JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 31, No. 5, September–October 2008

Classification of Two-Dimensional Fixed-Sun-Angle Solar Sail Trajectories

Stephen Wokes,* Phil Palmer,† and Mark Roberts‡
University of Surrey, Guildford, England GU2 7XH, United Kingdom

DOI: 10.2514/1.34466

The two-dimensional heliocentric trajectories for fixed-sun-angle solar sails are examined. The objective of this work is to completely classify all possible motion in this framework. This is achieved via a mathematical reduction that results in a two-dimensional phase-space interpretation of the problem. For a given sail quality (lightness factor) and sun angle, this phase space shows all possible solar sail trajectories. This phase space is used to survey the possible motion. It is discovered that there is a trajectory that moves from one logarithmic spiral solution to the other. Another trajectory that experiences being instantaneously stationary within the solar system is also found. How the phase space varies according to the sail quality and the sun angle is also investigated.

I. Introduction

THE last decade has seen some significant advances in the field of solar sailing. On the practical side, we have seen the first attempted launch of a sail[§] and the first in-space test of a deployment system specifically designed for a solar sail [1]. On the academic side, there have been numerous papers published in various areas such as construction design [2], attitude control [3,4], optimization of trajectories [5,6], and the development of novel and useful applications for solar sails [7,8]. This paper reexamines the fundamental dynamics of the 2-D, sun-centered, fixed-sun-angle solar sail. The current knowledge about the system can summarized as follows.

Tsu [9] ignored the radial solar radiation pressure (SRP) force and discovered that the system has some analytic solutions: the logarithmic spiral solutions. Bacon [10] showed that these solutions still exist when the radial SRP force is included, and London [11] showed that a solar sail with a fixed sun angle could move along these trajectories. London also calculated the times to move along these trajectories and hence the optimum sail orientation to minimize travel times. Kiefer [12] considered patching together logarithmic spirals and Kepler ellipses for solar observation missions. Wesseling [13] constructed a two-variable asymptotic solution about the (tightly bound) logarithmic spiral. Although the analysis was focused around the logarithmic spiral, Wesseling's work gave the first glimpses of the general motion. Van der Ha and Modi [14] also used a two-variable method to obtain asymptotic expressions for the orbital elements when the initial eccentricity was small. Vulpetti [15] found and explored the solar sail trajectories in which the orbital angular momentum about the sun is reversed. These trajectories are of relevance in this paper and appear within the classification of all 2-D fixed-sun-angle solar sail trajectories in Sec. V.

Recently, the focus of 2-D solar sail dynamics research has moved onto what can be achieved when the sun angle is varied (e.g., [16–18]). There has also been considerable work on the optimization of variable-sun-angle trajectories, either trying to achieve the desired mission as fast as possible [19] or using the lowest quality of sail possible [17]. Such research is out of the scope of this paper, because we only consider the fixed-sun-angle case. However, it may be interesting to reconsider the trajectories from such previous work within the framework presented in this paper.

II. System

We shall use a 2-D sun-centered inertial frame with the standard polar coordinates r and θ (the true longitude; that is, the angle measured from some inertial reference direction). The sail is modeled as a line, and the only forces acting are the sun's gravity and the SRP force. The sun is modeled as a point source and the sun angle α is constant. Solar sails are typically designed so that only one side of the solar sail faces the sun [20], and so the choice of α will be restricted to between -90 and $+90$ deg.

The magnitude of the radial and tangential components of the SRP force can be expressed as

$$F_{\text{SRP radial}} = \frac{m(\eta+1)\mu}{r^2} \tag{1}$$

$$F_{\text{SRP transverse}} = -\frac{m\eta\xi\mu}{r^2} \tag{2}$$

where μ is the gravitational parameter of the sun, r is the radial distance of the sail from the sun, m is the mass of the sail, and η and ξ can be expressed as [21]

$$\eta + 1 = \tfrac{\beta}{2}\cos\alpha(1-p_s) + \beta\cos^2\alpha[p_s\cos\alpha + \tfrac{1}{3}p_d] \tag{3}$$

$$-\eta\xi = \beta\cos\alpha\sin\alpha[p_s\cos\alpha + \tfrac{1}{3}p_d] \tag{4}$$

where p_s is the fraction of incoming photons specularly reflected, p_d is the fraction of incoming photons diffusely reflected, and β is the lightness factor of the sail. The lightness factor is related to the sail loading σ by

$$\beta = \frac{\tilde{\sigma}}{\sigma}$$

where $\tilde{\sigma}$ is the critical sail loading. If a sail has the critical sail loading and is perfectly specularly reflecting, then (if the sail is flat on) the SRP force perfectly cancels the gravitation force. For our sun, the critical sail loading has a value of 1.53 g m^{-2}.

It is assumed that the sail does not degrade, and therefore the constants on the right-hand sides of Eqs. (3) and (4) are all constant. The system can be parameterized by α, β, p_s, and p_d, but it is mathematically simpler to use η and ξ. Figure 1 shows the relationship between these two different forms of parametrization

Received 6 September 2007; revision received 29 November 2007; accepted for publication 1 December 2007.

*Ph.D. Student, Department of Mathematics; s.wokes@surrey.ac.uk. Student Member AIAA.

