Realistic Earth Escape Strategies for Solar Sailing

Malcolm Macdonald* and Colin R. McInnes†
University of Glasgow, Glasgow, Scotland G12 8QQ, United Kingdom

With growing interest in solar sailing comes the requirement to provide a basis for future detailed planetary escape mission analysis by drawing together prior work, clarifying and explaining previously anomalies. Previously unexplained seasonal variations in sail escape times from Earth orbit are explained analytically and corroborated within a numerical trajectory model. Blended-sail control algorithms, explicitly independent of time, which provide near-optimal escape trajectories and maintain a safe minimum altitude and which are suitable as a potential autonomous onboard controller, are then presented. These algorithms are investigated from a range of initial conditions and are shown to maintain the optimality previously demonstrated by the use of a single-energy gain control law but without the risk of planetary collision. Finally, it is shown that the minimum sail characteristic acceleration required for escape from a polar orbit without traversing the Earth shadow cone increases exponentially as initial altitude is decreased.

Nomenclature

a	=	semimajor axis
$\boldsymbol{a}_s$	=	sail acceleration
a_{s_c}	=	sail characteristic acceleration
E	=	eccentric anomaly
e	=	eccentricity
f	=	modified equinoctial element
g	=	modified equinoctial element
L	=	modified equinoctial element
N	=	out-of-plane sail acceleration force
$\boldsymbol{n}_s$	=	sail normal vector
p	=	semilatus rectum
R	=	radial sail acceleration force
r_p	=	radius of pericentre
T	=	transverse sail acceleration force
t	=	time
$\boldsymbol{v}$	=	$[v_x \ v_y \ v_z]$, sail velocity vector
W	=	weight of control law
α	=	sail pitch angle
$\tilde{\alpha}$	=	cone angle of force vector, from sun-sail line, along which locally optimal control maximises sail force vector
β	=	angle between orbit plane and plane defined by sun and velocity vectors
δ	=	sail clock angle
ε	=	ecliptic plane inclination/obliquity of ecliptic
$\boldsymbol{\lambda}$	=	locally optimal unit vector for variation of orbit element ($= [\lambda_x \ \lambda_y \ \lambda_z]$)
μ	=	Earth's gravitational parameter ($3.986032 \times 10^{+14}$ m^3 s^{-2})
ν	=	true anomaly

Introduction

SOLAR sailing is increasingly being considered by both ESA and NASA for future science missions. With the absence of reaction mass arises the potential for new high-energy mission concepts, such as sample and return missions to the terrestrial planets or near-Earth objects.[1,2] A range of applications for solar sailing have been extensively discussed since the 1960s; however, planet-centered tra-jectory analysis has been largely restricted to escape maneuvers or relatively simplistic orbit maneuvering, such as Lunar flyby or orbit inclination change.[3–5] In the early 1960s Sands and Fimple investigated solar-sail planetary escape trajectories using initially circular orbits and analytical techniques, which necessitated many simplifying approximations, including a fixed solar position and omitting orbit restrictions such as negative altitude.[6,7] Fimple used a locally optimal energy-gain control strategy, based on the earlier generalized conclusion by Irving that such a strategy is sufficiently close to the optimal for most low-thrust propulsion systems.[8] Lawden mathematically showed for a low-thrust motor that little advantage was to be gained by implementation of a more complex thrust program than aligning the motor thrust with the velocity vector.[9] In the late 1970s this generalized statement about low-thrust propulsion was shown to hold for solar sailing, with an overall efficiency drop of between 1 and 3.5% for escape from high Earth orbits.[10] In 1978 Sackett and Edelbaum presented optimal Earth subescape and orbit transfer maneuvers for solar-sail propulsion, using a method of orbit averaging to reduce the number of computations needed while still giving good performance estimates through the solution of a two-point boundary-value problem.[11,12] During this work, a characteristic rapid increase in orbit eccentricity was noted in most trajectories, and it was found that often the optimal solution resulted in a negative altitude perigee passage. A minimum altitude constraint was not included in the optimization tool; however, a penalty function was developed to ensure an adequate radius of perigee was maintained throughout the trajectory without significant loss of optimality. Several papers have since been published discussing the use of solar-sail propulsion for Earth escape using locally optimal techniques, and although the inclusion of a rotating sunline has become commonplace, much of the analysis performed continues to make significant simplifications, such as implementing a spherical gravity model or neglecting shadow, third-body gravity effects or even minimum perigee altitudes.[13,14] Recently, however, a more complete Earth escape trajectory analysis has been performed by Leipold, where the preceding perturbations were included, while continuing to neglect a minimum altitude constraint.[15] We also note that recent work has produced extremal steering strategies for simulation and optimization of Earth–moon transfer trajectories using solar sailing. This work resulted in the solution of a weak stability boundary problem and generated realistic orbital mechanics solutions for the transfer, with the inclusion of all relevant perturbations.[16]

With the increased interest in solar sailing from the science community, it becomes necessary to clarify previous anomalies observed in mathematical models and eliminate the simplifications made earlier in order to aid future analysis of solar-sailing missions, while also stressing the engineering complexity of planetary escape by sail propulsion. Slight variations, up to 7%, exist for escape times depending on the launch date through the Earth year. However,

*Research Assistant, Department of Aerospace Engineering; m.macdonald@aero.gla.ac.uk.
†Professor, Department of Aerospace Engineering.

no adequate explanation has been offered for the presence of this variation.[13-15] The number of eclipse events or the rate of energy gain by the sail have both been suggested as possible explanations. It is noted in Ref. 13 that the variation appears to fall as sail acceleration increases. However, it is hypothesized in Ref. 15 that an increase in sail acceleration can extenuate the variation, presumably as a greater proportion of the much shorter trajectory will now be in shadow. Furthermore, for geocentric spiral trajectories to both the lunar distance and a subescape point, the time of flight tends to be minimum for orbits within the ecliptic plane. It will be shown in this paper that these two factors are related by the derivation of an optimal inclination for solar-sail maneuvering and that the effect is amplified at low sail accelerations.[17]

The paper will discuss a simple, autonomous solution to the problem of planetary escape, while maintaining a minimum pericenter altitude, such as to remain above the upper atmosphere or to expedite the sail orbit out of this region and then maintain so thereafter. Following the generation of algorithms to provide safe planetary escape trajectories; Earth escape trajectories without Earth occultation of the sail are examined. This is a potentially beneficial scenario for attitude control, thermal, and other subsystem design.

Optimal Inclination for Planet-Centered Solar Sailing

Using a derivation of the locally optimal energy-gain steering law, the rate of energy variation is shown to be related to both the sail acceleration and the orbit inclination, thus confirming the presence of a theoretically optimal inclination. Time until Earth escape is investigated over a range of sail accelerations and inclinations using an independently derived solar-sail trajectory model, hence corroborating the effect of orbit inclination and consequently time of year on escape time. Additionally, the effect of introducing Earth eclipse is investigated in order to understand and quantify the effect this might have on escape times throughout the year.

To derive the locally optimal energy gain control law, following Ref. 15, the definition of a sun–vector coordinate system is required; the system is illustrated in Fig. 1. The origin of the sun–vector coordinate system is defined as the spacecraft center of mass and aligns the positive X_{sun} axis with the instantaneous direction of the sun. The Z_{sun} axis is defined as the cross product of the velocity vector and the sun unit vector, with the Y_{sun} axis completing the right-hand Cartesian coordinate system. Thus the Z_{sun} axis velocity component is always zero in the sun–vector coordinate system. An advantage of this system is that the sail normal vector and the direction of the sun directly define the sail pitch angle. The sail clock angle is taken from the Y_{sun} axis toward the projection of the sail normal into the plane defined by the Y_{sun} and Z_{sun} axis, as also seen in Fig. 1.

