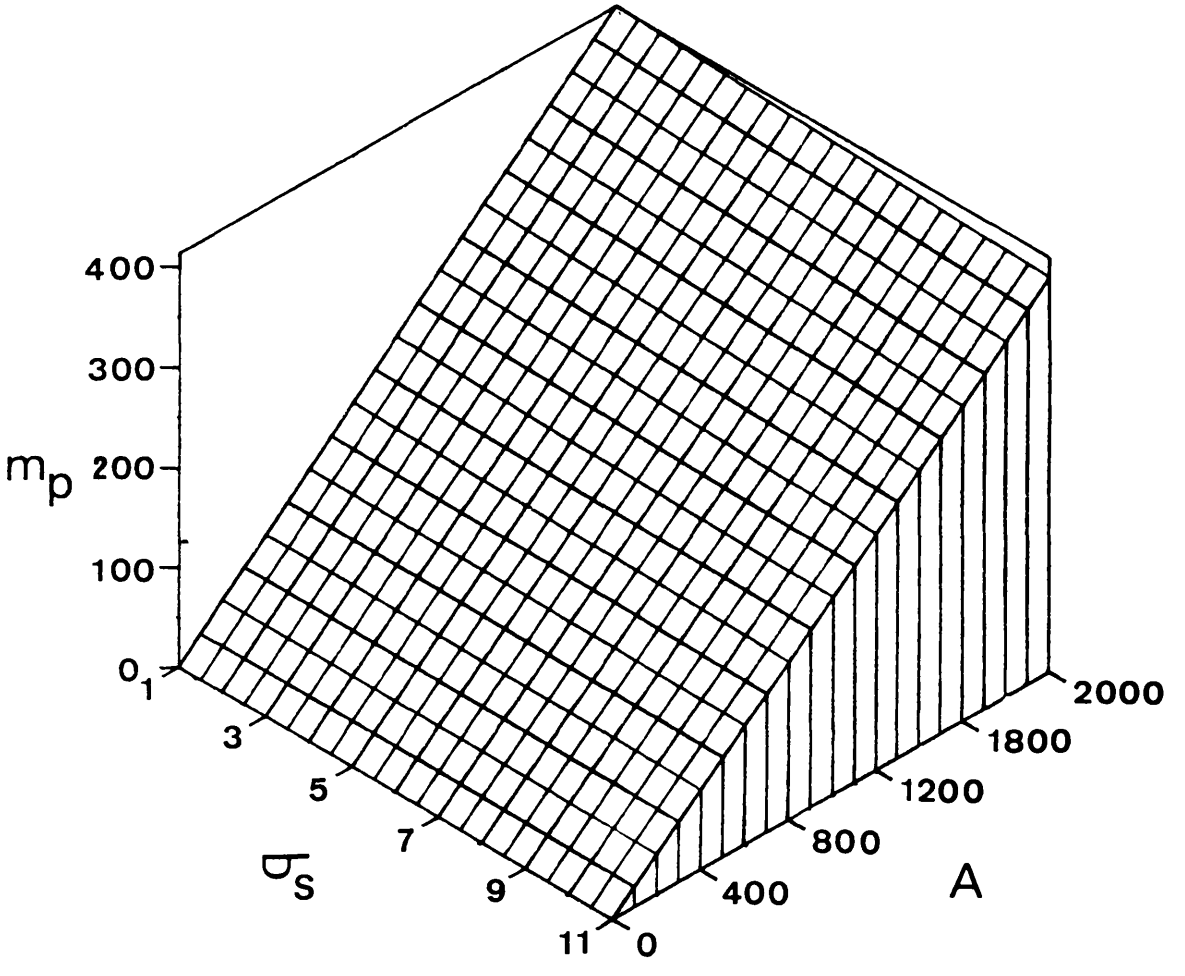
is required. A trade-off between payload mass, sail area and the sail total mass per unit area is shown in Figure 7.6. It can be seen that for low payload masses extremely small sails are required. Since station-keeping may be performed using trims on the sail attitude no propellant is required and the mission lifetime is limited only by the lifetime of the payload and the sail. This is also of importance for potential payloads as chemical exhaust plumes may corrupt solar spectral data. Since the solar sail is stationary, instrument pointing can be achieved more easily than with an orbiting ballistic spacecraft. Furthermore, the spacecraft position and velocity may be determined with greater accuracy, as would be required for accurate doppler observations of spectral lines, such as those to be carried out by the Soho spacecraft, Domingo et. al (1985).

Other space science missions utilising three-body solar sail stationary solutions would require a solar sail to be stationed near the L_2 point to investigate the far regions of the geomagnetic tail. Since the sail would be stationary temporal and spatial variations in the observations may be de-convolved. The full three-dimensional structure of the tail may then be investigated by transferring from in-plane to out-of-plane stationary points. Finally, applications for communications purposes have been made by Forward (1991) as an extension of the 'statite' concept, as discussed in section 1.5.3, to large geocentric distances where a full three-body analysis is required.

7.3.2 Stationary Solutions in the Earth-Moon system

The staging scheme proposed in section 7.3.1 may also be applied to the Earth-Moon system. To reduce long spiralling times from low

Figure 7.6



Surface of payload mass (kg) as a function of the sail area (m^2) and the sail total mass per unit area (gm^{-2}).

Earth orbit to low lunar orbit the sail may be parked on a suitable level surface of constant sail loading near the lunar $L_{1,2}$ points. Typically, a total spacecraft mass per unit area of 3 gm^{-2} is required for the level surfaces at these points. Small trims in the sail loading may be used to compensate for the motion of the Sun-line. The payload would then be transferred to low lunar orbit by an impulse from a chemical motor using space storable propellants. From a stationary point out of the plane of the system regions of high lunar lattitude are easily accessible.

The lunar L_2 point has been proposed as a location for repeater stations for lunar far-side communications. Such a communications path is essential for any future scientific investigation and utilisation of the lunar far-side. In order that both the lunar far-side and the Earth are visible from the repeater, the spacecraft must be forced to follow a ballistic Lagrange point halo orbit, as discussed in section 7.3.1. Alternatively the repeater may be displaced from the L_2 point out of the plane of the system using a low thrust propulsion system, Vonbun (1968). The lifetime of the spacecraft is again dictated by the onboard propellant mass. A similar scheme is also possible using the solar sail Lagrange point halo orbit discussed in section 6.9. With an out-of-plane displacement distance of 3.5×10^3 km both the lunar far-side and the Earth are visible from the solar sail halo orbit. The in-plane motion takes the spacecraft 1.1×10^4 km from the L_2 point so that it moves in a 20^o arc about the position of the L_2 point in the lunar sky with a period of 29.53 days. The sail position in the sky is therefore relatively constant when viewed from the lunar surface. Since no propellant is required the repeater lifetime is limited only by the lifetime of the payload and the sail. The required sail

characteristic acceleration of 0.2 mms^{-2} corresponds to a total spacecraft mass per unit area of 46 gm^{-2}. Therefore, even for a low performance sail with a total mass per unit area of 5 gm^{-2} a 10^3 kg payload may be deployed using a relatively small 158x158 m square sail.