†Deputy Director, Surrey Space Centre; P.Palmer@surrey.ac.uk. Member AIAA.

‡Professor, Department of Mathematics; m.roberts@surrey.ac.uk.

§Data available online at http://www.planetary.org/programs/projects/solar_ sailing/ [retrieved 16 April 2008].

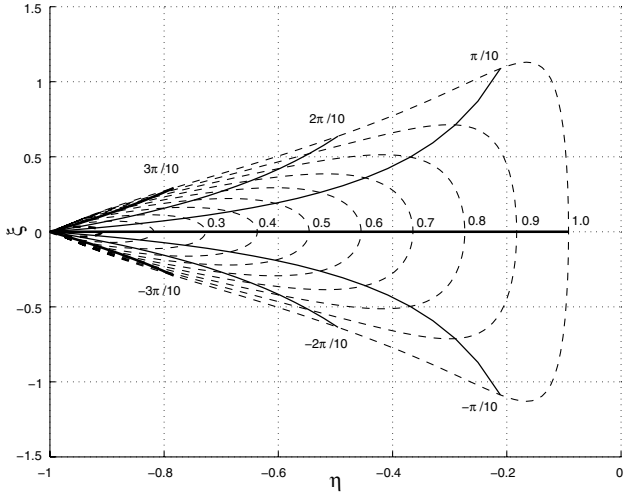

Fig. 1 Variation of η and ξ with respect to β and α for $p_s = 0.8272$ and $p_d = -0.0164$.

when $p_s = 0.8272$ and $p_d = -0.0164$ (these values are taken from Wie [3], in which they are stated to be representative of currently obtainable values). In Fig. 1, the dashed curves correspond to varying α with β as a specific constant (from 0.1 to 1.0). The solid curves correspond to varying β with α as a specific constant [from $-(3\pi/10)$ to $3\pi/10$], with the thick line along the $\eta = 0$ axis corresponding to $\alpha = 0$.

Considering the forces in the radial direction gives

$$\ddot{r} - \frac{h^2}{r^3} = \frac{\eta\mu}{r^2} \tag{5}$$

and considering the forces in the transverse direction gives

$$\dot{h} = -\frac{\eta\xi\mu}{r} \tag{6}$$

where $h = r^2\dot{\theta}$ is the specific orbital angular momentum, and the dots denote differentiation with respect to time. Equations (5) and (6) are the equations of motion for the 2-D fixed-sun-angle solar sail. This system is three-dimensional, with the variables being r, $\dot{r}$, and h. The angular momentum h is used to reduce the dimensions of the problem by one (using θ and $\dot{\theta}$, the system is four-dimensional).

A. Restrictions on η and ξ

It is clear that Eqs. (5) and (6) are unchanged if the sign of ξ and the sign of h are both reversed. The resultant trajectory is a mirror image of the original, with one moving clockwise and the other moving anticlockwise. Therefore, the dynamics when ξ is negative can be deduced from the case when ξ is positive. Hence, this paper restricts itself to the situations in which ξ is positive.

Further, we restrict attention to situations in which $\eta < 0$, which corresponds to the total radial force always being toward the sun. The dynamics of situations when $\eta \geq 0$ are beyond the scope of this paper.

B. Logarithmic Spirals

No general analytic solution or constants of motion to the preceding equations of motion have ever been found. However, the equations have been solved for a very specific set of initial conditions. These solutions were found by Bacon [10] (and later, London [11]) and are known as the logarithmic spiral solutions, taking the form

$$r(\theta) = r_0 e^{\theta \tan\chi}$$

where χ is the angle between the solar sail velocity vector and the transverse direction and is called the spiral angle of the trajectory. These trajectories can be traveled along in either direction: either spiraling outward and escaping the solar system or inward and falling into the sun.

The relationship between χ and α is given by

$$\frac{\sin\chi\cos\chi}{2 - \sin^2\chi} = \xi \tag{7}$$

which is from McInnes [21], but reexpressed in terms of ξ. This can be further simplified to

$$\tan\chi = \frac{4\xi}{1 \pm \sqrt{1 - 8\xi^2}} \tag{8}$$

III. Reduced System

Changing the independent variable from time to θ and using the substitution

$$v = \frac{h^2}{\mu r} \tag{9}$$

one can derive the following equation for v:

$$v'' = \frac{-\eta\xi(2\eta\xi + v')}{v} - \eta - v$$

If w is defined as

$$w = \frac{h\dot{r}}{\mu} \tag{10}$$

then by considering the differentiation of Eq. (9) with respect to θ, the equations of motion become

$$v' = -2\eta\xi - w \tag{11}$$

and

$$w' = -\frac{\eta\xi w}{v} + \eta + v \tag{12}$$

This system is two-dimensional, and hence this change of variables has reduced the original system [given by Eqs. (5) and (6)]. Dividing Eqs. (11) and (12) by η reveals that η just scales the variables v and w. It is ξ that dictates the dynamics of the v–w phase space. This shall be discussed fully in Sec. VI.