The sail normal vector is then described by Eq. (1) as

$$\mathbf{n}_s = \begin{pmatrix} \cos\alpha \\ \sin\alpha\cos\delta \\ \sin\alpha\sin\delta \end{pmatrix} \quad (1)$$
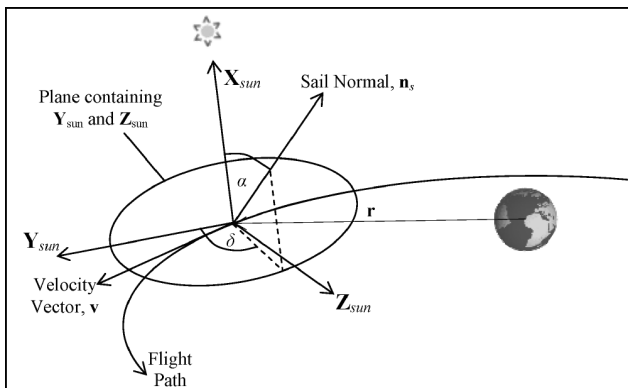
Fig. 1   Sun–vector coordinate system, with the plane normal to X_{sun} illustrated.

Note that when the velocity vector and the sun unit vector are parallel, this coordinate system is not defined; thus, the system is used only for the development of theory and not for orbit propagation.

Locally Optimal Energy Gain Control

This sail control strategy has been widely published in many forms and is based on the projection of the sail acceleration vector onto the velocity vector, hence maximizing the energy rate of change of the trajectory at any given point.[1,7,10,13-15,17]

Using the sun–vector coordinate system and following Ref. 15, we define the function to be maximized in Eq. (2):

$$F(\alpha, \delta) = \mathbf{a}_s \cdot \mathbf{v} \quad (2)$$

The sail acceleration vector is defined as $\mathbf{a}_s = a_{s_c}(\mathbf{n}_s \cdot \mathbf{X}_{sun})^2 \times \mathbf{n}_s$; thus combining with Eq. (1); we obtain the sail acceleration vector;

$$\mathbf{a}_s = a_{s_c}\cos^2\alpha \begin{pmatrix} \cos\alpha \\ \sin\alpha\cos\delta \\ \sin\alpha\sin\delta \end{pmatrix} \quad (3)$$

From the definition of the coordinate system, recall that the Z_{sun} axis velocity component is zero; thus,

$$F(\alpha, \delta) = a_{s_c}\left(v_x\cos^3\alpha + v_y\cos^2\alpha\sin\alpha\cos\delta\right) \quad (4)$$

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

$$\frac{\partial F(\alpha, \delta)}{\partial\delta} = -a_{s_c}v_y\cos^2\alpha\cdot\sin\alpha\cdot\sin\delta = 0 \quad (5)$$

$$\frac{\partial F(\alpha, \delta)}{\delta\alpha} = -a_{s_c}\big[3v_x\cos^2\alpha\sin\alpha$$
$$+ v_y\cos\delta(2\cos\alpha\sin^2\alpha - \cos^3\alpha)\big] = 0 \quad (6)$$

Rearranging Eq. (6), with $\cos^3\alpha \neq 0$ or $\alpha \neq 0$ and $\cos\delta \neq 0$ or $\delta \neq 90$ deg, 270 deg, and $v_y \neq 0$ gives

$$a_{s_c}\left(\tan^2\alpha + \frac{3v_x}{2v_y\cos\delta}\tan\alpha - \frac{1}{2}\right) = 0 \quad (7)$$

Solving for α and δ, from Eqs. (5) and (7),

$$\alpha_{1,2} = \arctan\left[-\frac{3v_x}{4v_y\cos\delta} \pm \sqrt{\frac{1}{2} + \left(\frac{3v_x}{4v_y\cos\delta}\right)^2}\right] \quad (8)$$

$$\delta_{1,2} = \arcsin(0) \Rightarrow \delta_1 = 0, \delta_2 = 180\,\text{deg} \quad (9)$$

As obtained in Ref. 15, Eqs. (8) and (9) both have two solutions. Equation (8) allows the optimal sail pitch angle to be found for the special case where $v_x = 0$. This angle is often quoted in literature as the optimal fixed-sail pitch angle:

$$\alpha_{\text{opt}} = \left|\arctan\left(\pm\frac{1}{\sqrt{2}}\right)\right| \cong 35.264\,\text{deg} \quad (10)$$

The solution for the sail clock angle, Eq. (9), states that an optimal steering law is achieved if the sail normal vector, the velocity vector, and the sun vector are all within the same plane. The sun vector and the velocity vector orientations cannot be altered or optimized; therefore, the optimal condition defined by Eq. (9) can be achieved only by aligning the sail normal vector within the plane defined by the other vectors, thus requiring a fixed-sail clock angle of $\delta = 0$ or 180 deg.

The preceding optimal energy control law can also be derived directly from the variational equation of the semimajor axis and

can thus be called the semimajor axis control law.[17] The variational equation of the semimajor axis[18] is given in Eq. (11):

$$\frac{da}{dt} = \frac{2a^2}{\sqrt{\mu p}} [R \quad T \quad N] \begin{bmatrix} e \sin v \\ (1 + e \cos v) \\ 0 \end{bmatrix} \quad (11)$$

We see from Eq. (11) that the rate of change of semimajor axis depends only on the radial and transverse perturbing accelerations and not on the out-of-plane acceleration. It therefore follows that in order to maximize the rate of change of semimajor axis, and hence orbit energy, the sail force should ideally be oriented entirely within the orbit plane. However, the orbit plane and the plane defined by the velocity and sun vectors are coincident only if the sail orbit lies within the ecliptic plane. When the sail orbit is outside the ecliptic plane, an angle exists between the orbit plane and the velocity/sun vector plane, denoted by the angle β. It is thus not always possible to maximize the sail force within both required planes at all times. To ensure the local maximum rate of energy change, the sail force vector is optimized such that the maximum sail force is directed along the orbit velocity vector. The sail pitch angle is found using a standard optimization process, given in Eq. (12) (Ref. 10) as

$$\tan \alpha = \frac{-3 \cos \alpha + \sqrt{9 \cos^2 \tilde{a} + 8 \sin^2 \tilde{a}}}{4 \sin \tilde{a}} \quad (12)$$

The derivation of $\tilde{\alpha}$ and the application of Eq. (12) will be discussed later within this paper. If the orbit plane is not coincident with the ecliptic plane, this optimization process rotates the sail normal vector out of the orbit plane toward the plane defined by the velocity and sun vectors, hence generating an out-of-orbit-plane force. As noted, the rate of change of semimajor axis depends only on the orbit perturbations within the plane of motion, and hence the generation of an out-of-plane sail force reduces optimality. Thus, we define the optimal orbit inclination such that the plane of motion is coincident with the ecliptic plane.

The definition of the ecliptic plane as the optimal plane for solar-sail in-plane orbit maneuvering allows us to explain the apparent seasonal variation of sail escape times from Earth orbit. A 7-deg orbit inclination at the northern hemisphere winter solstice results in an orbit inclination of 16.4 deg from the ecliptic plane. However, at the northern hemisphere summer solstice the inclination to the ecliptic is now 30.4 deg. Thus, the increased inclination with respect to the ecliptic plane should result in a greater out-of-orbit-plane force and hence an increased escape time for a June/July launch, as found in Refs. 13 and 14. As the sail acceleration is increased, the difference between escape time for June and December launch should decrease, as the number of orbits until escape is reduced, hence minimizing the effect of the out-of-plane sail force.

The defined optimal inclination holds true for the locally optimal variation of any orbit element or parameter where the rate of change is dependent on only the in-plane perturbing forces; these include eccentricity, radius of pericenter, and radius of apocenter.

