
This version is available at https://strathprints.strath.ac.uk/6375/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
Invariant manifolds and orbit control in the solar sail 3-body problem

Thomas J. Waters * and Colin R. McInnes †

Department of Mechanical Engineering, University of Strathclyde, Glasgow, G1 1XJ, Scotland, UK.

In this paper we consider issues regarding the control and orbit transfer of solar sails in the circular restricted Earth-Sun system. Fixed points for solar sails in this system have the linear dynamical properties of saddles crossed with centers; thus the fixed points are dynamically unstable and control is required. A natural mechanism of control presents itself: variations in the sail’s orientation. We describe an optimal controller to control the sail onto fixed points and periodic orbits about fixed points. We find this controller to be very robust, and define sets of initial data using spherical coordinates to get a sense of the domain of controllability; we also perform a series of tests for control onto periodic orbits. We then present some mission strategies involving transfer form the Earth to fixed points and onto periodic orbits, and controlled heteroclinic transfers between fixed points on opposite sides of the Earth. Finally we present some novel methods to finding periodic orbits in circumstances where traditional methods break down, based on considerations of the Center Manifold theorem.

I. Introduction and background

The notion of a solar sail as a practical and flexible space propulsion option is gathering momentum and there are a number of proposed missions utilizing the special properties of solar sails. Two examples of such are the exploration of the far solar system (see for example Dachwald1) and the Geosail concept (see for example McInnes et al.2). This enthusiasm is partly due to advances in materials science, making construction of the sails feasible, and partly due to the emerging understanding of the full dynamical possibilities of solar sails. In particular, the solar sail in the setting of the circular restricted 3-body problem (CR3BP) has produced new and interesting families of orbits which are beyond the capabilities of conventional spacecraft. This is seen most readily in the fixed points: the classical Sun-Earth CR3BP has five isolated fixed points, the Lagrange points, whereas the solar sail CR3BP has a 3-parameter family of fixed points. It is this additional freedom due to the sail parameters which makes the solar sail CR3BP such an interesting generalisation of the 3-body problem. This paper seeks to further extend our understanding of solar sails in the CR3BP by considering the issues of controlling the sail’s position or orbit and transferring the sail between fixed points, periodic orbits, and the Earth itself.

The linear stability properties of solar sails in the CR3BP were first examined in McInnes et al.3 and McInnes,4 where the nature of the fixed points and the linear stability of same were discussed. It was found that there are continuous surfaces of fixed or equilibrium points, and that these points are linearly unstable; we give below a brief recap of the important features of this analysis. The fact that fixed points in the x-z plane have center eigenvalues at linear order can be used to describe (unstable) periodic orbits about the fixed point, and the authors5 have used the Lindstedt-Poincaré method to find families of periodic orbits about fixed points high above the ecliptic plane. There an application of such periodic orbits was suggested for Earth observation: a large amplitude orbit with period one year would counter the effects of the variation in the Earth’s axis of rotation when viewed from the poles. In Baoyin and McInnes6 the authors show how periodic orbits about fixed points on the x-axis may also be described. However the instability inherent in these points and orbits means a controller must be used to maintain these states over long times.

*Research Fellow, thomas.waters@strath.ac.uk, member AAS
†Professor, colin.mcinnes@strath.ac.uk, member AIAA

American Institute of Aeronautics and Astronautics
The natural mechanism of control for a solar sail is in variations of the sail's orientation, and in McInnes and Lawrence and Piggott some preliminary investigations of the robustness of this control mechanism in the CR3BP were presented. In Baoyin and McInnes control in the elliptical 3-body problem was considered, and in Bookless and McInnes and McInnes control of solar sails in 2-body problems was examined. In this paper we will describe an optimal controller which minimizes a cost function, the gains being found from solutions of the algebraic Ricatti equation. This approach is very well known and is used broadly in the literature, and we refer the reader to, for example, Ogata. We consider the efficacy of this controller for fixed points and periodic orbits in Sections II and III respectively. There we present a number of tests to determine the relative controllability of different points and orbits, to answer questions such as: in which direction of displacement from the fixed point is it hardest to control from? How large an orbit may we control directly onto from the fixed point? Which fixed points have the largest domain of controllability?

As we shall see there is a very large set of fixed points and periodic orbits, and presenting the controllability of each would be tedious and distract from the overall behavior. Instead, where appropriate we will consider nominal parameter values, where a fixed point in the $x$-$z$ plane is specified by the parameters ($\beta, \gamma$). As nominal ($\beta, \gamma$) values we will choose $\beta = 0.05174$ for the following reasons: 1) this is a very moderate value within the feasibility of near-future technology, placing the application of this analysis very much in the near term; 2) this parameter value displays all the relevant features and thus is a representative value; and 3) this precise value was used in along with a $\gamma$ value of 0.8092 rad to describe orbits of period one year and some applications of same.

In Section IV we examine a possible mission made up of different legs, each requiring its own control scheme. This mission consists of transferring from the Earth’s vicinity to a high altitude fixed point, transferring from this point onto a periodic orbit, and finally the controlled transfer from a fixed point on the $L_1$ side of the Earth to one on the $L_2$ side. These transfers combine the control strategy outlined above with the use of invariant manifolds. Invariant manifolds are a well known tool in finding low-cost transfers, see for example Gómez et al. and Koon et al. and an advantage in using invariant manifolds for transferring solar sails is that the sail orientation is fixed along the manifold and thus we minimize the maneuvering demands on the sail. We hope to demonstrate with this representative mission the robustness and flexibility of the controlled solar sail. Also, we note how orbits and transfers with solar sails are ‘free’, that is there is no propellant needed for the life of the mission besides that required to insert the sail into its initial orbit.