7.4 Conclusions

It has been shown that solar sail spacecraft offer opportunities for unique space science missions both in geocentric and heliocentric space. Observations of solar active regions by synchronously orbiting solar sails and investigations of the full three-dimensional structure of the solar magnetic field using heliocentric halo orbits are possible. Similarly, for geocentric halo orbit missions unique investigations of the geomagnetic tail are possible.

Along with pure space science missions the artificial three-body stationary solutions may be used for the staging of interplanetary transfers. By eliminating the final planetocentric spiral shorter total transfer times appear to be possible as do lower mass sail designs, configured for interplanetary flight only. Other applications for communications purposes are possible with solar radio disk avoidance at the Earth-Sun L_1 point and lunar farside communications at the Earth-Moon L_2 point.

The unique aspect of many of these missions is a primary driver for the adoption of solar sailing for an actual mission. For example, only a solar sail spacecraft can make truly continuous out-of-plane observations of the Sun using a heliocentric halo orbit. Other low thrust propulsion systems can achieve such a trajectory but have a limited mission lifetime due to their finite onboard propellant mass.

Therefore, if such an out-of-plane mission was in fact scientifically desirable the only viable option would be to use solar sail spacecraft.

8. CONCLUSIONS AND FUTURE PROSPECTS

8.1 Review

The objective of this thesis has been to demonstrate that solar sail spacecraft may execute unique, advanced trajectories which have potential applications for space science missions. The various families of trajectories have been investigated and their stability characteristics studied. Although large sub-families of unstable trajectories were found to exist simple control schemes have been developed to ensure asymptotic stability. These families of advanced trajectories have also been extended by patching individual trajectories together.

A detailed model of the solar radiation pressure exerted on a planar solar sail has been constructed in chapter 3 by considering the Sun as an extended source of radiation. The resulting form of the solar radiation pressure is found to deviate from the inverse square form assumed in all previous studies of solar sail dynamics. This deviation leads to instabilities for solar sails in stationary and circular orbital configurations. This is of particular importance for heliosynchronously orbiting and polar stationed spacecraft. The effect of temporal variations of the solar luminosity has also been investigated. By modelling the variations with a simple sinusoidal function Mathieu's equation was obtained, with its well known conditions for stability and instability.

Following the investigation of polar stationed and heliosynchronous orbits the heliocentric halo orbit was developed in chapter 4. By pitching the sail such that a component of the solar radiation pressure force was directed out of the orbital plane, a

displaced, out-of-plane circular orbit was obtained. The geometry and period of the halo orbit were found to be a function of the sail loading parameter and the Sun-sail pitch angle. These functional relationships were obtained by considering the halo orbit to be a stationary solution to the dynamical equations in a co-rotating reference frame with an arbitrary co-rotation period. Extensions of the heliocentric halo orbit were obtained by patching four perpendicular halo orbits together to form a 'cubic' trajectory and by patching to a ballistic Keplerian orbit. By using a linear perturbation analysis stable and unstable families of halo orbits were identified. For the unstable families a simple feedback control to the sail pitch was found to ensure asymptotic stability. However, by allowing small trims in the sail loading parameter shorter damping timescales were obtained.

Using a similar co-rotating reference frame analysis the geocentric halo orbit was developed in chapter 5. By again directing a component of the solar radiation pressure force out of the orbital plane, a displaced out-of-plane circular orbit was obtained. It was demonstrated that geocentric halo orbits may be patched together or patched to ballistic Keplerian orbits to form complex new trajectories. The annual rotation of the Sun-line was shown to induce periodic perturbations which excited the unstable modes of the unstable halo orbit family. However, a simple feedback control was shown to stabilize these orbits.

The extension from two-body to three-body dynamical systems was made in chapter 6 where artificial stationary solutions in the Earth-Sun and Earth-Moon systems were investigated. By again considering stationary solutions in a co-rotating reference frame

extensions of the five classical Lagrange points were obtained for the Earth-Sun system. Rather than individual stationary points, extended surfaces around the classical Lagrange points were found to exist. It was then demonstrated that these new stationary solutions are unstable, but that it was in principle possible to obtain stability using a general feedback control to the sail attitude. Similarly, in the Earth-Moon system stationary solutions were obtained by trimming the sail loading parameter to compensate for the motion of the Sun-line during the synodic lunar month. At the lunar L_2 Lagrange point another periodic out-of-plane trajectory was found by linearising the dynamical equations about this point. The inherent instability of this trajectory necessitates the use of active control.

Finally, in chapter 7 possible applications of these advanced trajectories were discussed. Several scientifically interesting missions were identified and preliminary studies of the trade-off between payload mass and sail design made. Using twin solar sails deployed on heliocentric halo orbits stereoscopic solar observations were shown to be possible. Similarly, for geocentric applications a twin solar sail mission for geomagnetic tail observations was identified.

8.2 Dynamics and Control

The analysis of the various families of solar sail trajectories has assumed a perfectly reflecting, planar sail. Although this is a suitable approximation for a first order analysis any further detailed studies will require an accurate force model. Such force models exist from the JPL comet Halley design studies, however no such model exists for state-of-the-art solar sail design. Therefore, a model of an advanced solar sail design with a lightweight structure and a low mass per unit

area is required. Using a finite element simulation and estimating real material properties a force model which includes detailed effects such as thermal re-emission and billowing of the sail shape may be obtained. Similarly, it has been assumed that the sail attitude may be controlled with exact precision. Other simulations of the sail attitude control are therefore required to obtain a model of the responses to attitude control commands and the static pointing accuracy.

The dynamic models used for the heliocentric halo orbits and three-body stationary solutions are suitable and accurate. However, although the annual rotation of the Sun-line has been included in the analysis of the geocentric halo orbit dynamics, a more complete dynamic model is required. The Sun-line rotation, along with lunar and solar perturbations must be included in the dynamical equations and an asymptotic series solution obtained. By including these perturbations the resulting deviations from the nominal displaced circular halo orbit may be more fully understood. Using a more accurate nominal trajectory a more appropriate control scheme may be designed with smaller control accelerations.