Earth Escape Time, Sail Acceleration, and Orbit Inclination

The time until Earth escape was determined using modified equinoctial elements in the equations of motion, which have been studied, validated, and utilized by many previous authors for similar applications.[13,14,17,19–21] The equations of motion are propagated using an explicit, variable step size Runge–Kutta formula, the Dormand–Price pair, with relative and absolute error tolerances of 10^{-9} ensuring minimal truncation error.[22] The initial orbit inclination is defined with respect to the equatorial plane, such that orbits within the ecliptic plane have $i = \varepsilon \cong 23.439$ deg. The inclination is varied from 0 to 179 deg because the equations of motion are singular for inclinations of 180 deg only and remain defined through the point of escape, allowing the generation of the time until escape rather than a defined subescape point. The singularity at inclination 180 deg can be handled by appropriate redefinition.

The orbit model initially assumes a perfect reflector with no perturbations other than the sail itself, whereas the sun is assumed a parallel point source. Initially Earth shadow is also neglected allowing the nature of the relationship between inclination and sail efficiency to be seen without the background effects generated by orbit perturbations. Earth shadow is then introduced to provide a comparison, as sail propulsion efficiency is reduced for orbits within the ecliptic plane, because of the large fraction of time spent in Earth's shadow.[15] Such a reduction in sail efficiency could be expected to influence the escape times, hence altering the optimal inclination. However this is shown not to be true.

Figure 2 shows the time until escape from geostationary orbit (GEO) radius for a range of sail characteristic accelerations, from 0.15 to 2.0 mm s^{-2}. We see that for sail accelerations of 0.75 mm s^{-2} and greater that the minimum escape time corresponds to an orbit inclination within the ecliptic plane, as predicted. The minimum is visible on the surface plot as a groove on the otherwise reasonably smooth surface. However, for sail accelerations below 0.75 mm s^{-2} the minimum is not evident, though the orbits near the optimal inclination do tend to be the quickest to escape. This breakdown is caused by the relatively low level of sail acceleration compared to local gravity, and as a result the optimal inclination effect is lost during the high number of orbit revolutions required to gain escape energies. The breakdown in the predicted relationship between sail performance and orbit inclination is reflected by the much more irregular nature of the surface plot at low sail accelerations. We also note from Fig. 2 that the irregular surface continues into higher sail accelerations for inclinations between 45 and 90 deg; the reason for this remains unclear, and no satisfactory explanation could be derived, however, calculation error was eliminated as a possible cause.

As the orbit inclination increases from zero to ε the angle β, between the orbit plane and the plane defined by the sun/velocity vectors, decreases to zero, hence the optimal inclination. As inclination then continues to increase to 2ε, β increases symmetrically with $i < \varepsilon$. This symmetry is reflected in the time until escape, seen by taking a section through Fig. 2 at characteristic acceleration of 0.75 mm s^{-2} in Fig. 3. As the inclination increases, β continues to increase toward a maximum of 180-deg ε at inclination 180 deg. We see in Fig. 3 the time until escape rises as the orbit inclination tends toward 180 deg.

Furthermore, we note a change in orbit inclination can be as influential on escape time as a modest increase in sail acceleration of up to 0.25 mm s^{-2}; this is an important consideration that should be taken into account in the early stages of any solar-sail planet-centered mission analysis.

Figures 2 and 3 show the exact relationship between sail performance and orbit inclination and clearly shows an optimal sail inclination of $i = \varepsilon$. However, when passing through the Earth's shadow cone no propulsion is provided, and sail propulsion efficiency has been shown to be lower for orbits within the ecliptic plane.[15] Therefore, while the basic orbital mechanics suggests an
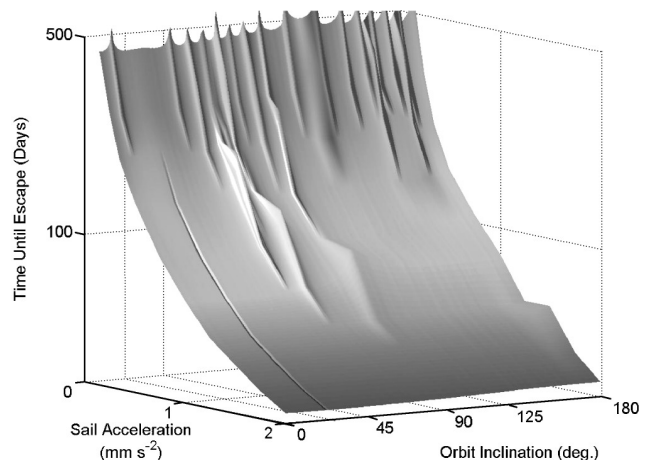
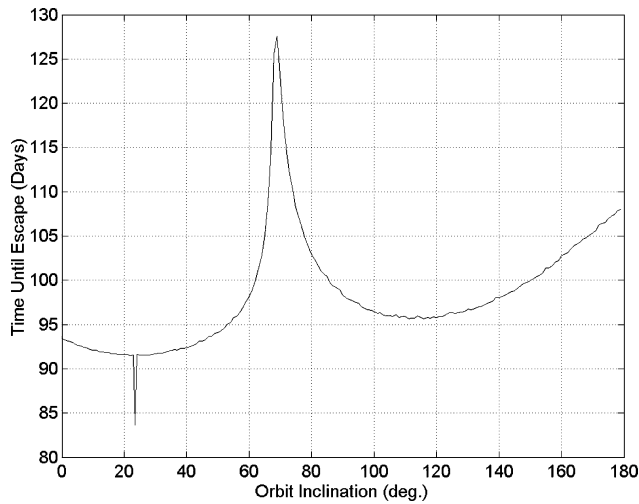**Fig. 2 Solar-sail escape time from GEO radius, without shadow effects.**

Fig. 3 Section through Fig. 2 corresponding to escape time for sail characteristic acceleration 0.75 mm s^{-2} vs inclination, without shadow effects.
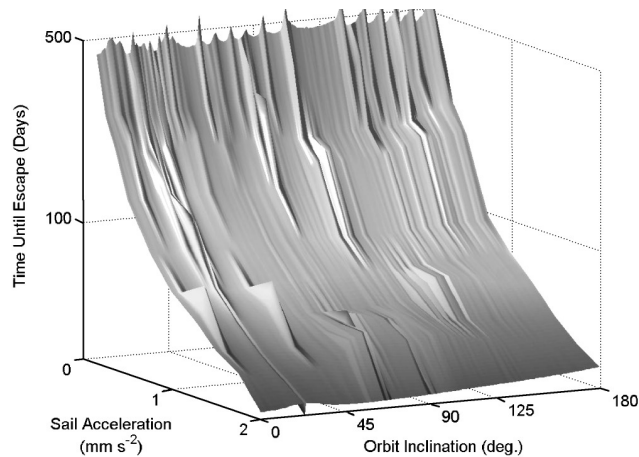

Fig. 4 Solar-sail escape time from GEO radius, with shadow.

inclination within the ecliptic plane to be optimal, the introduction of Earth shadow could be expected to alter this.

We see in Fig. 4 the time until escape from GEO radius for a range of sail characteristic accelerations, from 0.15 to 2.0 mm s^{-2}, with Earth shadow effects included. The Earth's shadow cone is modeled to include both umbra and penumbra occlusions. Figure 4 shows that shadow does not alter the optimal inclination, with the surface channel still visible at $i = \varepsilon$, yet we also notice that the time until escape is increased for orbits near to the ecliptic plane, because of the presence of shadow in this region. Figure 4 also shows the surface to be much more uneven than before, with the surface remaining irregular up to much higher sail accelerations. The irregular surface structure has been noted in the shadow-free case to be an indicator that the relationship between orbit inclination and sail performance is starting to breakdown. We note once again however that a change in orbit inclination can still be as influential on escape time as an increase in sail acceleration. Furthermore, the symmetrical nature of escape times about the ecliptic plane is now much more visible than in Fig. 2, and the increase in escape times for increasing orbit inclination can once again be seen.