Observe that v can also be written as

$$v = \frac{h}{\mu} v_\theta$$

where $v_\theta = r\dot{\theta}$ is the transverse component of the velocity vector. Combined with the definition of w, this reveals that the v–w space is a hodograph, scaled by the factor h/μ (see Battin [22] for an introduction to hodographs). The factor is not constant (h varies), and so one cannot recover the magnitude of the sail's velocity from v and w. However, the flight angle (which is the same as the spiral angle, but need not be fixed) is given by the ratio w/v. In particular, this means that when $h > 0$, the sign of $\dot{r}$ is the same as the sign of w. However, if the angular momentum h is negative, the sign of $\dot{r}$ is opposite to the sign of w.

This draws our attention to the importance of the sign of h (or, equivalently, the sign of $\dot{\theta}$). To further demonstrate this, consider a point in v–w space. The equations of motion dictate how this point moves with increasing θ. However, the sign of h determines if θ is increasing or decreasing. The θ evolution of any nonsingular point (see Sec. IV.C) is unique, but the sign of h determines which direction is the future and which is the past.

In the next section, we shall start to investigate the dynamics of a general v–w phase space. However, let us first consider the simplest case: that of Kepler motion. Kepler motion corresponds to $\eta = -1$ and $\xi = 0$ and, in this case, the v–w plane is as shown in Fig. 2. The nullclines ($v' = 0$ and $w' = 0$) intersect at (1, 0), giving the only equilibrium point. The example trajectory shows the path of a spacecraft on an orbit with eccentricity 0.3. The v–w representation

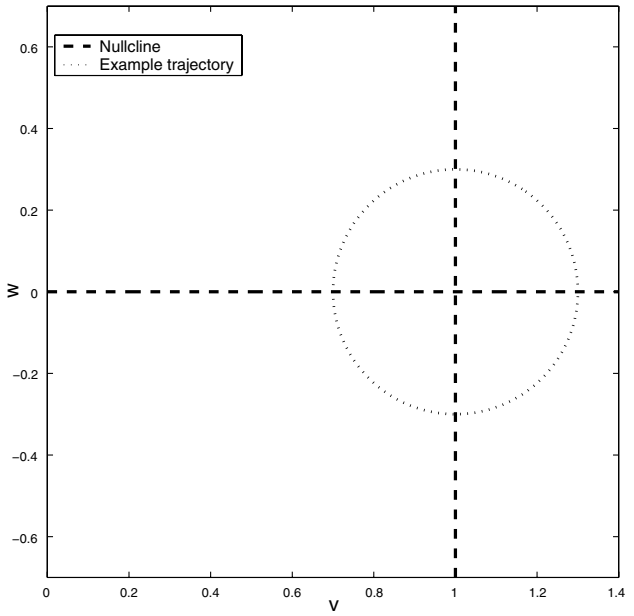

Fig. 2 A v–w plot when $\eta = -1$ and $\xi = 0$.

of any elliptic orbit with eccentricity e is a circle centered at (1, 0) with radius e. Hence, the fixed point (1, 0) corresponds to Keplerian circular motion. Contact with the $v = 0$ line in this v–w space corresponds to the spacecraft escaping the sun.

Regardless of η, when $\xi = 0$, the system is essentially Keplerian, with the sun's gravitational parameter altered to $-\eta\mu$. The v–w phase space of such systems is as shown in Fig. 2, except that the vertical nullcline is situated at $-\eta$ rather than 1.

IV. Structure of a v–w Space

To illustrate the structure of a v–w space, let us consider the situation when $\eta = -0.75$ and $\xi = 0.2$, which corresponds to a lightness factor of about 0.47. These values are chosen so that all the relevant features of the phase space can be seen on one figure (Fig. 3). This phase plot will be used repeatedly in this paper.

Two equilibrium points exist where the nullclines cross. These positions are located at $(\tilde{v}_1, -2\eta\xi)$ and $(\tilde{v}_2, -2\eta\xi)$ in the v–w space, where

$$\tilde{v}_1 = -\frac{\eta}{2}(1 - \sqrt{1 - 8\xi^2}) \qquad \tilde{v}_2 = -\frac{\eta}{2}(1 + \sqrt{1 - 8\xi^2})$$

The first equilibrium point at $\tilde{v}_1$ is always a saddle, and the second equilibrium at $\tilde{v}_2$ is always a source (with respect to increasing θ, not necessarily the same as increasing time), with the eigenvalues usually having an imaginary component (as is the case in Fig. 3). Section VI considers the case when this is no longer true.

The stable and unstable manifolds of the first equilibrium point cut the phase space into three distinct regions, which are labeled 1 to 3 in Fig. 3. A curve in the v–w space cannot move between the distinct regions, because trajectories are not allowed to intersect one another, except at singular points (see Sec. IV.C). The two seemingly disjoint parts of region 1 are connected through the origin (see Fig. 4), which will be explained in Sec. IV.C.