The simple control schemes used with the heliocentric and geocentric halo orbits have been designed to demonstrate that the unstable families of these trajectories may be controlled. A more detailed and realistic control analysis is required which gives faster damping timescales and is designed for specified accuracies. Furthermore, it was found that extremely accurate sail pointing was required to generate the necessary control accelerations. Rather than using the entire sail for orbit control a separate system of vanes would allow accurate orbit control, with the main sail used only for generating the acceleration required to establish the halo orbit. Once

deployed on the halo orbit a linear quadratic regulator may be desirable for station-keeping purposes to minimise drift of the spacecraft from the nominal trajectory. However, to prevent continuous sail attitude trimming, a control with positional deadbands may be required. The size of the deadbands would be dependent on the station-keeping accuracy required which in turn is dependent on the particular mission being undertaken.

A dynamic problem related to the three-body stationary solutions is that of solar sail rendezvous. It was proposed that by stationing a solar sail spacecraft on a level surface of constant sail loading long planetocentric spiral times may be eliminated. However, for sample return missions the sail may be required to maneouvre to rendezvous with the payload returning from the planetary surface. Such maneouvres may be investigated by a linearisation of the dynamical equations about the stationary solution. Optimal control techniques may then be applied to obtain time optimal rendezvous trajectories.

8.3 Mission Analysis

More detailed studies of the trade-off between the sail mass per unit area, payload mass and sail area are required. Then it will be possible to define more accurately the range of possible payload mass that may be injected onto heliocentric and geocentric halo orbits with various orbital parameters. To define such a range of payload mass however, requires an estimate of currently achievable sail material and structural mass per unit areas. Similarly, the investigation of potential space science missions will be required to define the necessary geocentric and heliocentric halo orbit parameters. This then defines the required sail loading and so defines the required range of

payload mass. Therefore, for a given mission a final sail design may be iterated to between the bounds imposed by the scientific requirements and the technological constraints.

The three-body staging technique for sample returns requires a global optimisation study to minimise the total mission duration for a given payload. That is, the individual elements of the mission may be optimised but this is not necessarily equivalent to optimising the entire mission. The main parameter for such a mission is the mass of the returned sample. Therefore, the surface sample mass must be investigated as a function of the total mission duration and the sail design.

For a given mission the scientific payload will have a requirement for a certain accuracy for the instrument pointing and for orbit station-keeping. While accurate payload pointing may be accomplished by a separate payload attitude control mechanism the required station-keeping accuracy determines the required accuracy of the solar sail orbit control scheme. For certain types of instrumentation, such as magnetometers and particle flux detectors, there is little pointing accuracy required. However, for optical or x-ray detectors a high pointing accuracy may be required.

Other mission analysis requirements for a solar sail mission are similar to those for conventional space science missions. The thermal control of the payload is of importance for close heliosynchronous missions as is electrostatic charging for geomagnetic tail missions. Furthermore, for a chosen set of onboard experiments the total power budget must be calculated which then defines the area of the solar array. Similarly, the telemetry rate for data returns defines the size of the communications dish and other parameters of the link budget.

8.4 Recommendations

Before any advanced solar sail trajectory may be utilised for an actual mission a high performance solar sail with a dimensionless sail loading parameter of order unity must be designed. The current level of thin film technology and large space structure design must be assessed so that an estimate of sail material and structural mass per unit areas can be obtained. Only when an advanced, realistic sail design has been developed can the studies discussed in sections 8.2 and 8.3 be undertaken.

Another necessary requirement for the utilisation of these advanced trajectories is an actual flight demonstration of solar sailing. Such a demonstration is required to validate designs and to flight test sail materials and other hardware. Such a demonstration has been partially undertaken with the test of samples of sail material onboard the Soviet Almaz satellite, Shvartzburg (1991). If a small flight test of solar sailing is undertaken then the unique aspects of the advanced solar sail trajectories investigated in this thesis may be a primary driver for the adoption of solar sailing for an actual mission.

APPENDIX A. RADIATIVE TRANSFER

A.1 Specific Intensity

For a totally general radiation field the properties of the field are a function of both position and time. Furthermore, at a particular position within the radiation field the properties have a distribution in direction and frequency. The properties of a radiation field may be completely described in a macroscopic sense by the specific intensity of the field. The specific intensity $I_\nu(\mathbf{r},\mathbf{n};t)$ of the radiation at position $\mathbf{r}$ and time t propagating in direction $\mathbf{n}$ with frequency ν is defined to be the energy dE transported across a directed surface element $d\mathbf{A}$ in time dt into a solid angle $d\Omega$ about direction $\mathbf{n}$ in the frequency range $(\nu,\nu+d\nu)$, viz

$$dE = I_\nu(\mathbf{r},\mathbf{n};t)\ (\mathbf{n}.d\mathbf{A})d\Omega dt d\nu \qquad (A.1)$$

where $\mathbf{n}.d\mathbf{A}$ is the projected surface area normal to the direction of propagation $\mathbf{n}$.

From a microscopic perspective the radiation field is composed of individual photons which may be described by the photon number density function $\Psi_\nu(\mathbf{r},\mathbf{n};t)$. This function is defined such that $\Psi_\nu(\mathbf{r},\mathbf{n};t)d\Omega d\nu$ is the number of photons per unit volume at position $\mathbf{r}$ at time t and in a frequency range $(\nu,\nu+d\nu)$ propagating with speed c into a solid angle $d\Omega$ about direction $\mathbf{n}$. Therefore, the number of photons crossing a surface element $d\mathbf{A}$ in time dt is given by $\Psi_\nu(\mathbf{r},\mathbf{n};t)(\mathbf{n}.d\mathbf{A})d\Omega d\nu(cdt)$. Since each photon has energy $h\nu$, where h is the Planck constant, the energy transported is given by

$$dE = ch\nu\ \Psi_\nu(\mathbf{r},\mathbf{n};t)\ (\mathbf{n}.d\mathbf{A})d\Omega dt d\nu \qquad (A.2)$$

Comparing equations (A.1) and (A.2) it is clear that the specific intensity is related to the photon number density function by

$$I_\nu(\mathbf{r},\mathbf{n};t) = ch\nu\ \Psi_\nu(\mathbf{r},\mathbf{n};t) \qquad (A.3)$$

The photon number density function may then be related to the photon distribution function $\Psi_\nu(\mathbf{r},\mathbf{n};t)$. This function is defined such that the number of photons per unit volume in frequency range $(\nu,\nu+d\nu)$ and with momenta in the range $(\mathbf{p},\mathbf{p}+d\mathbf{p})$ is given by $\Psi_\nu(\mathbf{r},\mathbf{n};t)d^3p$, where $\mathbf{p}=(h\nu/c)\mathbf{n}$ is the photon momentum. Using the relation $d^3p=p^2dpd\Omega$ and equating photon numbers in a unit volume of space using the functions $\Psi_\nu(\mathbf{r},\mathbf{n};t)$ and $\Psi_\nu(\mathbf{r},\mathbf{n};t)$ it is found that

$$\left\{\frac{h^3\nu^2}{c^3}\right\}\ \Psi_\nu(\mathbf{r},\mathbf{n};t)\ d\nu d\Omega = \Psi_\nu(\mathbf{r},\mathbf{n};t)\ d\nu d\Omega \qquad (A.4)$$