As sail acceleration is increased, we see in Fig. 5 the difference between escape time at optimal inclination, and worst-case inclination is confirmed to fall. This is analogous to the seasonal variations in Earth escape times found previously and confirms that this variation will reduce as sail acceleration is increased.
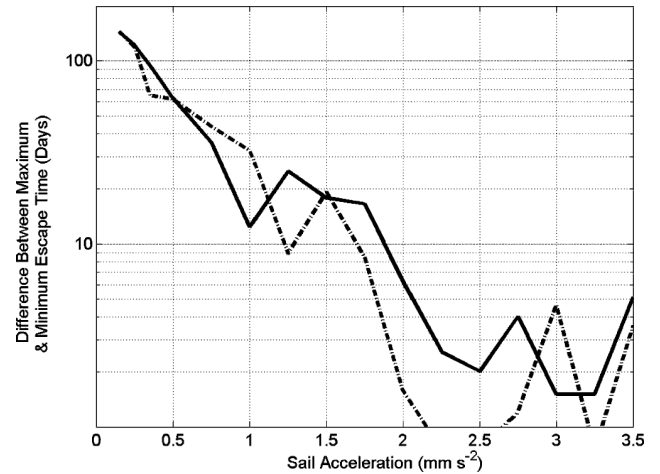

Fig. 5 Difference between maximum and minimum escape time: ——, shadow case and – – –, without shadow.

Near-Optimal Earth Escape with Minimum Altitude Constraint

Because of the high number of orbit revolutions that typify low-thrust planet-centered trajectories, accurate mission analysis is hindered as a result of the computational difficulties of generating optimal transfers. There is thus a requirement to generate planet-centered trajectories that are computationally simple and provide near-optimal orbits, hence the popularity of locally optimal solutions. The solution of a two-point boundary-value problem has been show to provide optimal orbits; however, it requires many simplifications and assumptions to be made, while still requiring a large amount of computational time in order to generate sufficient accuracy and as such is not practical for direct implementation in a real-time scenario. Locally optimal sail control has the advantage that the sail control angles can be calculated independent of time, hence making the system suitable for onboard autonomous sail control, as the control system could potentially be made into a closed loop with onboard navigation.

Using only the semimajor axis control law, we find that as orbit energy tends toward positive values the eccentricity tends to rapidly increase during the final few orbits prior to escape, resulting in a corresponding rapid decrease in perigee altitude. However, because of the nature of solar-sail propulsion about a planet, it is possible to gain energy for only half an orbit, as the sail travels away from the sun. Thus, if the sail is slightly below the energy required to escape a planet's gravity field at the end of this half-orbit the sail requires another pass of the planet before escape. If the radius of perigee is less than the radius of the planet, this will result in a collision with the planet, as has been seen in previous work where a negative altitude was noted prior to escape.[17] It thus follows that we require altering the locally optimal strategy used such that negative altitudes no longer become possible; in effect we wish to set a minimum radius of perigee.

Locally Optimal Steering Laws

The rate of change of any orbital element can be calculated and hence a locally optimal control law generated. These control laws maximize the instantaneous rate of change of the element and provide the sail orientation in closed analytical form. Following Ref. 17, a radius of pericenter control law can thus be generated in a similar manner to the semimajor axis control law already discussed, through use of Eq. (12) and the rate of change of pericenter given in Eq. (13) as

$$\frac{dr_p}{dt} = \frac{da}{dt}(1 - e) - a\frac{de}{dt} \qquad (13)$$

The relative perturbing force on each of the radial, transverse, and normal (RTN) axes can thus be found in the form of a unit vector along which we require to maximize the sail force in order to

maximize the local rate of change. The unit vector is a function of the orbit elements and explicitly independent of time. The vector for locally optimal increase of semimajor axis and radius of pericenter are defined in Eqs. (14) and (15), respectively, in modified equinoctial elements. The locally optimal decrease is found by taking the negative of Eqs. (14) and (15). We see from Eq. (14) that the vector for locally optimal variation of semimajor axis comes directly from Eq. (11). Equation (15) is derived by writing Eq. (13) in a similar form to Eq. (11) and then converting into modified equinoctial elements.

$$\lambda_a^{RTN} = \begin{bmatrix} e \sin v \\ (1 + e \cos v) \\ 0 \end{bmatrix} = \begin{bmatrix} f \sin L - g \cos L \\ 1 + (f \cos L + g \sin L) \\ 0 \end{bmatrix} \quad (14)$$

$$\lambda_{r_p}^{RTN} = \begin{bmatrix} (f \sin L - g \cos L) \begin{bmatrix} \dfrac{2\left(1 - \sqrt{f^2 + g^2}\right)}{1 - f^2 - g^2} \\ -\sqrt{f^2 + g^2} \end{bmatrix} \\ \left\{ \dfrac{2\left(1 - \sqrt{f^2 + g^2}\right)(1 + f \cos L + g \sin L)}{1 - f^2 + g^2} \right. \\ \left. -\left(\dfrac{f \cos L - g \sin L}{\sqrt{f^2 + g^2}} + \cos E \right) \right\} \\ 0 \end{bmatrix} \quad (15)$$

The sail pitch and Eq. (12) are both defined along the sail–sun line; it is thus required that we convert the unit vector from RTN axes into sun-centered axes, bearing in mind that the sail is in a planet-centered orbit and not a heliocentric orbit. With conversion of λ into the sun-centered coordinate system, we can define $\tilde{\alpha}$ using Eq. (16). The sail clock angle is derived directly from λ in the sun-centered coordinate system, using Eq. (17):

$$\tilde{\alpha} = \arccos(\lambda_x) \quad (16)$$

$$\delta = \arccos\left(\frac{\lambda_z}{\sqrt{\lambda_y^2 + \lambda_z^2}} \right) \quad (17)$$

Application of Eq. (12) gives the sail pitch angle required for either the locally optimal variation of semimajor axis or radius of pericenter, depending on the locally optimal vector that is utilized. The sail clock angle is found directly from the unit vector once it has been converted into a sun-centered coordinate system as no optimization is required; the sail acceleration magnitude is not related to the sail clock angle. The sail force is hence maximized along the direction defined by λ in Eqs. (14) or (15).

We can thus generate the sail control angles that provide the locally optimal variation in radius of pericenter or semimajor axis; by coupling this information, we can generate an algorithm that avoids planetary collision while maintaining all of the benefits of locally optimal sail control laws.

Blending Sail Control Laws

The blending of locally optimal control laws has previously been discussed for low-thrust orbit transfers where no constraint is placed upon the thrust vector orientation, such as orbit transfers by solar electric propulsion (SEP).[23] Additionally, the rudimentary technique of blending control laws for geocentric solar-sail orbit transfers has previously been introduced.[17]

The blending of control laws is accomplished by initially calculating the unit vector, in sun-centered coordinates, along which the force should be maximized in order to maximize the rate of change of each individual element being blended, obtaining a separate unit vector for each control law. We then compute the blended vector by applying Eq. (18) as

$$\lambda_b = \frac{\sum_k W_k \lambda_k}{\left| \sum_k W_k \right|} \quad (18)$$

where k represents each individual control law being blended and subscript b indicates the blended vector. From λ_b we define $\tilde{\alpha}$ using Eq. (16); thus, application of Eq. (12) gives the blended locally optimal sail pitch angle. The sail clock angle is again derived directly from the optimal thrust vector, using Eq. (17).

A significant difference between the approach taken for SEP-type propulsion and that previously outlined for solar sailing is that the weight functions are to be independent of time. Prior blending methods have used optimization techniques to set the weight functions for each control law; thus, the weightings are given as a function of time from start epoch. However, this means that although the individual control laws are a function of only the orbit elements, the final blended optimal force vector is a function of time hence negating the benefits of defining each control law independent of time. Defining the weight functions as functions of only the orbital elements offers additional benefits. Because the sail control angles are now only a function of the osculating elements, the control system is able to adjust for small unforeseen orbit perturbations or perturbations that cannot currently be accurately modeled because of a lack of real-world knowledge, such as sail wrinkles or sail degradation caused by radiation. As such, the system would potentially be suitable as an onboard autonomous controller, significantly reducing the amount of data in the uplink telecom budget. The sail would thus require only its current position rather than an entire new set of control angles, reducing uplink data requirements.