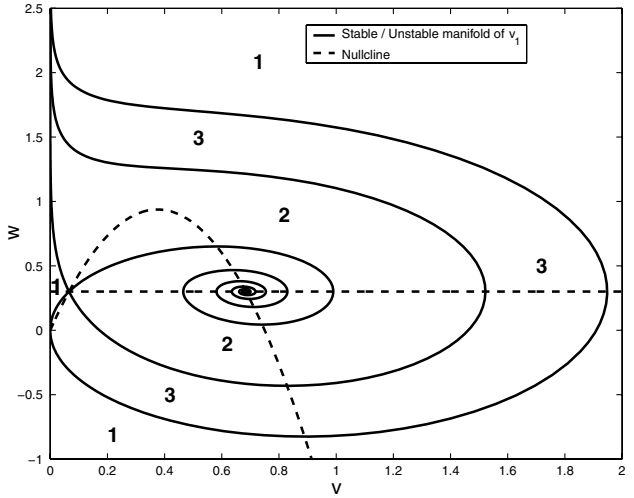

Fig. 3 A v–w plot when $\eta = -0.75$ and $\xi = 0.2$, with the distinct regions labeled 1 to 3.

A. Equilibrium Points

By considering $\mathrm{d}r/\mathrm{d}\theta$, one can find an analytic form of the motion corresponding to the equilibrium point at $\tilde{v}_i$: namely,

$$r = r_0 \exp\left(\frac{-2\eta\xi\theta}{\tilde{v}_i}\right)$$

Therefore, the equilibrium points of the v–w phase space correspond to the logarithmic spiral solutions. The preceding equation gives the spiral angle χ as

$$\tan\chi = \frac{-2\eta\xi}{\tilde{v}_i} = \frac{4\xi}{1 \pm \sqrt{1 - 8\xi^2}}$$

which is the same as Eq. (8). This result is far quicker to derive using the v–w framework. This equation shows that logarithmic spiral solutions are no longer possible when $\xi > 1/(2\sqrt{2})$.

An interesting discovery using the v–w space is that there is a heteroclinic orbit linking the two logarithmic spiral solutions.

B. Flow of the v–w System

The flow of the v–w system is dictated by the nullclines. Figure 5 shows the nullclines of the phase space (dashed lines, as before), and the direction of the flow is given by the black arrows. When a point is above/below the horizontal nullcline, v' is negative/positive, and when it is above/below the quadratic nullcline, w' is positive/negative. The shape of the flow suggests that there is a source at both the origin and at the second equilibrium point $(\tilde{v}_2, -2\eta\xi)$, whereas there is a sink at $(0, +\infty)$. This will have relevance in Sec. IV.C.

The reader is again reminded that the preceding results about the flow of the system are with respect to increasing θ, which may not be the same as increasing time.

C. Singular Points

The original equations of motion (5) and (6) have a singularity at $r = 0$, and the transformation to v–w space is the zero map if $\dot{\theta} = 0$.

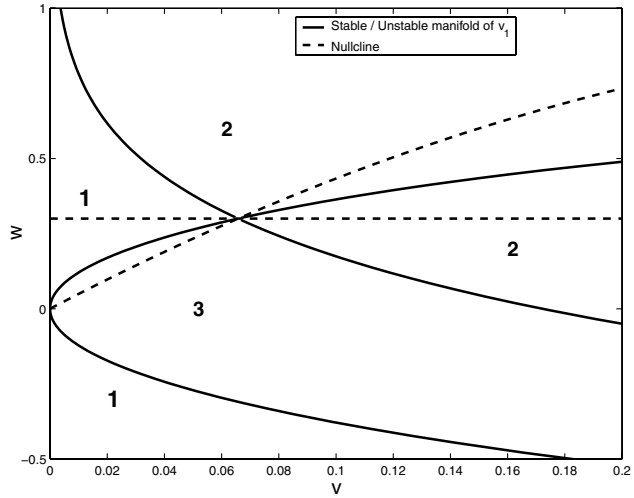

Fig. 4 Close-up of Fig. 3 about the origin.

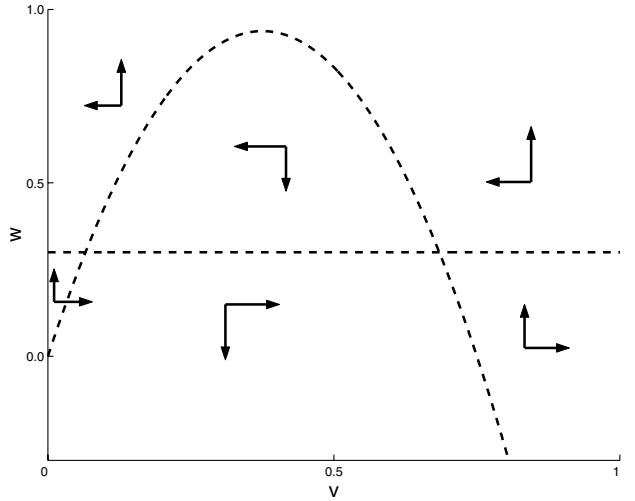

Fig. 5 Nullclines (dashed lines) and direction of the flow (shown by arrows) for the v–w space when $\eta = -0.75$ and $\xi = 0.2$.