Therefore, the specific intensity of the radiation field is given in terms of the photon distribution function as

$$I_\nu(\mathbf{r},\mathbf{n};t) = \left\{\frac{h^4\nu^3}{c^2}\right\}\ \Psi_\nu(\mathbf{r},\mathbf{n};t) \qquad (A.5)$$

For a black body radiation field, as the Sun may be assumed to be, the photon distribution function is a time independent, homogeneous and isotropic scalar function

$$\Psi_\nu = \frac{2}{h^3}\ (e^{(h\nu/kt)} - 1) \qquad (A.6)$$

such that I_ν is simply the Planck function. Therefore, by knowing the photon distribution function of the radiation field Ψ_ν, and so the specific intensity, the properties of the entire black body radiation field are known. These derived quantities, such as the flux and the

radiation pressure tensor, may be calculated by taking angular moments of the specific intensity.

A.2 Conservation of Specific Intensity

A useful property of the specific intensity function is that, for a time independent radiation field propagating in a vacuum, it is a conserved quantity along the direction of propagation.

Consider now radiation propagating in direction $\mathbf{n}$ at positions $\mathbf{r}$ and $\mathbf{r}+l\mathbf{n}$, and so at times t and t+(l/c). At position $\mathbf{r}$ the number of photons crossing surface element dA_1 in time dt into solid angle $d\Omega_1$ in the frequency range $(\nu,\nu+d\nu)$ is given by $c\Psi_\nu(\mathbf{r},\mathbf{n};t)dA_1 d\Omega_1 dt d\nu$. Similarly, the number of these photons dN_1 which also cross surface element dA_2 at $\mathbf{r}+l\mathbf{n}$ must be contained within a solid angle dA_2/l^2, viz

$$dN_1 = c\Psi_\nu(\mathbf{r},\mathbf{n};t)\ dA_1(dA_2/l^2)dtd\nu \qquad (A.7)$$

At position $\mathbf{r}+l\mathbf{n}$ the number of photons dN_2 crossing surface element dA_2 which have also crossed surface element dA_1 must be contained within solid angle dA_1/l^2, viz

$$dN_2 = c\Psi_\nu(\mathbf{r}+l\mathbf{n},\mathbf{n};t+(l/c))\ dA_2(dA_1/l^2)dtd\nu \qquad (A.8)$$

Since no photons have been created or destroyed between $\mathbf{r}$ and $\mathbf{r}+l\mathbf{n}$ the number of photons must be conserved so that $dN_1=dN_2$. Multiplying each side of equation (A.8) by $h\nu$ it is therefore found that

$$I_\nu(\mathbf{r},\mathbf{n};t) = I_\nu(\mathbf{r}+l\mathbf{n},\mathbf{n};t+(l/c)) \qquad (A.9)$$

so that the specific intensity is conserved along the direction of

propagation of the radiation. As a consequence of this conservation the value of $I_\nu(\mathbf{r},\mathbf{n};t)$, and so the radiation field characteristics, may be obtained from remote measurements of the energy falling in a given time on a known collecting area from a source subtending a known solid angle.

A.3 The Radiative Energy Flux and Pressure Tensor

The radiation flux $\mathbf{F}(\mathbf{r};t)$ is defined to be the vector such that $\mathbf{F}(\mathbf{r};t).d\mathbf{A}$ gives the nett rate of flow of radiative energy across a directed surface element $d\mathbf{A}$ in all frequencies. The nett number of photons crossing $d\mathbf{A}$ in unit time and in the frequency range $(\nu,\nu+d\nu)$ from all solid angles is given by

$$N = \left\{ \oint_{4\pi} c\Psi_\nu(\mathbf{r},\mathbf{n};t)\, \mathbf{n}\, d\Omega d\nu \right\}. d\mathbf{S} \qquad (A.10)$$

Therefore, multiplying by $h\nu$ the monochromatic radiative energy flux is obtained. Using equation (A.3), so that $I_\nu(\mathbf{r},\mathbf{n};t)=ch\nu\Psi_\nu(\mathbf{r},\mathbf{n};t)$, and integrating over the entire frequency spectrum equation (A.10) may be written as

$$\mathbf{F}(\mathbf{r};t) = \int_0^\infty \oint_{4\pi} I_\nu(\mathbf{r},\mathbf{n};t)\, \mathbf{n}\, d\Omega d\nu \qquad (A.11)$$

The radiation pressure tensor, or radiation stress tensor, $\mathbf{P}(\mathbf{r};t)$ is defined such that P^{ij} is the nett rate of transport, per unit area of surface oriented normal to j^{th} coordinate axis, of the i^{th} component of momentum. The number of photons with frequency ν propagting in direction n^j and crossing unit surface area in unit time is $c\Psi_\nu(\mathbf{r},\mathbf{n};t)n^j$ with each photon transporting momentum $(h\nu/c)n^i$ in the i^{th} direction.

Therefore, integrating over all solid angles the monochromatic radiation pressure tensor is obtained as

$$P_\nu{}^{ij}(\mathbf{r};t) = \oint_{4\pi} \Psi_\nu(\mathbf{r},\mathbf{n};t) \left\{\frac{h\nu n^i}{c}\right\} (cn^j)\, d\Omega \qquad (A.12)$$

However, from equation (A.3) $I_\nu(\mathbf{r},\mathbf{n};t)=ch\nu\Psi_\nu(\mathbf{r},\mathbf{n};t)$ so that by integrating over the entire frequency spectrum equation (A.12) may be written in dyadic form as

$$\mathbf{P}(\mathbf{r};t) = \frac{1}{c}\int_0^\infty \oint_{4\pi} I_\nu(\mathbf{r},\mathbf{n};t)\, \mathbf{nn}\, d\Omega d\nu \qquad (A.13)$$

Therefore, for a given specific intensity equation (A.13) may be used to obtain the radiation pressure at any point and in any direction in the radiation field. Furthermore, by knowing the surface geometry of any body placed within the radiation field the resulting radiation pressure force exerted on the body may be obtained.