Finally in Section V we present some novel methods for finding periodic orbits based on considerations of the Center Manifold theorem, and show how this is useful for cases when the traditional Lindstedt-Poincaré method breaks down, for example at fixed points in the $x$-$y$-$z$ space. In Section VI we give some conclusions.

First we briefly recap the important features of the fixed points in the solar sail CR3BP as described in.

### A. Background

We will follow the conventions set out in McInnes, and consider a rotating coordinate system in which the primary masses are fixed on the $x$-axis with the origin at the center of mass. The $z$-axis is the axis of rotation and the $y$-axis completes the triad. We chose our units to set the gravitational constant, the sum of the primary masses, the distance between the primaries, and the magnitude of the angular velocity of the rotating frame to be unity; in this system therefore the unit of distance equals 1.496 × 10^8 km and the unit of velocity equals 29.8 km/s. We shall denote by $\mu = 3 \times 10^{-5}$ the dimensionless mass of the smaller body $m_2$, the Earth, and the radius of the larger body $m_1$, the Sun, is given by $1 - \mu$ (see Figure 1).

Denoting by $r$, $r_1$ and $r_2$ the position of the sail w.r.t. the origin, $m_1$ and $m_2$ respectively, the solar sail’s equations of motion in the rotating frame are

$$\frac{d^2 r}{dt^2} + 2\omega \times \frac{dr}{dt} = a - \omega \times (\omega \times r) - \nabla V \equiv F,$$

with $\omega = \hat{z}$ and $V = -[(1 - \mu)/r_1 + \mu/r_2] \frac{dr_i}{dt} r_i = |r_i|$. Letting $X = (r, \dot{r})^T$ we can write the equations of motion as the first order nonlinear system

$$\dot{X} = f(X; u),$$

where $u$ is the parameter set and a dot denotes differentiation w.r.t. time. These differ from the classical
equations of motion in the CR3BP by the radiation pressure acceleration term
\[ a = \beta \frac{(1-\mu)}{r_1^2} (\hat{r}_1 \cdot n)^2 n, \]  
where \( \beta \) is the sail lightness number, and is the (dimensionless) ratio of the solar radiation pressure acceleration to the solar gravitational acceleration. For the purpose of this paper we are assuming an ideal sail, although generalizations to non-ideal sails are straightforward.

Here \( n \) is the unit normal of the sail and describes the sail's orientation. We define \( n \) in terms of two angles \( \gamma \) and \( \phi \) w.r.t. the rotating coordinate frame,
\[ n = (\cos(\gamma) \cos(\phi), \cos(\gamma) \sin(\phi), \sin(\gamma))^T, \]  
where \( \gamma, \phi \) are the angles the normal makes with the \( x-y \) and \( x-z \) plane respectively (see Figure 1). Throughout the paper these angles will be given in radians.

Fixed points are given by the zeroes of \( F \) in (1), and we find a 3-parameter family of same specified by the triple \((\beta, \gamma, \phi)\), with each set of parameters defining two fixed points, one on the \( L_1 \) side of the Earth and one on the \( L_2 \) side (see Figure 3). When \( \gamma = 0 \) fixed points are in the \( x-y \) plane, and when \( \phi = 0 \) fixed points are in the \( x-z \) plane. By linearizing the system (1) about a fixed point and examining the linear system's eigenvalues, we find that fixed points in the \( x-z \) plane have the linear dynamical properties of centers crossed with saddles, and have the linear spectrum
\[ \{ \pm \lambda_1 i, \pm \lambda_2 i, \pm \lambda_r \}. \]  
On the other hand, fixed points in the \( x-y-z \) space have the properties of stable/unstable spirals crossed with saddles. As we are particularly interested in periodic orbits about fixed points built from the linear centers we will consider of primary interest therefore those fixed points in the \( x-z \) plane \((\phi = 0)\), however in Section V we will discuss also periodic orbits about fixed points outside of this plane.

The form of (5) tells us the structure of the invariant manifolds of the nonlinear system (2), and we may conclude the following (we will drop the term invariant):

\[ \{ \pm \lambda_1 i, \pm \lambda_2 i, \pm \lambda_r \}. \]  

---

**Figure 1.** The rotating coordinate frame and the sail position therein. The angles \( \gamma \) and \( \phi \) which the sail normal makes with respect to the rotating frame are also shown.

**Figure 2.** Projections in configuration space of portions of the invariant manifolds associated with a typical fixed point in the \( x-z \) plane (in this case the nominal point \( \beta = 0.05174, \gamma = 0.8092, \phi = 0 \)).
1. There is one stable and one unstable manifold associated with each fixed point, and they characterize the flow onto and away from the fixed point.

2. There is a four dimensional center manifold upon which trajectories oscillate about the fixed point with a combination of frequencies $\lambda_1$ and $\lambda_2$, in orbits which are a generalization of a set of classical curves known as ‘roses’. In particular, there are two 2-dimensional foliations of the center manifold on which the frequencies separate; these paraboloidal surfaces contain the periodic orbits in each frequency, $\lambda_1$ and $\lambda_2$.

The $\lambda_1$ family consists of roughly vertical orbits, and thus we will call them Halo type or $\mathcal{H}$-type; on the other hand the $\lambda_2$ family is made up of more horizontal orbits and thus we will call them Lyapunov type or $\mathcal{L}$-type (see Figure 2). In the authors described the $\mathcal{H}$-type family of periodic orbits, and in Section III we will describe the $\mathcal{L}$-type family also.

These manifolds are useful because they in a sense summarize the dynamics associated with each parameter set, and show at a glance an approximation to the general structure of solutions to the nonlinear system (1). In Figure 3 we show the important features of the fixed points in the $x$-$z$ plane, as determined from the analysis in.