The optimality of the blended system depends heavily on the weight functions applied in obtaining the blended locally optimal thrust vector. In general, appropriate definition is obtained through good engineering judgment, experience of the system, and some trial and error. Although it is possible to obtain planetary escape through the use of only the pericenter controller, this would result in a greatly increased escape time caused by the inefficiency of this controller in gaining orbit energy. As such, we wish to only use the pericenter control law when it is absolutely required. In Eqs. (19–21) we see the weight of each control law is a function of the radius of pericenter:

$$W_a = \exp(W_f / 10) \quad (19)$$

$$W_{r_P} = 2500 / \exp(W_f) \quad (20)$$

$$W_f = \left(r_P \big/ 10^6 - 2.5 \right) \quad (21)$$

As the pericenter drops toward undesirable values, the pericenter controller becomes more prominent than the energy-gain controller, and as pericenter increases the energy-gain controller becomes more prominent. Furthermore, in order to ensure a rapid changeover between the control laws a set of exponential weight functions are employed. A rapid changeover between controllers is desirable in this scenario as a result of the rapid nature of pericenter decrease toward the end of the escape trajectory. In a more general sense a difficulty encountered in generating transfer trajectories with more than one control law is that the controller can become stuck in a dead-band region, where it is caught between selection of each control law and the orbit parameters alter very little. However, this characteristic can be turned into an advantage, where blended control laws have been used to generate stationkeeping algorithms for potential future solar-sail missions, such as GeoSail,[24] Geostorm,[25] and a Mercury Sun-Synchronous Orbiter.[26]

Initial Orbit Selection

Low-cost launch options are somewhat limited and tend to place the payload into a prohibitive orbit for solar-sail performance, as a result of Earth's steep gravity-well, short-orbit periods that require rapid slew maneuvers and the residual upper atmosphere affecting solar sails up to 1000–1500 km and perhaps beyond at times of solar maxima. However, for completeness it is necessary to consider nonoptimal initial orbits caused by parallel applications in-orbit about other planetary bodies, such as a sample return mission.[2] Earth escape from high energy orbits is however of practical interest, for example, a piggyback launch opportunity to a 72-h Earth orbit with a

future science mission, similar to the INTEGRAL[‡] spacecraft, with a perigee altitude of 10,000 km well above the upper atmosphere would provide an attractive initial orbit for solar sail operations. It was found however that even such high-energy orbits can have a rapid reduction in perigee and could still traverse the upper atmosphere. Three potential options are presented as initial orbits for solar sail Earth escape; the first is from geostationary transfer orbit (GTO), the second is from GEO, and the third a 1000-km-altitude polar orbit. The three potential orbits cover a wide range of initial conditions, and thus a comprehensive test case for the control algorithms and weight functions is presented.

The trajectory model once again uses modified equinoctial elements in the equations of motion and unless otherwise stated includes third-body gravity perturbations caused by the sun and the moon, models the Earth as a nonspherical body up to the 18th order, models the sail as a nonideal reflector (89%) utilizing an optical force model[11] and the sun as a uniformly bright finite disk. The model also includes both Earth and lunar shadow, while differentiating between umbra and penumbra.

Escape from GTO

GTO has been identified by many studies as a potential starting orbit for solar-sail missions, particularly by the DLR ODISSEE concept.[27] GTO is taken to be similar to the Ariane 5 auxiliary payload ring (ASAP) delivery,[28] giving a perigee altitude of 560 km. At this altitude the solar sail will experience air drag and aerodynamic torque; as such, the blended-sail control law is altered so that when the sail altitude is below 1000 km the sail is continually slewed to maintain a minimum drag, edge on, profile to the atmosphere. This minimum profile approach will have the additional benefit of significantly reducing gravity-gradient effects across the sail surface, which will aid attitude control system design. The sail moves through the atmosphere with negligible aerodynamic loading on the sail structure and allows GTO to be considered as a realistic initial orbit, although sail slew rates are still required to be high. This addition to the sail control strategy is adopted only for GTO escape trajectories; however, it would be valid for any high eccentricity orbit with pericenter inside the upper regions of the planetary atmosphere, such as a Molniya orbit.

The standard Ariane 5 midnight launch places the payload on an orbit with a sun-pointing apogee. We consider this midnight option along with the nonstandard midday launch, which places the payload on an orbit with a sun-pointing perigee. An ASAP launch would be an auxiliary payload, and the GTO orbit alignment would be defined by the primary payload requirements. It was found that standard midnight launches result in the sail striking the planet before escape for all sail accelerations above 0.3 mm s^{-2} when the semimajor axis controller is used exclusively. This is shown at the left-hand side of Fig. 6, where the escape duration drops to zero days, indicating a planetary collision. The midday launches however do not repeat this when using the single controller. Instead we find Earth collision occurs only for high sail characteristic accelerations once the total number of orbit revolutions prior to escape is small. We see from Fig. 6 that when the blended controller, previously described and by Eqs. (19–21), is used the trajectory no longer strikes the planet for either the midday or midnight launch options. Comparison of the blended control law with the energy gain control law shows only a small increase in escape time for midday launches as a result of raising and then maintaining perigee altitude above the upper atmosphere. Thus, we can conclude that the inclusion of the additional element of sail law to reduce aerodynamic loads on the sail has negligible impact on sail escape performance, yet has a potentially significant impact on reducing sail loads.

Selecting the specific case of sail characteristic acceleration 2.0 mm s^{-2}, we can examine in detail the behavior of the control laws and the effect of the weight functions. The control system is visualized in Fig. 7, where we see the single controller allows the
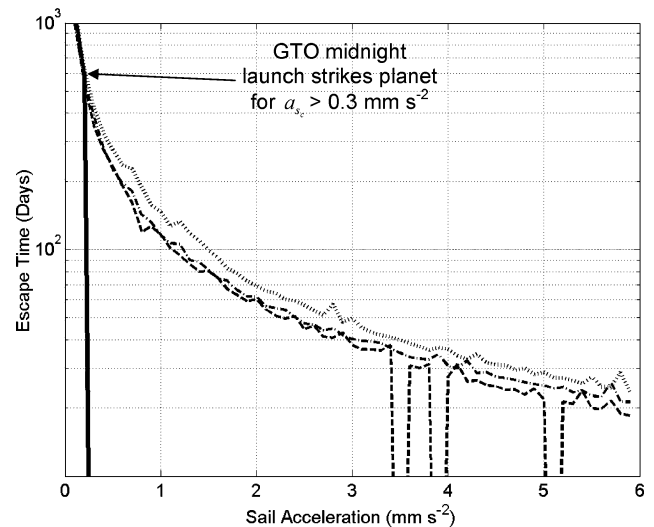
Fig. 6   GTO escape times for midnight and midday launch using semimajor axis control only (——— and – – –, respectively) and blending controllers for midnight and midday launch (· · · and ·–··–, respectively), where escape time is given as zero trajectory intersects planet.
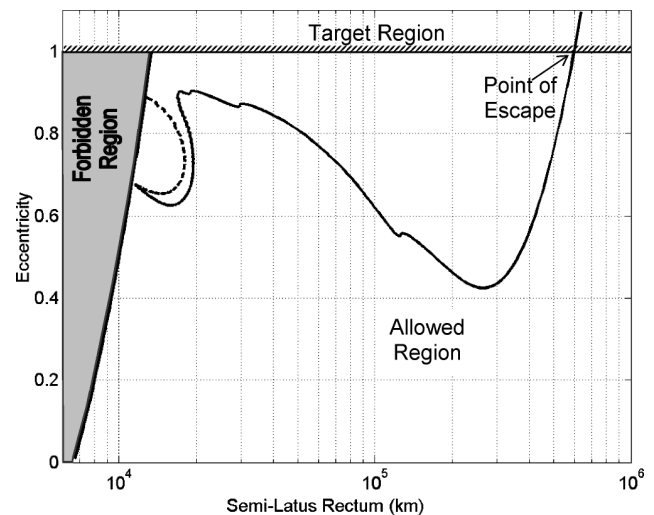

Fig. 7   Visualization of single (. . . .) and blended (———) control laws for escape spirals from midnight GTO. Forbidden region of phase space corresponds to a minimum altitude of 200 km.

orbit to enter the forbidden phase space on the semilatus rectum-eccentricity phase space diagram, whereas the blended controller realizes it is approaching this region and acts to avoid it, hence safely navigating the solar sail toward the target region and Earth escape.