APPENDIX B. REFERENCE FRAMES

B.1 Inertial Systems

Consider a system of n interacting particles in an inertial reference frame I. By Newton's second law the force exerted on the j^{th} particle of mass m_j at position $\mathbf{r}_j$ may be written as

$$\mathbf{f}_j = m_j \frac{d^2\mathbf{r}_j}{dt^2} , \quad \mathbf{r}_j \in \mathbb{R}^3 \quad (j=1,n) \tag{B.1}$$

The total mass M of the system may be written as

$$M = \sum_{j=1}^{n} m_j \tag{B.2}$$

Similarly, the centre of mass $\mathbf{R}$ of the system of particles may be defined in the inertial frame I as

$$\mathbf{R} = \frac{\sum_{j=1}^{n} m_j \mathbf{r}_j}{\sum_{j=1}^{n} m_j} \tag{B.3}$$

If equation (B.1) is now summed over all the n particles of the system it is found that

$$\sum_{j=1}^{n} m_j \frac{d^2\mathbf{r}_j}{dt^2} = \sum_{j=1}^{n} \mathbf{f}_j^{ex} + \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{f}_{ij}^{in} \tag{B.4}$$

where $\mathbf{f}_j^{ex}$ (j=1,n) are the external forces applied in the inertial frame I and $\mathbf{f}_{ij}^{in}$ (i,j=1,n) are the internal interaction forces between the i^{th} and j^{th} particles. However, by Newton's third law $\mathbf{f}_{ij}^{in}=-\mathbf{f}_{ji}^{in}$ (i,j=1,n) so that the last term of equation (B.4) vanishes. Therefore if the

system is isolated, so that there are no external forces acting on the system, equations (B.3) and (B.4) give

$$M \frac{d^2\mathbf{R}}{dt^2} = 0 \Rightarrow \mathbf{R} = \mathbf{R}_{o1} + \mathbf{R}_{o2}t \qquad (B.5)$$

where $\mathbf{R}_{o1}$ and $\mathbf{R}_{o2}$ are constants of the motion. Therefore, the centre of mass of the system moves with constant velocity with respect to the inertial frame I.

A Galilean transformation $\mathbf{R}'=\mathbf{R}-(\mathbf{R}_{o1}+\mathbf{R}_{o2}t)$ may then be used to transform to a new inertial frame I' so that the centre of mass of the system now lies at the origin. If $\mathbf{r}'_j$ is the position vector of the j^{th} particle with respect to the centre of mass of the system then $\mathbf{r}_j=\mathbf{R}'+\mathbf{r}'_j$ so that dynamical equations become

$$\frac{d^2\mathbf{r}'_j}{dt^2} = \frac{1}{m_j}\mathbf{f}_j \quad , \quad (j=1,n) \qquad (B.6)$$

Therefore, taking $\mathbf{R}'=0$ through the Galilean transformation the dynamical equations of the system may be formulated in an inertial frame with an origin at the centre of mass.

B.2 Rotating Systems

Although Newtonian dynamics applies in purely inerital reference frames, rotating systems require the use of non-inertial reference frames. In such frames the dynamical equations contain additional 'inertial forces' due the non-inertial nature of the system.

Consider a coordinate frame S which rotates with an angular velocity Ω about an arbitrary direction $\mathbf{n}$ with respect to the inertial frame I. For a particle P with position vector $\mathbf{r}$ fixed in S the infinitesimal change in $\mathbf{r}$ with respect to I during a rotation of $\delta\theta$ is

given by $\delta\theta(\mathbf{n}\times\mathbf{r})$. If the particle position vector $\mathbf{r}$ also changes in the rotating frame during the infinitesimal rotation, the total change in position measured in the inertial frame I is given by

$$\delta\mathbf{r}|_I = \delta\mathbf{r}|_S + \delta\theta\mathbf{n}\times\mathbf{r} \tag{B.7}$$

Therefore, given that $\Omega=(d\theta/dt)$, the rate of change of $\mathbf{r}$ in I is obtained from the limit of $\delta\mathbf{r}/\delta t$ as $\delta t\to 0$, viz

$$\frac{d\mathbf{r}}{dt} = \frac{\partial\mathbf{r}}{\partial t} + \boldsymbol{\Omega}\times\mathbf{r} \quad , \quad \boldsymbol{\Omega} = \left\{\frac{d\theta}{dt}\right\}\mathbf{n} \tag{B.8}$$

where the partial derivative is evaluated in the rotating frame S and the full derivative in the inertial frame I. If equation (B.8) is now written as an operator from the rotating frame to the inertial frame, acting on an arbitrary vector, viz

$$\frac{d}{dt}(*) = \frac{\partial}{\partial t}(*) + \boldsymbol{\Omega}\times(*) \tag{B.9}$$

Then, the inertial acceleration may be obtained from

$$\frac{d}{dt}\left\{\frac{d\mathbf{r}}{dt}\right\} = \frac{\partial}{\partial t}\left\{\frac{d\mathbf{r}}{dt}\right\} + \boldsymbol{\Omega}\times\left\{\frac{d\mathbf{r}}{dt}\right\} \tag{B.10}$$

Therefore, substituting equation (B.8) in equation (B.10) and collecting terms the acceleration, as observed from the inertial frame, is given by

$$\frac{d^2\mathbf{r}}{dt^2} = \frac{\partial^2\mathbf{r}}{\partial t^2} + 2\boldsymbol{\Omega}\times\frac{\partial\mathbf{r}}{\partial t} + \frac{\partial\boldsymbol{\Omega}}{\partial t}\times\mathbf{r} + \boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\mathbf{r}) \tag{B.11}$$

For uniform rotation $(\partial\boldsymbol{\Omega}/\partial t)=0$ so that the transverse acceleration term vanishes. The remaining terms are then the particle acceleration in

the rotating frame, the coriolis acceleration and the centrifugal acceleration respectively. Therefore, for a particle of mass m the dynamical equation in the rotating frame S is given by

$$\frac{\partial^2 \mathbf{r}}{\partial t^2} + 2\boldsymbol{\Omega} \times \frac{\partial \mathbf{r}}{\partial t} + \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}) = \frac{1}{m} \sum_{j=1}^{n} \mathbf{f}_j \qquad (B.12)$$

where the external forces $\mathbf{f}_j$ (j=1,n) exerted on particle P are applied from the inertial frame I.

APPENDIX C. DYNAMICS AND CONTROL OF LINEAR SYSTEMS

C.1 Solution of Linear Systems

Many natural and artificial systems may be described in terms of systems of linear ordinary differential equations. More often though systems of non-linear differential equations are required to fully define the dynamics. However, once a particular solution of the non-linear system is known the dynamics in the neighbourhood of this solution may be described by a linear system.