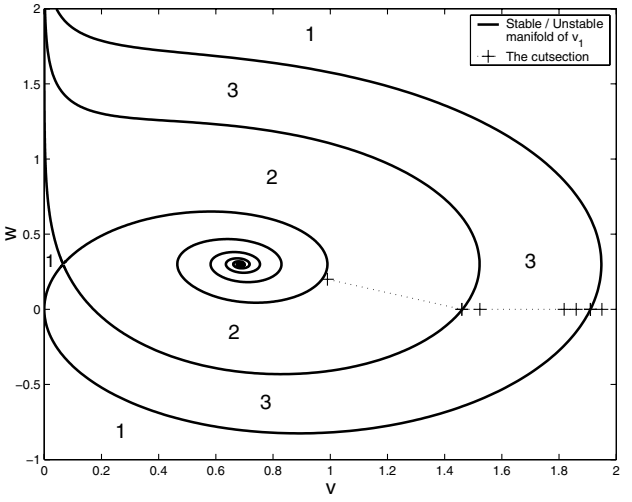

Fig. 6 Cut section used to classify the possible motion, with distinct regions labeled 1 to 3.

Therefore, it should not be surprising that the v–w equations of motion (11) and (12) are singular when $v = 0$.

When $\xi = 0$, any point on the $v = 0$ axis is an attainable singularity and corresponds to the spacecraft escaping the solar system ($r \to \infty$). However, when $\xi \neq 0$, it can be shown that there are only two attainable singular points. The first is at the origin, labeled SP_0, and the second is at $(0, +\infty)$, labeled SP_∞.

Both singular points correspond to $\dot{\theta} = 0$, but at SP_0, this occurs with a finite value of r, whereas at SP_∞, this happens with $r = \infty$. Therefore, SP_0 corresponds to the angular momentum changing sign [Eq. (6) insists this happens once, at most] and SP_∞ corresponds to solar system escape.

Consideration of the equations of motion shows that the singularity SP_∞ is attained with a finite value of θ. This means that such solar system escape trajectories escape with a finite limit to their true longitude. This is distinctly different from logarithmic spiral solutions, which have no such limit to their true longitude.

It can be shown that the singularity at SP_0 is attained within a finite value of time as well as θ. Because the original equations of motion (5) and (6) are still valid at this singularity, a curve that passes through SP_0 must be at least continuous. By considering the dominant terms in the equations of motion, one can show that a v–w curve moves through SP_0 in a quadratic way (e.g., $v = Kw^2$), with K as a constant. (There is an single important exception to this, which will be discussed in Sec. V.B.) The same technique shows that curves approach SP_∞ on logarithmic paths of the form $w = -K \log(v)$.

The quadratic nature of all curves as they pass through SP_0 is what allows paths to move between the seemingly disjointed parts of region 1, because the stable manifold of $\tilde{v}_1$ must also be quadratic through this point. In contrast, paths in region 3 that pass through SP_0 return back to region 3.

V. Specific Classification of Possible Trajectories

In this section, we relate trajectories in the v–w space to their corresponding trajectories in real space. By doing this, we construct a complete classification of all possible trajectories when $\eta = -0.75$ and $\xi = 0.2$. Later, in Sec. VI, we show how the classification can be generalized for different values of η and ξ.

Let us suppose that we are given a starting point in the v–w phase space and that the angular momentum is positive. We shall see that the qualitative nature of the future and past history of the resulting trajectory depend on whether the starting point it is regions 1, 2, or 3. To show this, we shall use the cut section shown in Fig. 6, with the marks along the cut section denoting the starting places. (Although not visible in the Fig. 6, there are two marks where the cut section crosses from region 1 to region 3, each being very close the manifold and on opposite sides of it.) The corresponding plots are shown in Figs. 7–18. Each figure shows either a v–w plot or a real-space plot, with axes x and y denoting Cartesian coordinates.

The real-space plots are scaled relative to the initial starting radius. This is done because the whole trajectory scales up with the choice of the initial radius. This is the same in the 2-D Kepler case; the curve is not defined by the eccentricity and argument of perigee alone: a notion of distance (i.e., the semimajor axis) is needed.

In the following subsections, v_0 and w_0 denote the initial values of v and w for the graphs plotted.

A. Region-1 Trajectories

Figures 7–9 show the two sample trajectories from the cut section that are within region 1. The forward-time evolution of any trajectory in region 1 leads to solar system escape with positive angular momentum (attains SP_∞). Its backward-time evolution passes through SP_0, thus changing the sign of its angular momentum, then escapes the solar system with negative angular momentum (again attaining SP_∞). During the craft's entire history, the sign of $\dot{r}$ changes only once, and this cannot occur at the same time as when its angular momentum transits through zero.

These trajectories are the generalization of the hyperbolic paths in Kepler motion. In the v–w space, trajectories far from the origin and away from $v = 0$ are nearly circular about the origin, as they would be in the Kepler case. However, when approaching $v = 0$, the v–w paths are forced toward SP_∞ (via SP_0 if $v = 0$ is approached with negative w).

The first trajectory in region 1 shows what looks like (in the real-space plot) a standard hyperbolic trajectory. The second trajectory in region 1 shows the effect of moving v_0 and w_0 closer to the manifold dividing regions 1 and 3. (They are so close to the manifold that in Fig. 7 that the trajectory is only visibly distinct from the manifold

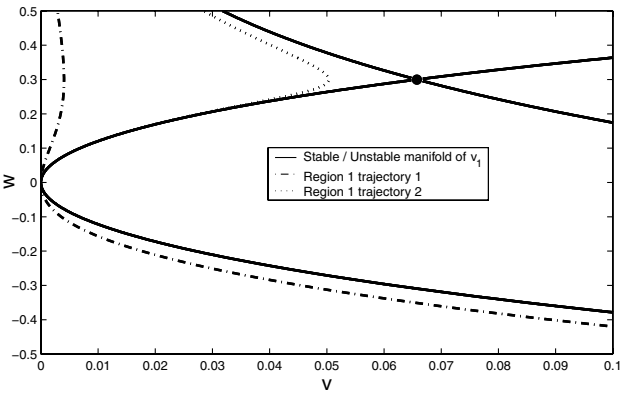

Fig. 7 A v–w plot of trajectories 1 and 2 in area 1.