Escape from GEO

GEO represents perhaps the most attractive initial orbit, with a large orbit radius well outside the steep gravity well and air drag associated with low Earth orbit (LEO). Furthermore it is attainable at relatively low cost as a Delta IV auxiliary payload (ESPA).[29] It has been assumed previously that the issues of air drag and aerodynamic torque on a solar sail need not be considered for an escape spiral beginning at GEO.[29] However, it has been found that this assumption breaks down for the locally optimal energy-gain control law at high sail accelerations when the number of orbits until escape is low, causing a rapid variation in eccentricity and hence pericenter altitude during the short escape spiral. Thus, it is required that we use the blended control system. It is also found that this low perigee passage occurs just prior to a reduction in the number of orbits required for escape, as seen in Fig. 8.

We see in Fig. 8 the required time until escape from GEO against sail characteristic acceleration and the typical exponential drop-off
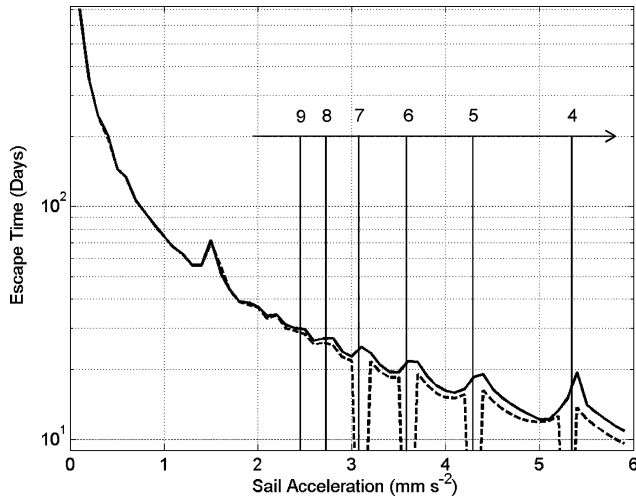
Fig. 8 Escape time from GEO using semimajor axis control only (....) and blended control (——), where escape time is given as zero trajectory intersects planet. Orbit revolutions prior to escape indicated.

Fig. 9 GEO escape using blended-sail control: ...., unperturbed model, where only a perfectly reflecting sail acts to alter the orbital elements.

rate is clearly visible and corresponds well with previous work.[15] However, we also note in Fig. 8 that the exponential drop-off is mixed with a short-period oscillation, seen as maxima and minima within the exponential curve. Each maxima corresponds to a reduction by one in the number of orbits required to reach escape energy. This jagged curve is a unique characteristic of solar-sail propulsion caused by the inability of a solar sail to gain orbit energy while traveling toward the sun. Hence if a sail falls just short of escape energy as it reaches the maximum distance from the sun, it must then complete a half-revolution before gaining the required orbit escape energy. A small increase in sail acceleration will, however, result in the sail acquiring escape energy just before this maximum turning point and time until escape thus appearing reduced. In reality the exact locations of these spikes in escape time would be difficult to predict and hence take advantage of, or conversely ensure against encountering, as a result of trajectory model uncertainties, calculation errors, and launch date uncertainties. Thus, such maxima and minima would make advanced mission planning awkward, as the exact escape epoch would be difficult to predict; hence, ensuring escape for an optimal planetary transfer trajectory would be problematical and require an in-built margin in the planetary escape phase. We see in Fig. 8 when sail acceleration is low both controllers provide almost identical results and up to a characteristic acceleration of 3 mm s^{-2} the escape time are similar. At no time does the blended controller allow the sail to pass below the 1000-km-altitude limit.

Figure 9 shows the escape time using the blended-sail control in a model containing orbit perturbations and one neglecting all perturbations other than the idealized sail thrust vector. Once again the weight functions defined in Eqs. (19–21) are used. We see in Fig. 9 that the blended-sail control system is able to adjust sail orientation with respect to time, thus correcting for a different set of perturbations from those within the original design scenario. The ability of the control system to adapt in the presence, or absence, of perturbations not originally considered is caused by the nature of the individual control laws, where the desired sail orientation is defined by the current orbital elements and not by a stored data file as would be required if attempting to follow a true-optimal trajectory. Thus, if the sail is not where we originally predicted the onboard system automatically adjusts, correcting for the unforeseen perturbation, while maintaining the near-optimal nature of the original trajectory. This self-correcting feature of the control system offers the potential to reduce the required uplink telemetry, as only the current sail state vectors are required, rather than an entire new set of control angles.

Earth Escape from 1000-km Polar Orbit

A high polar orbit within the LEO environment can be achieved as a dedicated low-cost launch, for example, through use of a
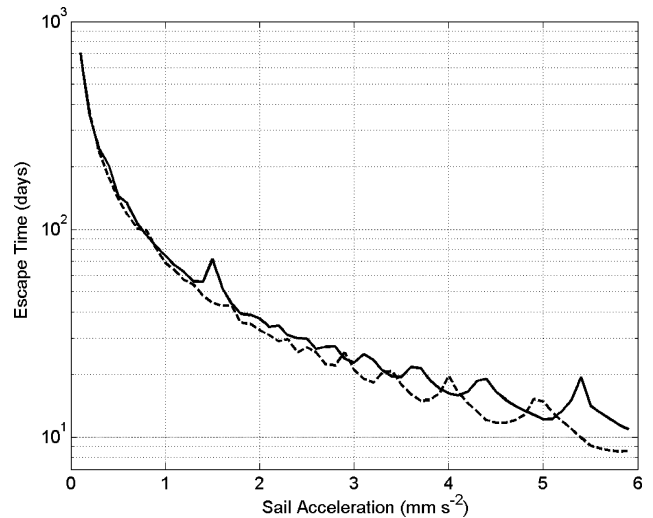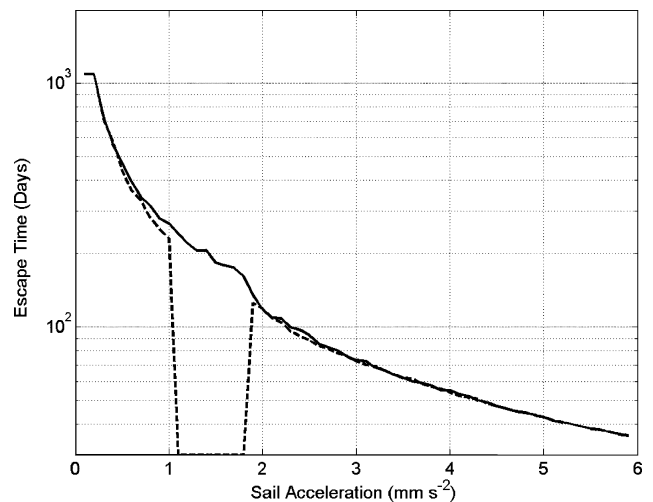
Fig. 10 Escape times from a 1000-km polar orbit. Semimajor axis control (....) and blended control (——) are both presented. Note that the escape time for a sail of acceleration of 0.1 mm s^{-2} was found to exceed five years, and maintaining calculation accuracy made calculation of escape time prohibitive.