For a general non-linear n-dimensional dynamical system with a state vector $\mathbf{X}$ and input vector $\mathbf{U}$ the dynamical equation may be written as

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(\mathbf{X},\mathbf{U};t) \quad , \quad \mathbf{X} \in \mathbb{R}^n \tag{C.1}$$

If a particular solution $\mathbf{X}^*$ exists with a piecewise continuous input $\mathbf{U}^*$ equation (C.1) may be linearised about this solution. New variables in the neighbourhood of $\mathbf{X}^*$ may be defined as $\mathbf{X}=\mathbf{X}^*+\mathbf{x}$ and $\mathbf{U}=\mathbf{U}^*+\mathbf{u}$. Therefore, the nonlinear system may now be written as

$$\frac{d(\mathbf{X}^*+\mathbf{x})}{dt} = \mathbf{F}(\mathbf{X}^*,\mathbf{U}^*;t) + \sum_{j=1}^{\infty} \frac{1}{j!} \left.\frac{\partial^j \mathbf{F}}{\partial \mathbf{X}^j}\right|_{\mathbf{X}=\mathbf{X}^*,\ \mathbf{U}=\mathbf{U}^*} \mathbf{x}^j + \sum_{j=1}^{\infty} \frac{1}{j!} \left.\frac{\partial^j \mathbf{F}}{\partial \mathbf{U}^j}\right|_{\mathbf{X}=\mathbf{X}^*,\ \mathbf{U}=\mathbf{U}^*} \mathbf{u}^j \tag{C.2}$$

If $|\mathbf{x}|$ and $|\mathbf{u}| \ll 1$ then all the terms of order two or greater may be ignored in equation (C.2) so that a linear variational system is obtained, viz

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \quad , \quad \mathbf{A} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{X}}\right|_{\mathbf{X}=\mathbf{X}^*,\ \mathbf{U}=\mathbf{U}^*} \quad , \quad \mathbf{B} = \left.\frac{\partial \mathbf{F}}{\partial \mathbf{U}}\right|_{\mathbf{X}=\mathbf{X}^*,\ \mathbf{U}=\mathbf{U}^*} \tag{C.3}$$

with **A** the system matrix and **B** the input distribution matrix. It should be noted that if $\mathbf{X}^*$ is a time varying solution, rather than a stationary solution, then the system and input distribution matrices will have time varying coefficients. If however, $\mathbf{X}^*$ is time independent, or may be transformed to a time independent form using a suitable coordinate system, then equation (C.3) defines a set of first order linear, ordinary differential equations with constant coefficients.

The state transition matrix $\Phi(t_1,t_o)$ for equation (C.3) is defined by the matrix differential equation

$$\frac{d}{dt}\Phi(t_1,t_o) = \mathbf{A}\,\Phi(t_1,t_o) \quad , \quad \Phi(t_o,t_o) = \mathbf{I} \tag{C.4}$$

so that $\Phi(t_1,t_o)=e^{\mathbf{A}(t_1-t_o)}$. The solution of equation (C.3) with a null input is then simply

$$\mathbf{x}(t_1) = \Phi(t_1,t_o)\,\mathbf{x}(t_o) \tag{C.5}$$

so that any initial state vector may be propagated forwards or backwards in time once the state transition matrix is known. Furthermore, with a non-zero input it may easily be shown that the general solution of equation (C.3) is given by

$$\mathbf{x}(t_1) = \Phi(t_1,t_o)\,\mathbf{x}(t_o) + \int_{t_o}^{t_1} \Phi(t_1,\tau)\,\mathbf{B}\,\mathbf{u}(\tau)\,d\tau \tag{C.6}$$

where the integral term in equation (C.6) represents the cumulative effect of the input vector on the system dynamics. Therefore, for a known input $\mathbf{u}(t)$, $t_o<t<t_1$ and state transition matrix $\Phi(t_1,t_o)$ the solution of the linear system is known.

C.2 Stability Analysis

The stability characteristics of linear dynamical systems may be classified according to the system eigenvalues which are obtained from the characteristic polynomial. Substituting an exponential solution of the form $\mathbf{x}=\mathbf{x}_o e^{st}$ into equation (C.3) with a null input, a set of homogeneous linear equations are obtained, viz

$$(s\mathbf{I} - \mathbf{A})\mathbf{x}_o = \mathbf{0} \tag{C.7}$$

The non-trivial solution of equation (C.7) requires that the determinant of the matrix $(s\mathbf{I}-\mathbf{A})$ vanishes. The resulting characteristic polynomial $P(s)=0$ is then of the form

$$P(s) = \sum_{j=o}^{n} A_{n-j}\, s^j = A_o \prod_{j=o}^{n} (s-s_j)\ ,\ s_j = \sigma_j + i\omega_j \quad (j=1,n) \tag{C.8}$$

where the fundamental theorem of algebra allows $P(s)$ to be written as the product of the complex roots s_j $(j=1,n)$. The form of the solution of the linear system, and so the stability of the solution to the non-linear system, can now be determined from the sign of $\sigma_j=\mathrm{Re}(s_j)$, $(j=1,n)$.

Definition C.1 Unstable $(\sigma_j>0)$

A solution $\mathbf{X}^*$ of equation (C.1) is stable if

$$\exists \eta>0:\ |\mathbf{X}(t_o)-\mathbf{X}^*(t_o)| < \eta \Rightarrow \lim_{t\to\infty} |\mathbf{X}(t)-\mathbf{X}^*(t)| \to \infty \quad \forall \mathbf{X}(t) \tag{C.9}$$

Definition C.2 Asymptotic Stability $(\sigma_j<0)$

A solution $\mathbf{X}^*$ of equation (C.1) is asymptotically stable if

$$\exists \eta>0:\ |\mathbf{X}(t_0)-\mathbf{X}^*(t_0)| < \eta \Rightarrow \lim_{t\to\infty} |\mathbf{X}(t)-\mathbf{X}^*(t)| = 0 \quad \forall \mathbf{X}(t) \qquad (C.10)$$

Definition C.3 Lyapunov Stability ($\sigma_j=0$)

A solution $\mathbf{X}^*$ of equation (C.1) is Lyapunov stable if

$$\forall \epsilon>0\ \exists \delta>0:\ |\mathbf{X}(t_0)-\mathbf{X}^*(t_0)| < \delta \Rightarrow |\mathbf{X}(t)-\mathbf{X}^*(t)| < \epsilon \quad \forall \mathbf{X}(t) \qquad (C.11)$$

Along with these three definitions related to the sign of $Re(s_j)$ two other supplementary definitions are used, viz

Definition C.4 Poincaré Stability

Let S^* be the phase trajectory of $\mathbf{X}^*(t)$. Then, $\mathbf{X}^*(t)$ is a Poincaré stable solution of equation (C.1) if

$$\forall \epsilon>0\ \exists \delta>0:\ |\mathbf{X}(t_0)-\mathbf{X}^*(t_0)| < \delta \Rightarrow \sup \text{dist}(\mathbf{X},S^*) < \epsilon \qquad (C.12)$$

where the supremum of dist($\mathbf{X}$,S^*) is the least, usually perpendicular distance, from $\mathbf{X}$ to the nearest point on the phase trajectory S^*.