![Figure 3](image.png)

**Figure 3.** The relevant parameters for fixed points in the $x$-$z$ plane (negative $z$ is not shown due to symmetry in the $x$-$y$ plane). On top we show the $\beta, \gamma$ level curves that specify the position of the fixed point; below we show the values of the linear frequencies $\lambda_1, \lambda_2$. We also show the positions where the linear frequencies are in resonance with $\lambda_1 = 2\lambda_2$ (dashed line). In both, the position of the Earth and $L_1, L_2$ are shown, with the fixed points going to $L_1/L_2$ in the limit $\beta \to 0$ or $\gamma \to \pi/2$. The shaded areas of the plot represent regions where equilibrium is not possible.

### II. Optimal control onto fixed points

Let the subscript $e$ denote a value at an equilibrium or fixed point, such that $F(r_e) = 0$. Letting $r \to r_e + \delta r$ where $\delta r = (\delta x, \delta y, \delta z)^T$, we write the first order linearized system as $\delta \dot{X} = A.\delta X$, with
\[ \delta X = (\delta r, \delta \dot{r}), \quad A = \begin{pmatrix} 0 & I \\ -M & \Omega \end{pmatrix}, \quad \Omega = \begin{pmatrix} 0 & 2 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_{ij} = \frac{\partial F_i}{\partial r_j} \]. \quad (6)

As mentioned above and shown in,\(^5\) \(A\) will admit at least one positive real eigenvalue implying the fixed point \(r_e\) is unstable.

The mechanism of control is to allow the angles \(\gamma\) and \(\phi\) to vary. Some authors consider allowing \(\beta\) to vary however this is practically inefficient and so we will discount it here. Letting \(u = (\gamma, \phi)^T \rightarrow u_e + \delta u\), our controller linear system becomes

\[ \delta \dot{X} = A \delta X + B \delta u = A \dot{X} + B(-G \delta X) = (A - B G) \delta X = A_c \delta X, \quad (7) \]

where \(A = \partial f / \partial X, B = \partial f / \partial u\) and we have introduced the \(2 \times 6\) gain matrix \(G\) via a feedback control law. We seek to define \(G\) such that the controlled linear system \(\delta \dot{X} = A_c \delta X\) is asymptotically stable, that is, the eigenvalues of \(A_c\) all have negative real part.

As this is a multivariable control system it is overdefined and so we may use the freedom to minimize the cost function

\[ J = \int (\delta X^T Q \delta X + \delta u^T N \delta u) dt, \quad (8) \]

where \(Q, N\) are symmetric positive semi-definite and definite weighting matrices respectively. By standard control systems theory,\(^12\) this is solved by \(G = N^{-1} B^T P\), where the performance matrix \(P\) solves the algebraic Riccati equation,

\[ A^T P + PA + Q - PBN^{-1} B^T P = 0. \quad (9) \]

There are a number of methods for solving this equation, however the method we favor here consists of constructing the following Hamiltonian matrix

\[ H = \begin{pmatrix} A & -BN^{-1}B^T \\ -Q & -A^T \end{pmatrix}. \quad (10) \]

A Hamiltonian matrix is any of the form \(((A, M_1), (M_2, -A^T))\) where \(M_1, M_2\) are symmetric, and has the property that if \(\lambda\) is an eigenvalue, then so is \(-\lambda, \overline{X}\), and \(-\overline{\lambda}, \overline{X}\). We compute the eigenvalues of \(H\), and take the first \(n\) as those with negative real part. For each of these stable eigenvalues, compute the eigenvectors \(v\) in \(\mathbb{R}^{2n}\), and partition as \(v = (f_i, g_i)^T\) with \(f_i, g_i \in \mathbb{R}^n\). Define the matrices \(X_1 = (f_i)\), \(X_2 = (g_i)\), and let the performance matrix be \(P = X_2 X_1^{-1}\). From this we calculate \(G\), and this gives the optimally controlled system, according to the weights \(Q, N\).

The weights themselves \(Q, N\) we are free to choose as long as they satisfy the properties mentioned above. We will reduce this freedom to one parameter by letting \(Q = I_6\) and \(N = \rho I_2\) (although we note the possible advantages in weighting some elements of the diagonals over others). Thus by varying \(\rho\) we can choose whether the controller penalizes the control effort (\(\rho\) large) or the distance from equilibrium (\(\rho\) small) according to (8). Typically when control onto a point or orbit is required then \(\rho\) is set to be small, and in this section and Sections III and IV that follow we will let \(\rho = 10^{-4}\) unless stated otherwise. In Section V however we will discuss some control strategies with \(\rho\) large.

With this gain matrix we have forced the linearized controlled system to be asymptotically stable. Now we take this linear control scheme, and use it to define the nonlinear controlled system as

\[ \dot{X} = f(X; u_e + \delta u). \quad (11) \]

The question now reduces to the following: in the vicinity of the fixed point the linear system dominates and trajectories are controlled; how far from the fixed point may we place the sail in phase space such that the resultant trajectory remains controlled? This is equivalent to asking how far the linear approximation extends in phase space.

This ‘domain of controllability’ (DOC) in 6-dimensional phase space will vary from one fixed point to the next, and remembering that there is a 2-parameter family of fixed points in the \(x-z\) plane this means
there is a very large space to search. To give a crude approximation to the DOC we will therefore describe
sets of initial data of increasing distance from the fixed point using spherical coordinates, and using these we
integrate the controlled nonlinear system until one of two things happens: 1) the trajectory is not controlled
and escapes from the vicinity of the fixed point, or 2) the parameters \((\gamma, \phi)\) leave their allowed ranges
\[-\pi/2 < \gamma, \phi < \pi/2.\] (12)

This second condition is due to typical sail design of having only one reflective side and thus we must have \(n \cdot \hat{r}_1 > 0.\)

We do not claim that this test presents a definitive measure of the domain of controllability for each fixed
point; merely that the tests can provide us with a first method of comparing the relative controllability of
different fixed points, and suggesting the directions in which the controller is most inefficient.