Dnepr launcher,[§] or the new Arianespace Vega launcher[30] due for first launch post-2006. Several advantages have been identified that could make this an attractive option for future sail missions; however, as many problems as benefits would remain, and escape from a 1000-km polar orbit would be a significant engineering challenge.[29,31] The potential parallel applications at other planets however require that the control system be able to safely guide the sail to escape from such a low energy orbit.[2] We note that this control strategy was successfully adapted in Ref. 2 for Mercury capture and escape trajectories as part of a sample return mission study.

Escape times from a 1000-km polar orbit are shown in Fig. 10, where the initial orbit normal is aligned with the Earth–sun line and calculation start epoch set at vernal equinox. We see that the semimajor axis controller causes the sail trajectory to intersect the Earth for most sail characteristic accelerations between 1 and 2 mm s^{-2}. It is shown that the blended control system is able to steer the sail to escape without colliding with the Earth. Additionally, it is found that the escape times are within 5% of the semimajor axis control times, except close to the region when this controller breaks down

and safe escape times tend toward 10% longer in duration. Although the weights given in Eqs. (19–21) were once again used for this escape time scan, an additional condition was added such that the semimajor axis controller was used exclusively if eccentricity were less than 0.07 and perigee altitude were greater than 500 km. The additional conditions were found to improve optimality, which was compromised because of the low orbit energy of LEO. A similar approach was found to be helpful in generating escape trajectories for sample return missions.[2]

Earth Escape Without Shadow

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

Using the blended control algorithms outlined in Eqs. (19–21) and the trajectory model already outlined within this paper, the required sail characteristic acceleration for escape from an Earth polar orbit at a range of altitudes was found (Fig. 11). We define the initial orbit such that the initial orbit normal is aligned with the Earth–sun line, and calculation start epoch is set at vernal equinox. The orbit model utilized only considers perturbations caused by the sail acceleration; the introduction of other perturbations such as gravitational harmonics or a more realistic sail force model significantly prolongs calculation time and from experience typically alters escape time by between 3 and 5%. Furthermore, it has been shown in this paper that the control system can correct for perturbations not included in the original design strategy. The sail acceleration at each altitude was incremented in steps of 0.01 mm s^{-2} until an escape trajectory was achieved without any shadow events, the initial altitude was then increased by 50 km, and the process repeated. In Ref. 17 a locally optimal control law for variation of the ascending node angle was presented. It was found, however, that introducing this into the blending equations produced an unnecessary complication within the control system and produced much slower escape times because of the tendency of this controller to cancel out any energy gains over the orbit period. Hence, only the blended control algorithms utilized earlier within this paper and in Eqs. (19–21) are utilized in this section.

The required sail characteristic acceleration for a range of initial altitudes from 800 to 25,000 km is shown in Fig. 11, where we see an exponential increase in sail acceleration requirements as altitude is decreased in order to maintain a shadow free escape trajectory. This exponential curve is analogous to the well-documented exponential reduction in escape time as initial altitude is increased for a given sail acceleration or the exponential reduction in escape time for a given altitude as sail acceleration is increased, as seen in prior figures within this paper. The corresponding minimum time for shadow free escape from each altitude is shown in Fig. 12. As would be expected from the exponential curve of required sail accelerations in Fig. 11, the minimum escape time for shadow-free trajectories is essentially independent of initial altitude, as the required sail acceleration varies exponentially thus maintaining a constant escape time. The mean escape time was found to be 141.46 days; the standard deviation in the escape time data is 6.1 days. In Ref. 15 a single shadow-free Earth escape trajectory is produced, using a locally optimal radius of apocenter control law, the initial altitude of this trajectory was 20,000 km, for a sail characteristic acceleration of 0.85 mm s^{-2}. From Fig. 11 we see that this point is above the presented curve, and hence the two results correspond well. Note the trajectory presented in Ref. 15 has an escape time of 146 days, again corresponding with Fig. 12.
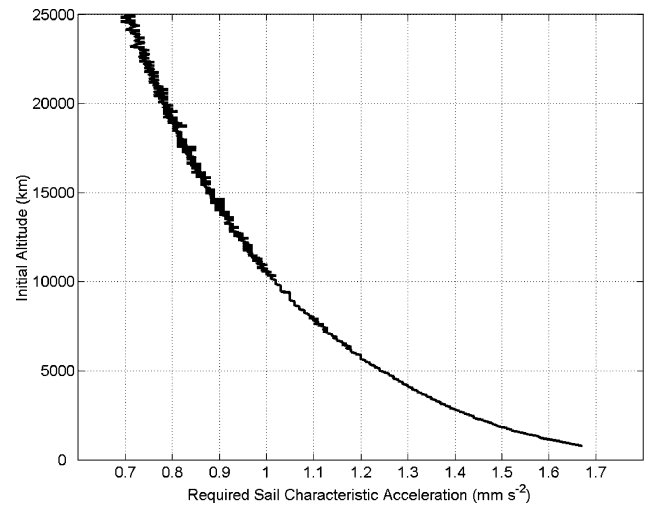

Fig. 11   Required sail characteristic acceleration for shadow-free Earth escape, from a polar orbit.
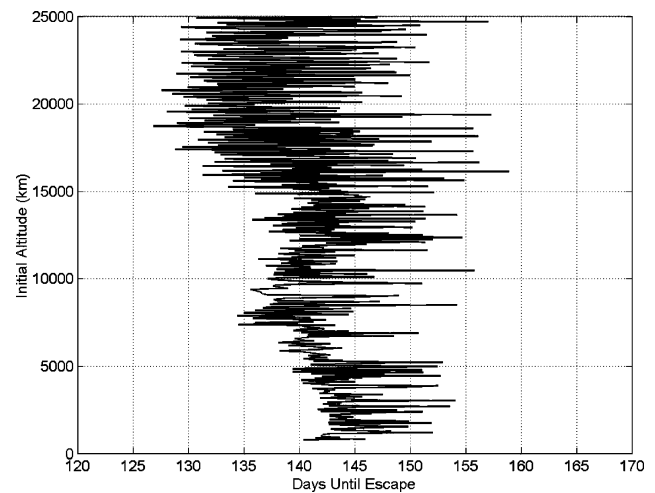

Fig. 12   Minimum shadow-free escape time from each altitude in Fig. 11.