Definition C.5 Marginal Instability

A solution $\mathbf{X}^*$ of equation (C.1) is marginally stable if

$$\forall \epsilon>0\ \exists M>0:\ |\mathbf{X}(t_0)-\mathbf{X}^*(t_0)| < \epsilon \Rightarrow |\mathbf{X}(t)-\mathbf{X}^*(t)| < Mt \quad \forall \mathbf{X}(t) \qquad (C.13)$$

so that the instability is restricted to linear growth.

C.3 Feedback Control

For an unstable system a feedback control scheme may be developed to ensure asymptotic stability by relating the input of the system to the output in a feedback loop. However, it must firstly be

demonstrated that the system may be controlled in such a manner.

Without loss of generality it may be assumed that $\mathbf{x}(t_o)=0$ in equation (C.6). Then, writing the state transition matrix as a power series, the solution at time t_1 may be written as

$$\mathbf{x}(t_1) = \int_{t_o}^{t_1} \left\{ \sum_{j=o}^{\infty} \frac{\mathbf{A}^j}{j!} (t_1-\tau)^j \right\} \mathbf{B}\, \mathbf{u}(\tau)\, d\tau \qquad (C.14)$$

Since the system and input distribution matrices are time independent equation (C.14) may be written as

$$\mathbf{x}(t_1) = \sum_{j=o}^{\infty} \mathbf{w}_j\, (\mathbf{A}^j\mathbf{B}) \qquad (C.15)$$

where the set of constant vectors $\mathbf{w}_j$ are defined by

$$\mathbf{w}_j = \int_{t_o}^{t_1} \frac{(t_1-\tau)^j}{j!} \mathbf{u}(\tau)\, d\tau \qquad (C.16)$$

Therefore, the solution $\mathbf{x}(t_1)$ is a linear combination of the columns of the infinite matrix $(\mathbf{B},\mathbf{AB},.....)$. However, it may demonstrated that

$$r(\mathbf{B},\mathbf{AB},....,\mathbf{A}^{N-1}\mathbf{B}) = r(\mathbf{B},\mathbf{AB},....,\mathbf{A}^{n-1}\mathbf{B}) \quad , \quad \forall N>n \qquad (C.17)$$

so that if $r(\mathbf{B},\mathbf{AB},....,\mathbf{A}^{n-1}\mathbf{B})<n$ then a full basis for the n-dimensional state space does not exist. Therefore, owing to the construction of the solution in equation (C.15) it can be seen that there will be at least one final state $\mathbf{x}^*(t_1)$ that cannot be attained.

Definition C.6 Controllability

The linear system (C.3) of dimension n is completely controllable

if for any given initial state $\mathbf{x}_0$ and final state $\mathbf{x}_1$ there exists a finite time t_1 and a piecewise continuous input $\mathbf{u}(t)$, $t_0<t<t_1$ such that $\mathbf{x}(t_1)=\mathbf{x}_1$. The condition for complete controllability is then that the controllability matrix $\mathbf{C}=(\mathbf{B},\mathbf{AB},\ldots,\mathbf{A}^{n-1}\mathbf{B})$ has full rank, (a rigorous proof is given by Barnett and Cameron (1985)).

Given that a system is completely controllable, feedback control may be used to ensure asymptotic stability for an unstable system. For a general output vector $\mathbf{y}=\mathbf{Cx}$ the input vector may be written as $\mathbf{u}=\mathbf{K}(\mathbf{r}-\mathbf{y})$ where $\mathbf{K}$ is the gain matrix and $\mathbf{r}$ is a reference input which is to be tracked. Substituting in equation (C.3) it is found that

$$\frac{d\mathbf{x}}{dt} = (\mathbf{A} - \mathbf{BKC})\mathbf{x} + \mathbf{BK}\,\mathbf{r} \qquad (C.18)$$

If the required system dynamics can be specified in terms of a new system matrix $\mathbf{A}^*$ then the required gain matrix can be obtained as

$$\mathbf{K} = \mathbf{B}^{-1}(\mathbf{A} - \mathbf{A}^*)\mathbf{C}^{-1} \quad , \quad \det\mathbf{B} \neq 0 \quad , \quad \det\mathbf{C} \neq 0 \qquad (C.19)$$

where the condition $\det\mathbf{C}\neq 0$ requires that all of the state varibles are known. Therefore, by choosing the new system matrix $\mathbf{A}^*$ to have eigenvalues which lie in the left hand complex plane asymptotic stability can be ensured. The required gain matrix is then obtained from equation (C.19) as a function of the new and old system matrices.

REFERENCES

Allen, C.W.: 1955, Astrophysical Quantities, pp. 138-139, Athlone Press, University of London.

Angrilli, F. and Bortolami, S.: 1990, ESA Jour. 14, 431.

Bacon, R.H.: 1959, American Jour. Physics. 27, 164.

Barnett, S. and Cameron, R.G.: 1985, Introduction to Mathematical Control Theory, Clarendon Press, Oxford.

Borja, M.R.: 1984, M.S. Thesis, Wright-Patterson Air Force Base, Ohio.

Born, M.: 1927, Mechanics of the Atom, pp. 212-235, G.Bell and Sons Ltd., London.

Christensen-Dalsgaard, J. et. al.: 1984, in Space Science Horizon 2000, ESA SP-1070, pp. 27-36.

Cotter, T.P.: 1973, LA-5231, Los Alamos Scientific Laboratory.

Domingo, V. et. al.: 1985, SOHO-Solar and Heliospheric Observatory, ESA SCI(85) 7.

Drexler, K.E.: 1979, AIAA 79-1418, 4th Princeton/AIAA Conference on Space Manufacturing Facilities, Princeton, New Jersey.

Dumont, R. et. al.: 1980, in Solid Particles in the Solar System, IAU Symposium 90, pp. 67-70, eds. Halliday, I. and McIntosh, B.A.

Dunham, D.W.: 1989, Handbook on Trajectories, Mission Design and Operations for the Spacecraft of the Solar-Terrestrial Science Project, Computer Sciences Corporation, Maryland.

Epstein, P.S.: 1916, Ann. Der Physik. 50, 489.

Farquhar, R.W. et. al.: 1980, Jour. Guidance and Control. 3, 549.

Fimple, W.R.: 1962, American Rocket Soc. Jour. 32, 527.

Forward, R.L.: 1981, Jour. Astronautical Sciences. 29, 73.

Forward, R.L.: 1984, Jour. Astronautical Sciences. 32, 221.

Forward, R.L.: 1986, AAS 86-409, 33rd Annual American Astronautical Society Meeting, Boulder, Colorado.