With this in mind our first test will be to let the displacement from the fixed point be \(\pm R\) in each
coordinate direction in phase space, that is \(\delta X = (\pm R, 0, 0, 0, 0, \ldots, 0, 0, 0, 0, \pm R),\) and steadily increase
\(R\) until we find a maximum value for each fixed point. Fixing \(\beta,\) this gives an \(R_{\text{max}} - \gamma\) plot as shown in
Figure 4. We can make four observations from this quick test:

1. Fixed points on the \(L_1\) side have a larger DOC than those on the \(L_2\) side (this is because the \(L_2\) side
fixed points move closer to the Earth as \(\beta\) increases, which acts as a strong perturber in contrast with
the \(L_1\) side, see Figure 3).

2. The largest value of \(R_{\text{max}}\) occurs at the value of \(\gamma\) which gives the fixed point which is highest over
the ecliptic plane, that is the largest \(z_e.\)

3. This value of \(R_{\text{max}}\) can be surprisingly large, at best 0.0035 in non-dimensional units, which is equiva-
   lent to a positional displacement of 500,000 km or a velocity displacement of 100 m/s.

4. From monitoring which trajectory first fails the test we see the controller is most sensitive to displace-
   ments in the \(x\)-direction.

![Figure 4. The maximum values of \(R\) in each coordinate direction in phase space for fixed points along the
\(\beta = 0.05174\) level curve. The largest \(R_{\text{max}}\) occurs at \(\gamma = 0.5\) and 0.8 for the \(L_1\) and \(L_2\) fixed points respectively.
The nominal fixed point \(\gamma = 0.8092\) on the \(L_1\) side is only moderately controllable.](image)

Based on the last point here, we consider as a second test fixing \(\delta x = 0,\) and examining how large the
displacements in the other directions may be. We separate the coordinate and velocity directions, and set
initial data as follows: the coordinate displacements are on a circle of radius \(R_1\) in the \(\delta y-\delta z\) plane, and the
velocity displacements are on a sphere of radius \(R_2\) in the \(\delta \dot{x}-\delta \dot{y}-\delta \dot{z}\) space. This leads to the displacement in
phase space having magnitude \(\sqrt{R_1^2 + R_2^2}.\) To define the values of the initial data on the circle and sphere
we use polar/spherical coordinates as follows:

\[
\begin{align*}
\delta y &= R_1 \cos(\theta_1), & \delta z &= R_1 \sin(\theta_1), \\
\delta \dot{x} &= R_2 \cos(\psi_2), & \delta \dot{y} &= R_2 \sin(\psi_2) \cos(\theta_2), & \delta \dot{z} &= R_2 \sin(\psi_2) \sin(\theta_2) \quad \text{with} \quad 0 \leq \theta_1, \theta_2 < 2\pi, \\
& & 0 \leq \psi_2 \leq \pi.
\end{align*}
\] (13)

6 of 15
American Institute of Aeronautics and Astronautics
The position of the initial data value is determined by incrementing the angles $\theta_1$ and $(\theta_2, \psi_2)$ by the fixed amounts $\Delta_1$ and $\Delta_2$ respectively. For example, when $\Delta_1 = \Delta_2 = \pi/2$, there are $4 \times 6$ distinct initial data points, however when $\Delta_1 = \Delta_2 = \pi/4$ there are $8 \times 26$.

Using these sets of initial data we integrate the controlled nonlinear system until the trajectory escapes or the sail tips over ($\mathbf{n}, \hat{r}_1 > 0$). We find that we may increase $R_1, R_2$ to quite large values and still maintain control onto the equilibrium point. For example, if $R_2 = R_1/10$, we may increase $R_1$ about the nominal fixed point ($\gamma_e = 0.8092$) up to a value of 0.0057, see Figure 5.

Finally, the natural question to ask is: which direction is the most controllable? That is, with zero initial velocity, how large a positional displacement may there be and still maintain control? The answer is the $+z$ direction, where we can control up to the order of $\delta z = 0.015 (2$ million km); this is quite remarkable remembering that $L_1$ is ‘only’ 0.01 (1.5 million km) away from the Earth (although this distance will be dwarfed by some of the orbits of the next section).

III. Optimal control onto periodic orbits

In this section we will use a simple discrete method to optimally control onto a periodic orbit. Firstly we choose the nominal orbit we wish to control onto. We see from (5) that there are two families of periodic orbits about fixed points in the $x$-$z$ plane, what we call $\mathcal{H}$-type and $\mathcal{L}$-type, and we have shown some of these orbits in Figure 2 and their frequency values in Figure 3. In,\textsuperscript{5} the authors describe a Lindstedt-Poincaré method to provide high order approximations to periodic $\mathcal{H}$-type solutions to the uncontrolled nonlinear system (2), and the same method applies to the $\mathcal{L}$-type family. A slight modification is to let the linear order solution be

$$
x_1 = k A_y \cos(\lambda_1 t + \xi_1), \quad y_1 = A_y \sin(\lambda_1 t + \xi_1), \quad z_1 = m A_y \cos(\lambda_2 t + \xi_2),
$$