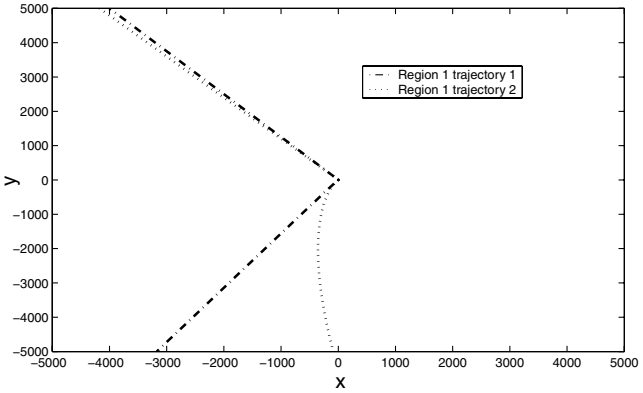

Fig. 8 Real-space plot of trajectories 1 and 2 in area 1.

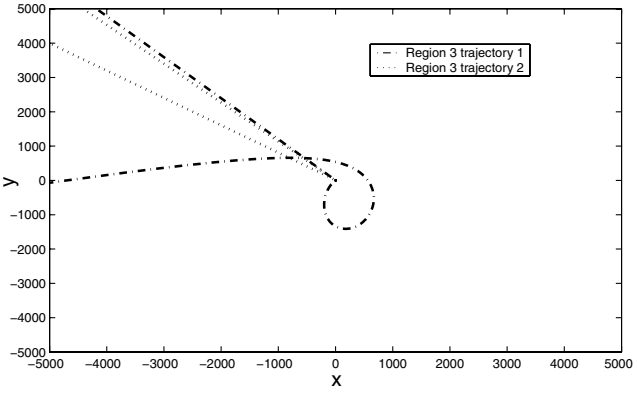

Fig. 11 Real-space plot of trajectories 1 and 2 in area 3.

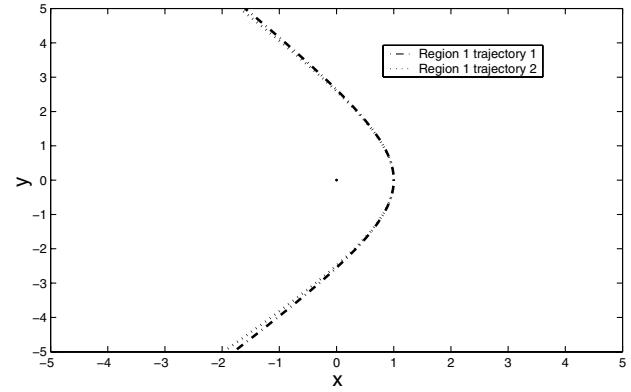

Fig. 9 Close-up of the real-space plot of trajectories 1 and 2 in area 1.

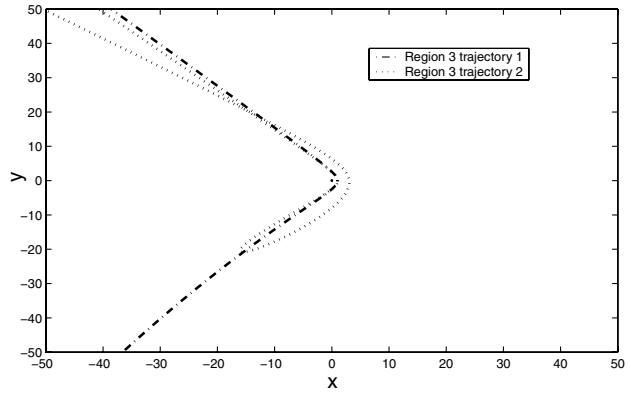

Fig. 12 Close-up of the real-space plot of trajectories 1 and 2 in area 3.

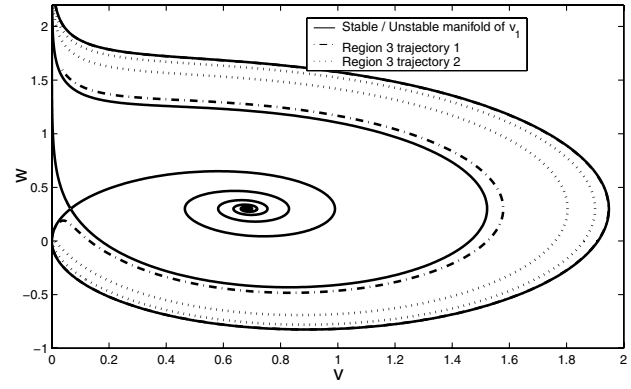

Fig. 10 A *v–w* plot of trajectories 1 and 2 in area 3.