To quantify the true effect of neglecting all orbit perturbations other than sail acceleration, we investigate the single case of escape from 20,000 km, as studied in Ref. 15. The orbit model we use is the model introduced earlier within this paper and includes gravitational harmonics up to and including the 18th-order terms, lunar and solar gravity, corrects the sail acceleration for the true solar distance, models the sun as a uniformly bright finite disc, and models the sail force using an optical force model from Ref. 1. Figure 11 indicates the sail acceleration required is approximately 0.8 mm s^{-2}; however, we must also account now for the inclusion of orbit perturbations and an optical sail force model; we thus increase sail characteristic acceleration to 0.85 mm s^{-2}. Using Eqs. (19–21), we propagate the escape trajectory for these initial conditions as in Fig. 13, where we see the escape trajectory viewed in a fixed sun-axis reference frame, looking from the sun toward the Earth. It is seen from Fig. 13 that at no time does the trajectory pass behind the Earth, and hence no terrestrial shadow events are recorded. This result is verified by analytical analysis of Earth, sun, and spacecraft position vectors and including a 2% addition to the Earth's radius, accounting for the increase in shadow size caused by the atmosphere. Figure 14 shows the orbit inclination and ascending node angles. We see that the ascending node angle initially increases slowly for the first 100 days, before then rapidly increasing for the final 40 days prior to escape on day 141, five days prior to the similar trajectory in Ref. 15, which used only a radius of apogee controller. We note that the minimum altitude of this trajectory is 2397.2 km, on the 116th day of the trajectory.
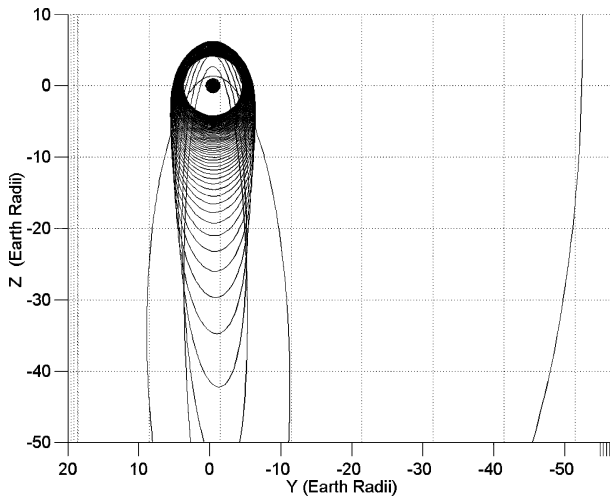
Fig. 13   Shadow-free escape from 20,000-km altitude seen from a fixed sun-line coordinate system, with sail acceleration of 0.85 mm s^{-2}.
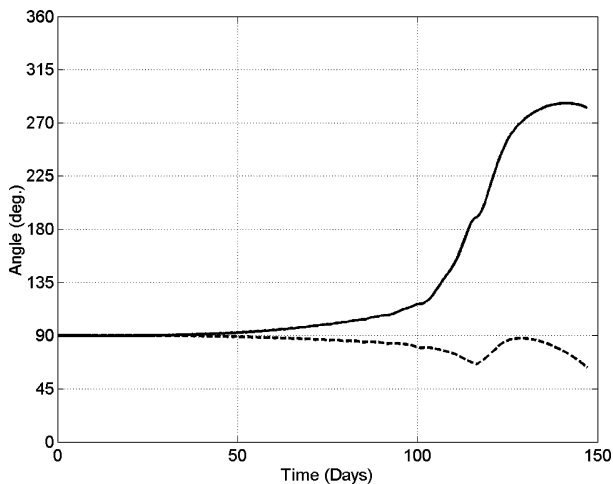

Fig. 14   Inclination (. . . .) and ascending node (———) angles for trajectory in Fig. 13.

Conclusions

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

A method of blending locally optimal control laws was presented that maintains the near-optimal nature of prior locally optimal energy gain controllers, but also maintains a safe minimum altitude through use of a pericenter control law. The algorithms presented are explicitly independent of time and have been shown capable of adapting to different perturbations from those included within the original design scenario. Thus, the control algorithms are potentially suitable as an autonomous onboard controller.

Finally, the required sail acceleration to escape from a polar orbit without Earth occultation of the sail/sun line was investigated. It was shown that the required sail acceleration increases exponentially as initial altitude is decreased. It was also therefore seen that the time until escape corresponding to the minimal sail acceleration requirement was largely independent of initial altitude, with an approximately duration of 142 days.

References

[1]McInnes, C. R., *Solar Sailing: Technology, Dynamics and Mission Applications*, Springer-Verlag, London, 1999.

[2]McInnes, C. R., Hughes, G., and Macdonald, M., "A Low-Cost Mercury Sample Return Mission with Solar Sails," *Aeronautical Journal*, Aug. 2003, pp. 469–478.

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

[4]Eguchi, S., Ishii, N., and Matsuo, H., "Guidance Strategies for Solar Sail to the Moon," *Advances in the Astronautical Sciences*, Vol. 85, Pt. 2, 1993, pp. 1419–1433.

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

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

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

[8]Irving, J. H., *Space Technology*, Wiley, New York, 1959.

[9]Lawden, D. F., "Optimal Escape from a Circular Orbit," *Astronautica Acta*, Vol. 4, June 1958, pp. 218–234.

[10]Green, A. J., "Optimal Escape Trajectories From a High Earth Orbit by Use of Solar Radiation Pressure," M.S. Thesis, T-652, Massachusetts Inst. of Technology Cambridge, MA, 1977.

[11]Sackett, L. L., "Optimal Solar Sail Planetocentric Trajectories," R-1113, Charles Stark Draper Lab., Inc., JPL-NASA Contract NAS 7-100 Final Rept., Sept. 1977.

[12]Sackett, L. L., and Edelbaum, T. N., "Optimal Solar Sail Spiral to Escape," Advances in Astronautical Sciences, Paper A78 31-901, Sept. 1978.

[13]Coverstone-Carroll, V., and Prussing, J. E., "A Technique for Earth Escape Using a Solar Sail," Advances in Astronautical Sciences, Paper 99-333, Aug. 1999.

[14]Coverstone, V. L., and Prussing, J. E., "Technique for Earth Escape from Geosynchronous Transfer Orbit Using a Solar Sail," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 4, 2003, pp. 628–634.

[15]Leipold, M., "Solar Sail Mission Design," Ph.D. Dissertation, Technische Univ. München, DLR-Forschungsbericht 2000-22, Munich, 2000.

[16]Pagel, G., "Extremale Steuerstrategien für Sonnensegler am Beispiel von Bahntransferproblemen zum Erdmond," Ph.D. Dissertation, Technischen Univ., Berlin, May 2002 (in German).

[17]Macdonald, M., and McInnes, C. R., "Analytic Control Laws for Near-Optimal Geocentric Solar Sail Transfers," *Advances in the Astronautical Sciences*, Vol. 109, No. 3, 2001, pp. 2393–2413.

[18]Roy, A. E., *Orbital Motion*, Inst. of Physics, Bristol, England, U.K., 1998, Chap. 6.

[19]Giacaglia, G. E. O., "The Equations of Motion of an Artificial Satellite in Nonsingular Variables," *Celestial Mechanics*, Vol. 15, March 1977, pp. 191–215.

[20]Walker, M. J. H., Ireland, B., and Owens, J., "A Set of Modified Equinoctial Elements," *Celestial Mechanics*, Vol. 36, Aug. 1985, pp. 191–215.

[21]Betts, J. T., "Optimal Interplanetary Orbit Transfers by Direct Transcription," *Journal of Astronautical Sciences*, Vol. 42, No. 3, 1994, pp. 247–268.

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

[23]Kluever, C. A., "Simple Control Laws For Low-Thrust Orbit Transfers," Advances in Astronautical Sciences, Paper 98-203, Feb. 1998.

[24]Macdonald, M., McInnes, C. R., Alexander, D., and Sandman, A., "GeoSail: Exploring the Magnetosphere Using a Low-Cost Solar Sail," *Proceedings of Fifth IAA International Conference on Low-Cost Planetary Missions*, ESA, Noordwijk, The Netherlands, 2003, pp. 341–349.

[25]West, J. L., "NOAA/DOD/NASA Geostorm Warning Mission," Jet Propulsion Lab., JPL D-13986, Pasadena, CA, Oct. 1996.

[26]Leipold, M., and Wagner, O., "Mercury Sun-Synchronous Polar Orbits Using Solar Sail Propulsion," *Journal of Guidance, Control, and Dynamics*, Vol. 19, No. 6, 1996, pp. 1337–1341.

[27]Leipold, M, "ODISSEE—A Proposal for Demonstration of a Solar Sail in Earth Orbit," International Academy of Astronautics, Paper IAA-L98-1005, April 1998.

[28]"Ariane 5 ASAP User's Manual," Issue 1, Rev. 0, Arianespace, Paris, URL: http://www.arianespace.com/site/images/ASAP5-manual.pdf [cited 9 March 2003].

[29]McInnes, C. R., "Final Report: Solar Sail Option for the PROBA II Mission," ESA (ESTEC), Noordwijk, The Netherlands, Contract Rept. Contract 15003/00/NL/PB, Sept. 2001.

[30]"Vega User's Manual," Issue 0, Rev. 0, Arianespace, Paris, URL: http://www.arianespace.com/site/images/VEGAUsersManual.pdf [cited 9 March 2004].

[31]Murphy, D., and Wie, B., "Robust Thrust Sail Control Authority for a Scalable Sailcraft," American Astronautical Society, Paper AAS 04-285, 2004.