(14)

where $\lambda_1, \lambda_2$ depending on whether we wish to approximate the $\mathcal{H}$- or $\mathcal{L}$-type orbits respectively, and $k, m$ are coefficients made up of components of the eigenvectors. We have chosen to let the free parameter be the $y$-amplitude, $A_y$, as due to the (assumed) symmetry in the $x$-$z$ plane two members of different families will only touch when their $y$-amplitudes are the same. Also, members of the $\mathcal{H}$- and $\mathcal{L}$-type families may pass through vertical or horizontal in which case $A_x = 0$ or $A_z = 0$ respectively, which would lead to singular coefficients if these were chosen as free parameters. The Lindstedt-Poincaré (LP) method uses these linear solutions to build up approximations to periodic solutions to the nonlinear system in the following way: let
\[ \delta x = \varepsilon x_1 + \varepsilon^2 x_2 + \varepsilon^3 x_3 + \ldots \text{ etc.}, \]

where

\[ x_n = \sum_{j=0}^{n} A_j \cos[j(\lambda_n \omega t + \xi_j)] \]

and \( \omega = 1 + \varepsilon^2 \omega + 1 + \varepsilon^2 \omega^2 + \ldots \). Then using the free parameters \( \omega \), to switch off the secular terms in the \( n \)th order Taylor expansion of (2) about the fixed point, we solve a set of algebraic equations for the coefficients \( A_j \) (see\(^5\) for a more detailed discussion).

The result is a trigonometric Cosine and Sine series parameterised by \( A_y \) in the \((x, z)\) and \( y \) coordinates respectively. We find it most convenient to use \textit{these} analytic representations as nominal orbits to control onto, rather than the numerical orbits found from using a differential corrector to fine tune initial data, for the following reasons:

1. The numerical orbits can only be found for amplitudes up to the limit of the approximations providing convergent initial data (without using numerical continuation, which is not really an option for a 2 parameter family of fixed points); on the other hand, though the approximations do not provide convergent initial data they do provide a close enough approximation to the orbit itself for the controller to control onto, even for large amplitudes.

2. These analytic representations are \textit{exactly} periodic rather than the almost periodic numerical orbits.

The procedure therefore is straightforward: we split the nominal orbit into a series of points by incrementing the period, and at each define the constant gain matrix using the same methods in Section II as if the point on the orbit were a fixed point. Then the components of the gain at each moment are interpolated to provide the period, and at each define the constant gain matrix using the same methods in Section II as if the point

\( \text{Controller to bring us onto the nominal orbit?} \)

\( \text{Controller to bring us onto a periodic orbit?} \)

We will give the details of each test in turn.

1. The initial data point provided by the approximations is in the \( x-z \) plane (although we can easily choose somewhere else along the orbit), thus we set the initial coordinates to be on a circle in the \( x-z \) plane centered on the nominal initial data position, and the velocities to be on the sphere centered on the nominal initial velocity, as in (13). With this setup, we integrate the periodically controlled nonlinear system until the trajectory escapes or the sail tips over \((n, \hat{r}_1 > 0)\). With this method we can examine the controllability of the nominal periodic orbit; however there is a 4-parameter family of periodic orbits (2 parameters specifying the fixed point and the two families \( \mathcal{H} \) and \( \mathcal{L} \) about that point), so for the sake of brevity we will not present an exhaustive search here. Instead we will focus only on the nominal point \( \beta = 0.05174, \gamma_e = 0.8092 \). For \( \mathcal{H} \)-type orbits about this point, and letting \( R = R_1 = R_2 \), we find that as \( A_y \) increases from 0.001 to 0.005, the maximum value of \( R \) increases steadily from 0.0007 to 0.00085. On the other hand, \( \mathcal{L} \)-type orbits have a maximum \( R \) increasing steadily from 0.00085 to 0.00115, all other things being equal.

2. This is a useful question with regards transfer as we shall see in Section IV. We take as our initial data \( X = (x_e, y_e, z_e, 0, 0, 0)^T \), and integrate the controlled nonlinear system using the gain defined by a specified orbit either of \( \mathcal{H} \)- or \( \mathcal{L} \)-type with amplitude \( A_y \). This gives a largest amplitude orbit we may control onto from that fixed point, and in Figure 6 we show the largest amplitude of each family for fixed points along the \( \beta = 0.05174 \) level curve. Control onto the \( \mathcal{L} \)-type family is relatively unremarkable, the amplitude we may control onto hovering around the \( A_y = 0.005 \) mark mostly, with the orbits themselves being as expected. On the other hand, the \( \mathcal{H} \)-type family show a large dip

\( \text{American Institute of Aeronautics and Astronautics} \)
around the $\gamma = 0.7$ point, and this is due to a resonance effect: there is a fixed point along this $\beta$ level curve ($\gamma = 0.751$) where the linear frequencies satisfy $\lambda_1 = 2\lambda_2$ (see Figure 3). Now when we try to control onto the $H$-type orbit about this point, the following happens: the orbits the sail controls onto become more egg-like and develop a cusp at one end; passing through this cusp a doubly-periodic orbit emerges. We will not discuss further the nature of such orbits. We do not see this resonance in the $L$-type orbits.

In Figure 7 (a) we show an example of control onto periodic orbits from the fixed point. The orbits we may control onto have moderately large amplitudes, and for large amplitudes nonlinear effects become prominent: the orbits the control effort brings the sail onto become asymmetric and displaced from the nominal orbits of the LP approximation.

3. With this test we see very large amplitude orbits, and perhaps unsurprisingly equally strong nonlinear effects. The amplitudes of the orbits the controller brings the sail onto have especially large amplitudes when $\gamma$ is small, which corresponds to the sail facing onto the Sun. While the $y$-amplitudes may look moderate, the orbit is elongated in the $z$-direction and the resulting orbit can have a $z$-amplitude the order of 0.06 (10 million km)! This is really quite remarkable and is 6 times the distance of $L_1$ from the Earth (see Figure 7 (e)).