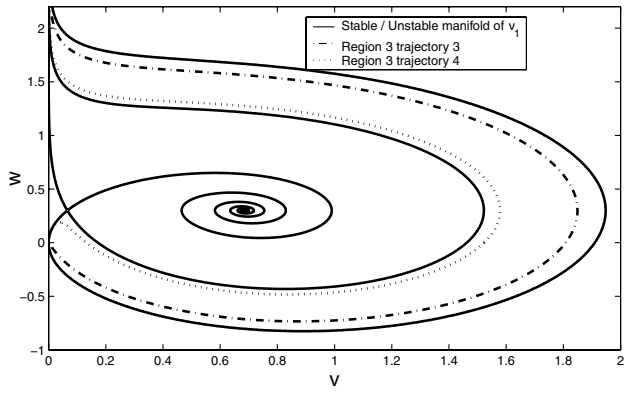

Fig. 13 A *v–w* plot of trajectories 3 and 4 in area 3.

when v_1 is approached.) The effect of the v–w representation getting close to the equilibrium point $\tilde{v}_1$ in Fig. 7 can be seen in the spiral shape of the trajectory in Fig. 8. In addition, the change in the sign of the angular momentum is more evident for the second trajectory.

B. Region-3 Trajectories

Figures 10–15 show the sample trajectories from the cut section that are within region 3. The forward-time evolution of any trajectory in region 3 leads to solar system escape with positive angular momentum, and its backward-time evolution leads to solar system escape with negative angular momentum (just as for points in region 1). However, during the entire trajectory, the sign of $\dot{r}$ changes three times. This is because the unstable manifold of $\tilde{v}_1$ crosses the $w = 0$ axis, thus forcing all region-3 trajectories to "dip under" the $w = 0$ axis.

These trajectories look extremely non-Keplerian, coming in from infinity, tracing a loop that does not include the sun, and then escaping off to infinity. These trajectories are the generalization of the angular momentum reversal trajectories discovered by Vulpetti [15].

The first trajectory from region 3 has v_0 and w_0 just inside the manifold separating regions 1 and 3. (They are so close to the manifold that in Fig. 10 that it is impossible to see the forward evolution.) In Fig. 11, its backward propagation looks as though it might escape the solar system, as it did for the second trajectory in region 1. But instead, it falls back on itself and swings past the sun with negative angular momentum. The same thing happens to the second trajectory in region 3, although at a smaller (relative) distance from the sun. In both of these figures, the backward propagation goes through zero angular momentum before it falls in on itself (i.e., $\dot{r} = 0$). However, the time between these events is noticeably smaller in the second trajectory.

The third trajectory in region 3 is a special unique trajectory in which $\dot{r} = 0$ occurs at precisely the same time as $h = 0$. The forward- and backward-time evolutions from this special point trace out

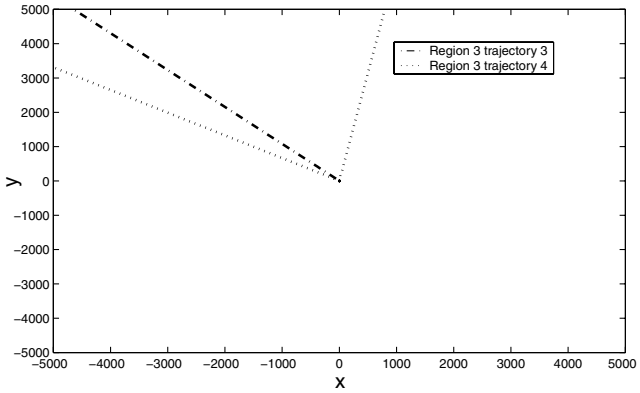

Fig. 14 Real-space plot of trajectories 3 and 4 in area 3.

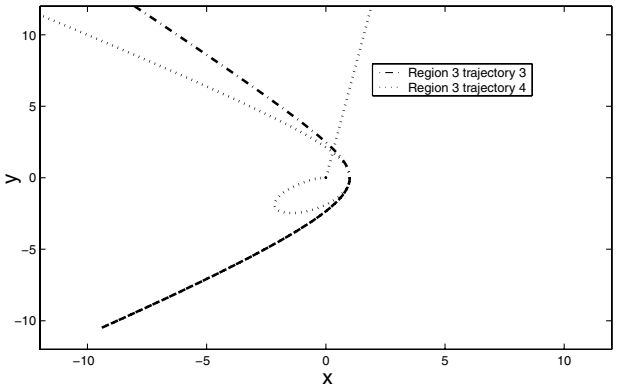

Fig. 15 Close-up of the real-space plot of trajectories 3 and 4 in area 3.

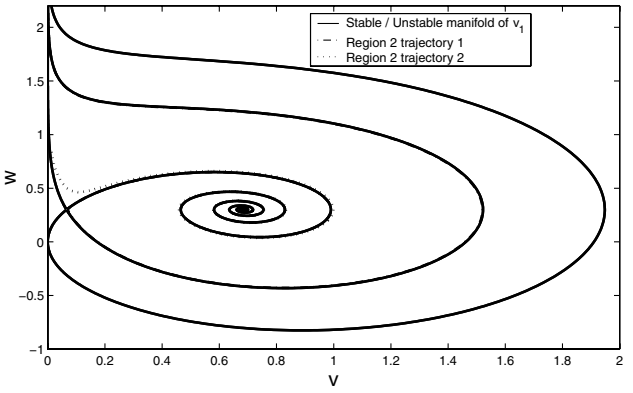

Fig. 16 A v–w plot of trajectories 1 and 2 in area 2.