Forward, R.L.: 1989a, AIAA 89-2545, AIAA/ASME/SAE/ASEE 25th Joint Propulsion Conference, Monterey, California.

Forward, R.L.: 1989b, AIAA 89-2546, AIAA/ASME/SAE/ASEE 25th Joint Propulsion Conference, Monterey, California.

Forward, R.L.: 1991, Jour. of Spacecraft and Rockets. 28, To be published.

Fox, H.G.: 1989, JHU/APL-SDO9252, Space Department, Applied Physics Laboratory, Johns Hopkins University, Baltimore, Maryland.

Friedman, L. et. al.: 1978, AIAA 78-7, AIAA 16th Aerospace Sciences Meeting, Huntsville, Alabama.

Friedman, L.: 1988, Starsailing-Solar Sails and Interstellar Travel, Wiley and Sons Inc., New York.

Garvey, J.: 1987, Foundation Astronautics Notebook 24/25, World Space Foundation, Pasadena, California.

Garwin, R.L.: 1958, Jet Propulsion. 28, 188.

Goldstein, H.: 1980, Classical Mechanics, Addison-Wesley, Massachusetts.

Gordon, B.J.: 1961, M.S. Thesis, Wright-Patterson Air Force Base, Ohio.

Green, J.: 1977, M.S. Thesis, Charles Stark Draper Laboratory Inc., Cambridge, Massachusetts.

Hill, J.R. and Whipple, E.C.: 1985, Jour. Spacecraft and Rockets. 22, 245.

Isayev, Y.N. and Kunitsyn, A.L.: 1972, Celestial Mechanics. 6, 44.

Jacobson, R.A. and Thornton, C.L.: 1978, Jour. Guidance and Control. 1, 365.

Johnson, G.E. et. al.: 1989, Q7844/JPA/Issue 1, Cambridge Consultants Ltd., Cambridge.

Jordan, D.W. and Smith, P.: 1987, Non-linear Ordinary Differential Equations, pp. 245-261, Calrendon Press, Oxford.

Kane, S.R.: 1982, Astrophysical Jour. 254, L53.

Kelly, H.J.: 1960, American Rocket Soc. Jour. 30, 947.

Kiefer J.W.: 1965, Proc. 15th International Astronautical Congress, 1, 383.

Lebedew, P.: 1902, Astrophysical Jour. 10, 155.

London, H.S.: 1960, American Rocket Soc. Jour. 30, 198.

Longsdon, J.M.: 1989, ISIS Jour. 80, 254.

MacNeal, R.H.: 1967, ARC-R-249, Astro Research Corporation, Santa Barbara, California.

MacNeal, R.H.: 1971, NASA contractors report CR-1745.

Malanin, V.V. and Repyakh, A.V.: 1974, Problemy Mekhaniki Upravlyayemogo Dvizheniya. 5, 99.

Mercader del Rio, L.: 1989, ESA Jour. 13, 27.

Mihalas, D. and Mihalas, B.W.: 1984, Foundations of Radiation Hydrodynamics, Oxford University Press, Oxford.

Nayfeh, A.: 1973, Perturbation Methods, chapter 6, Wiley and Sons, New York.

Nichols, E.E. and Hull, G.F.: Physical Review. 17, 26.

Oswald, G.: 1990, Private Communication.

Perret, A. et. al.: 1989, IAF 89-539, 40th International Astronautical Congress, Malaga, Spain.

Riviere, G.: 1977, Proc. 28th International Astronautical Congress, 1, 181.

Roy, A.E.: 1982, Orbital Motion, pp. 125-138, Hilger Co., Bristol.

Sackett, L.L.: 1977, NASA contractors report CR-157129.

Sackett, L.L. and Edelbaum, T.M.: 1977, AAS 78-31901, AAS/AIAA Astrodynamics Specialist Conference, Jackson Hole, Wyoming.

Sands, N.: 1961, American Rocket Soc. Jour. 31, 527.

Sauer, C.G.: 1976, AIAA 76-792, AIAA/AAS Astrodynamcis Conference, San Diego, California.

Sauer, C.G.: 1977, AAS 78-31903, AAS/AIAA Astrodynamics Specialist Conference, Jackson Hole, Wyoming.

Schuerman, D.W.: 1980, Astrophysical Jour. 238, 337.

Schwartz, S.J.: 1985, in Solar System Magnetic Fields, pp. 190-223, ed. Priest, E.R., Reidel Co., Dordrecht.

Shvartsburg, A.: 1991, Soviet Solar Sailing Group Press Release, CDBUI, Moscow.

Simmons, J.F.L. et. al.: 1985, Celestial Mechanics. 35, 145.

Staehle, R.L.: 1982, Jour. British Interplanetary Soc. 35, 327

Teeter, R.R.: 1977, NASA contractors report CR-158167.

Timnat, Y.M.: 1987, Advanced Chemical Rocket Propulsion, pp. 208-213, Academic Press, London.

Tsander, K.: 1924, From a Scientific Heritage, NASA TTF-541 (1967), quoting a 1924 report by the author.

Tsiolkovsky, K.E.: 1921, Extension of Man Into Outer Space, (also 1936, Symposium Jet Propulsion, United Scientific and Technical Press).

Tsu, T.C.: 1959, American Rocket Soc. Jour. 29, 442.

Van der Ha, J.C. and Modi, V.J.: 1977a, Jour. Astronautical Sciences. 25, 283.

Van der Ha, J.C. and Modi, V.J.: 1977b, Acta Astronautica. 4, 813.

Van der Ha, J.C. and Modi, V.J.: 1977c, AIAA 77-32, AIAA 15th Aerospace Sciences Meeting, Los Angeles, California.

Van der Ha, J.C. and Modi, V.J.: 1979, Celestial Mechanics, 19, 113.

Van der Ha, J.C.: 1980, DGLR 80-012, Walter Hohmann Symposium, Cologne.

Villers P.: 1960, M.S. Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Vonbun, F.O.: 1968, NASA technical note TN-D-4468.

Von Flotow, A.H.: 1989, A Solid State Heliogyro, Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Wenzel, K.P. et. al.: 1990, ESA Bulletin. 63, 21.

Wilson, R. et. al.: 1981, Science. 211, 700.

Wright, J.L.: 1974, BMI-NLVP-TM-74-3, Battelle Memorial Institute, Columbus, Ohio.

Wright, J.L. and Warmke, J.M.: 1976, AIAA 76-808, AIAA/AAS Astrodynamics Conference, San Diego, California.

Wright, J.L. and French, J.R.: 1987, Foundation Astronautics Notebook 23, World Space Foundation, Pasadena, California.

Wright, J.L: 1990, Space Sailing, Gordon and Breech, London, (to be published).

Zhukov, A.N. and Lebedev, V.N.: 1964, Cosmic Research. 2, 41.