![Figure 6. The largest amplitude orbits for members of the different families about fixed points on the $\beta = 0.05174$ curve on the $L_1$ side of the Earth. On the left we show control from the fixed point onto the orbit (test 2), and on the right control from the initial data provided by the LP approximations (test 3).](image)

We can conclude that with regards periodic orbits the $L$-type family are the more robust, and that combined with a controller the solar sail can be placed on large amplitude periodic orbits of varying period, position and inclination with small control effort.

### IV. Mission strategies and transfers

In previous sections we have attempted to give an overview of the effects of an optimal controller with regards fixed points and periodic orbits. In this section we will try and show how these methods can be used for a practical mission, and in the process we hope to demonstrate the range and flexibility of the controlled solar sail.

The mission we have in mind consists of the following stages:

1. Transfer of the solar sail from the Earth’s vicinity to the nominal fixed point $\beta = 0.05174$, $\gamma_e = 0.8092$ on the $L_1$ side of the Earth.

2. Moving from this fixed point onto a one-year orbit for an arbitrary time and then back to the fixed point.

3. Finally transferring from a fixed point on the $L_1$ side to a fixed point on the $L_2$ side of the Earth with the same $\beta$ value.

A possible application of this scenario is to facilitate transfer from the day-side of the Earth to the night-side for science missions, as well as the possible uses of high elevation fixed points and one year orbits already mentioned.
Figure 7. Some examples of control onto periodic orbits. In (a) we show the control onto \( L \)- and \( H \)-type orbits of amplitude \( A_y = 0.0035 \) from the nominal fixed point \( \beta = 0.05174, \gamma_e = 0.8092 \) as in test 2; in (b) and (c) we show the control effort for the \( L \)- and \( H \)-type resp. In (e) we show the control from initial data onto an \( H \)-type orbit with \( A_y = 0.0052 \) about \( \beta = 0.05174, \gamma_e = 0.8092 \) as in test 3, with the control effort shown in (d). In both (a) and (e) the Earth and \( L_1 \) are shown, and in (b),(c),(d) the solid line is \( \delta \gamma = \gamma - \gamma_e \) and the dashed line is \( \delta \phi \).

Transfer in the 3-body problem is most efficiently achieved by taking advantage of the dynamics of the problem setting via the invariant manifolds. As mentioned in Section I, fixed points in the \( x-z \) plane will (without control) have one stable mode and one unstable mode due to a negative and positive real eigenvalue respectively. As the linear eigenspaces are tangential to the invariant manifolds at the fixed points, we can numerically approximate the stable/unstable manifolds by integrating the system (2) in the direction of the stable/unstable linear eigenvectors either backwards or forwards in time. The solar sail 3-body problem shares the \( y \) and \( z \) symmetries of the classical 3-body problem, that is

\[
\begin{align*}
\{x, y, z, \dot{x}, \dot{y}, \dot{z}, t\} & \rightarrow \{x, -y, z, -\dot{x}, \dot{y}, -\dot{z}, -t\}, \\
\{x, y, z, \dot{x}, \dot{y}, \dot{z}, t\} & \rightarrow \{x, y, -z, \dot{x}, \dot{y}, -\dot{z}, t\},
\end{align*}
\]

(16) (17)

thus the stable manifold of a fixed point in the \( x-z \) plane is the image in the \( x-z \) plane under time reversal of the unstable manifold, and vice versa.

The practical use of these invariant manifolds is that if we can insert the sail onto the stable manifold of a fixed point in phase space then the dynamics of the problem will naturally and freely carry the sail onto (or at least to the close vicinity of) the fixed point; also, if a sail is placed uncontrolled at a fixed point, the unstable manifold describes how the sail will fall away from the point and where it will naturally tend towards. What’s more, transfers where the sail orientation is fixed (that is, along invariant manifolds) will be easier to manage then those which require variations in the orientation.

In searching for a candidate trajectory for the first leg of the mission, we simply take a series of points along the \( \beta \) level curve of varying \( \gamma \) values and integrate (2) backwards in time in the direction of the stable eigenvector; when we find one that passes close to the Earth this is a candidate for transfer. In Figure 8 we show how the stable invariant manifold of the fixed point \( \beta = 0.05174, \gamma = 0.872 \) passes close to the Earth and may be used for the transfer of the first leg. As the sail approaches the \( \gamma = 0.872 \) fixed point the sail is controlled onto the \( \gamma = 0.8092 \) fixed point. The control effort for these two stages is shown in Figure 9 (a) and (b).

The second leg of the mission has been discussed in Section III. The nominal orbit about the nominal fixed point is of \( H \)-type with amplitude \( A_y = 0.0026 \) (as discussed in$^5$), and we may control onto this
orbit from the fixed point as this amplitude is well within the acceptable range as shown in Figure 6. The trajectory is shown in Figure 8. We note that were transfer to the orbit the primary objective, it is easy to skip the intermediate stage of putting the sail at the fixed point and going directly onto the orbit, by simply controlling from a point on the stable manifold of the fixed point and going directly onto the orbit. Alternatively, trajectories in the stable invariant manifold of the periodic orbit which pass close to the Earth could be used, as mentioned in, although we will not pursue that option in this work. To return to the fixed point for the next stage is also typical as discussed in Section II (by simply controlling onto the nominal fixed point from some point on the orbit). The control effort for these two stages is shown in Figure 9 (c) and (d).

The final leg consists of a transfer from a fixed point on the side to a fixed point on the side of the Earth. With regards the stable/unstable invariant manifolds, there are two possible scenarios worth mentioning; they are homoclinic and heteroclinic paths. A homoclinic path joins an equilibrium point to itself, that is the unstable manifold is identical to the stable manifold and thus there is a trajectory that leaves the fixed point only to return to it. A heteroclinic path is a trajectory which joins two distinct equilibrium points, that is the unstable manifold of one point meets smoothly the stable manifold of another. Note these are asymptotic trajectories, and thus for practical reasons we must deal with their approximations by integrating the system (1) in the direction of the linear eigenvectors.

Such paths are useful, particularly the heteroclinic paths, as they essentially provide a free transfer; the dynamics of the 3-body setting do all the work. As for their existence in the solar sail 3-body problem, the authors in a separate work have displayed some homoclinic paths, and indeed homoclinic invariant manifolds of periodic orbits where each trajectory in a periodic orbit’s invariant manifold is a homoclinic path. Heteroclinic paths are more difficult to come by, and we leave to a further work a systematic search for such features.

We use instead what we will call a ‘controlled heteroclinic transfer’, where the unstable manifold of one fixed point passes close enough to another fixed point for the controller to finish the transfer. Finding candidate unstable manifolds is not as easy as it would at first appear, since (as already mentioned) a true heteroclinic path is an asymptotic one with the benefit that the sail will naturally slow down as it approaches the target fixed point. For non-heteroclinic paths on the other hand, the sail will be moving quickly as it approaches the target, and may be moving too quickly for the controller to work.
Instead we look for an unstable invariant manifold which seems to be escaping from the Earth, and then falls back towards the Earth in the vicinity of the $\beta$ level curve on the far side; then the sail is moving slowly at this turning point and the controller has a better chance. We find that the $\gamma = 1.322$ fixed point has an unstable invariant manifold with this property, thus we first need to bring the sail into the vicinity of this point. This is simply done using the methods of Section II, and the control effort required is shown in Figure 9 (e). Then, the controller is switched off and the sail is allowed to fall away on the unstable manifold, thus $\gamma$ is fixed (Figure 9 (f)). When the sail is moving slowly in the vicinity of the $L_2$ side fixed points, the controller is switched on again to bring the sail onto the $\gamma = 0.9$ fixed point (Figure 9 (g)). The trajectories for this third leg are shown in Figure 8.

In general the control effort over the length of the whole mission is relatively small, that is $\gamma$ and $\phi$ are largely fixed for most of the transfer. One of the drawbacks in using a solar sail for the missions outlined above is the relatively long time scale, as is typical of solar sails, bearing in mind that $2\pi$ units in the non-dimensionalized system represents one year. This however is a small price to pay considering there is no propellant consumption aside from that required to insert the sail onto its initial trajectory.

V. Alternative routes to periodic orbits

In this final section we present some alternative methods to finding periodic orbits using control. The first uses a large value of $\rho$ as given in Section II, and the second is consists of using a single variable control scheme to assign eigenvalues based on the Center Manifold theorem.

As we described in Section II, an optimal controller minimizes a cost function according to arbitrary weighting matrices $Q$ and $N$ as in (8). Letting $Q = I_6$ and $N = \rho I_2$ reduces the freedom of the weights to one parameter. Typically, when we wish to control to a fixed point without considering the control effort we may set $\rho$ to be small, and this is the procedure we have followed in this work so far ($\rho = 10^{-4}$). This strongly effects the eigenvalues of the controlled linear system $A_c$. If on the other hand we were to choose a large value of $\rho$, the controller would seek to minimize the control effort and not the distance from equilibrium, and thus the sail would orbit the fixed point rather than fall onto it (at least for moderate time scales).

We see this most clearly in the form the eigenvalues of $A_c$ take for $\rho = 10^4$ for example. If the eigenvalues of $A$, the uncontrolled linear system, are

$$\left\{ \pm \lambda_1, \pm \lambda_2, \pm \lambda_3 \right\},$$

Figure 9. The control effort required for the mission described in the text. Each leg of the mission is labeled as follows: (a) transfer from Earth to $\gamma = 0.872$, (b) control onto $\gamma = 0.8092$, (c) control onto periodic orbit, (d) control back onto $\gamma = 0.8092$, (e) control onto $\gamma = 1.322$ fixed point, (f) transfer to $L_2$ side along unstable manifold, and (g) control onto $\gamma = 0.9$ fixed point.
then the eigenvalues of $A_c$ will be approximately
\[\{-\epsilon \pm \lambda_1 i, -\epsilon \pm \lambda_2 i, -\lambda_r, -\lambda_r\}\], (19)
with $\epsilon$ small. This means that each mode is asymptotically stable, however the real modes will decay and since the real part of the complex modes is so small the imaginary parts will dominate. Thus the sail will orbit the fixed point according to the frequencies $\lambda_1, \lambda_2$. As previously pointed out, if these frequencies are not in ratio the sail will follow a dense ‘rose’-type orbit, akin to the Lissajous orbits about the classical collinear points. If the frequencies are in ratio then the orbit will close forming a possible multiply-periodic orbit.

The advantage of this method is that the sail will orbit the fixed point for long times with a minimum of control effort. The sail may eventually fall onto the point as the system is asymptotically stable, although weakly so. The orbit the sail follows is sensitive to the initial data, and depending on this the sail will follow an orbit in one frequency of a mixture of the linear frequencies. Nonetheless, the ‘center’ manifold of this controlled system will be close to the center manifold of the uncontrolled system, so the orbits in it will be similar. In a way this suggests the disadvantage of the optimal controller, that is we cannot specify the eigenvalues of the controlled system directly, merely how strongly asymptotically stable they should be.

An alternative is to consider a single variable controller with which we may explicitly assign the eigenvalues of the controlled linear system. We will show how we may use this to describe periodic orbits in a direct way, but first the method.

Choose a parameter $u^*$ to vary, either $\gamma$ or $\phi$. Let $B = \partial f / \partial u^*$, and the controlled linear system is $A_c = A - BG$ where $G$ is $1 \times 6$ now. Choose six eigenvalues $\lambda_i$ and construct the desired characteristic polynomial $(\lambda - \lambda_1) \ldots (\lambda - \lambda_6)$. Form the characteristic polynomial of $A_c$ in terms of the components of $G$, and equate the coefficients of these two polynomials. Then solve uniquely for the components of the gain $G$.

The natural question which follows is, what should we choose the eigenvalues to be? We take some guidance from (19) above and the Center Manifold Theorem, which we can state in loose terms as the following:

**Center Manifold Theorem** If the eigenvalues of a dynamical system linearised about a fixed point are stable (negative real part) and imaginary (zero real part), the stability or otherwise of the fixed point is determined by the system reduced to the center manifold.

What this essentially means is that stable modes die out, thus trajectories are in a sense pushed onto the center manifold. Once there, the stability of the fixed point may be determined by using normal forms to reduce the system to the center manifold. The benefit is that the dimension of the center manifold may be much smaller than that of the full system, greatly simplifying the analysis. See Carr\textsuperscript{17} or Troger and Steindl\textsuperscript{18} for a discussion.

We however are not so much interested in the stability of the fixed point as we may freely specify this. Instead, the idea is that if we choose the linear eigenvalues to satisfy the theorem then we should see the same behavior, i.e. trajectories being pushed onto the center manifold, and remaining there. There are three distinct advantages to using this method:

1. Fixed points in the $x$-$y$-$z$ space do not have center eigenvalues at linear order, thus the Lindstedt-Poincaré method cannot be used directly to find approximations to periodic orbits. We note that one alternative is to perturb the dynamical system so that the linear system does contain centers, with the perturbation being picked up at higher order. This is just like the halo orbits about the classical collinear points, where the linear system is perturbed to ensure the in-plane and out-of-plane frequencies are the same. This however becomes quite cumbersome, for example when using a differential corrector, as the nice symmetry of the orbits about points in the $x$-$z$ plane is lost.

Using the methods described above however, we can simply assign center eigenvalues at linear order using an appropriately chosen gain. We give an example of this below.

2. Not only can we assign the magnitude of the stable modes, but we may also assign the frequency of the center modes. This may be useful for specific missions, for example the one-year orbits described in.\textsuperscript{5}
3. Assigning only one pair of center modes removes the difficulty of choosing initial data so as to fall onto the $H$- or $L$-type orbits, as there is only one family and it attracts all trajectories (in the vicinity of the fixed point).

For example, consider the fixed point

$$\begin{align*}
\beta &= 0.05, \quad \gamma_c = 0.5, \quad \phi_c = 0.095, \\
x_e &= 0.98713, \quad y_c = 0.009855, \quad z_c = 0.0143.
\end{align*}$$

(20)

Linearising (2) about this point we see the linear eigenvalues are

$$\left\{-0.0047 \pm 1.319i, 0.025 \pm 0.768i, -0.548, 0.508\right\},$$

(21)

that is, a stable spiral, an unstable spiral and a saddle (note the saddle modes are not of equal magnitude as the symmetry of fixed points in the $x$-$z$ plane is lost). As mentioned there are no center eigenvalues to build up periodic orbits using the LP method. Instead, we assign the eigenvalues of the controlled linear system to be

$$\left\{\pm \lambda_c i, -\lambda_1, -\lambda_2, -\lambda_3, -\lambda_4\right\},$$

(22)

with $\lambda_i$ real and positive. Then, the real stable modes will die out and the trajectory will fall onto the 2-dimensional center manifold. For example, for the fixed point above we will let the eigenvalues be

$$\left\{\pm 0.76i, -0.5, -1, -1.3, -1.5\right\}$$

(23)

which gives a gain of $(-115.3, 20.26, 5.53, -87.8, -18.4, -45.15)$. We find that the controller works best when we assign the center frequency to be close to the imaginary part of one of the linear system’s eigenvalues; which one you choose distinguishes between the two families of orbits (although for small amplitudes there is more freedom to let the frequency differ).

The disadvantage of this approach is that we are not specifying a nominal orbit to control onto; the initial data determines the amplitude of the orbit the controller brings the sail onto. Not surprisingly therefore, this approach is sensitive to initial conditions. With a little experimentation though, large amplitude periodic orbits may be found about fixed points in the $x$-$y$-$z$ space, see Figure 10.

**VI. Conclusions**

We have examined the control characteristics of solar sails in the Earth-Sun circular restricted 3-body problem (CR3BP). By firstly designing an optimal controller for fixed points, we have shown that the sail may be controlled onto a fixed point from quite large displacements in phase space. Then, using a simple discrete method, we have described the optimal control onto periodic orbits and have shown that again the controller is very robust under a number of simple tests. To more fully demonstrate the capabilities of the controlled solar sail we devised a sample mission involving transfer from the Earth and between fixed points; the analysis of this sample mission combined the use of invariant manifolds of fixed points with an optimal controller. Finally we outlined some novel techniques to finding periodic orbits using the controller in situations where the Lindstedt-Poincaré method breaks down. To conclude, the solar sail combined with an efficient controller can enable a whole variety of interesting and useful orbits in the CR3BP with minimum cost and effort.
Acknowledgments

This work was funded by grant EP/D003822/1 from the UK Engineering and Physical Sciences Research Council (EPSRC).

References