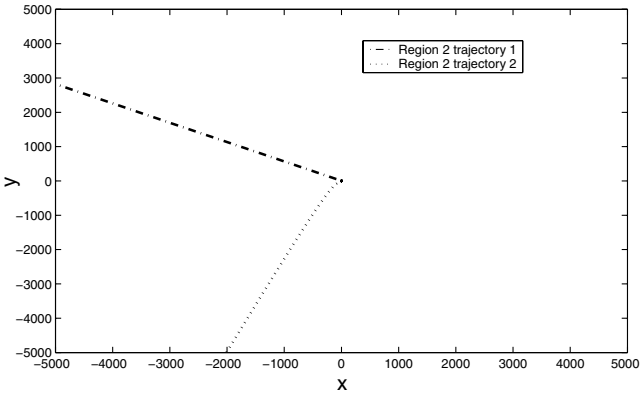

Fig. 17 Real-space plot of trajectories 1 and 2 in area 2.

exactly the same path in both real space and in v–w space. The v–w curve of this trajectory is the only curve that approaches SP_0 along a nonquadratic path. Instead, it approaches along a linear path, attains SP_0, and then bounces back along the path it came in on.

All v–w trajectories in region 3 wrap around this special trajectory, and hence all the curves beneath it are actually the same as the curves above it. This can be seen in the figures displayed. The first trajectory and fourth trajectory are actually the same trajectory in v–w space. All that has changed is the craft's starting point along the trajectory.

C. Region-2 Trajectories

Figures 16–18 show the sample trajectories from the cut section that are within region 2. Region 2 is different from the other two regions, because only the forward-time evolution leads to solar system escape. The backward-time evolution leads to the sail falling into the sun. This difference in behavior is because curves within region 2 of a v–w plot do not pass through SP_0 and thus their angular momentum cannot change sign.

During the entire history of the trajectory, the sign of $\dot{r}$ changes either twice or not at all. The reason for this is that the heteroclinic path between the two equilibrium points does not dip under the $w = 0$ axis. Trajectories in region 2 are the familiar spiral-like trajectories that are commonly associated with solar sails.

The first trajectory in region 2 has v_0 and w_0 just inside the manifold that separates regions 2 and 3. (They are so close that in Fig. 16 that it is impossible to see this trajectory.) In Fig. 17, the trajectory appears to be trying to fall in and change the sign of its angular momentum, as did the fourth trajectory of region 3. However, it never manages to transit through $h = 0$, and instead falls into the sun. The same thing happens with the second trajectory in region 2.

Falling into the sun corresponds to the crafts v–w representation spiraling into $\tilde{v}_2$. Thus, this terminal trajectory approximates the more tightly bound logarithmic spiral solution. This is clearly seen in the second trajectory, but in the first trajectory, we instead see motion corresponding to the more loosely bound logarithmic spiral of $\tilde{v}_1$. This is because the v–w representation of the first trajectory spends a long time near the first equilibrium point before finally falling in toward the second equilibrium point.

VI. Varying η and ξ

In Sec. IV, we used the parameter values $\eta = -0.75$ and $\xi = 0.2$ to illustrate various aspects of a general v–w phase space. In Sec. V, we classified all possible motion when $\eta = -0.75$ and $\xi = 0.2$. In this section, we examine how much of this changes when η and ξ vary, although we still impose the restrictions set out in Sec. II.A.

As mentioned in Sec. III, it is ξ that determines the structure of the v–w phase space; η is just a scaling factor. This can be seen if the v–w equations of motion (11) and (12) are rewritten as

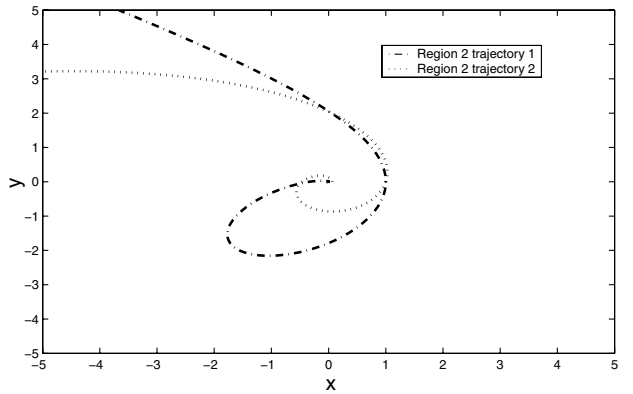

Fig. 18 Close-up of the real-space plot of trajectories 1 and 2 in area 2.

$$\frac{v'}{-\eta} = 2\xi - \frac{w}{-\eta}$$

and

$$\frac{w'}{-\eta} = \frac{\xi w}{v} - 1 + \frac{v}{-\eta}$$

Therefore, two phase spaces with the same value of ξ but different values of η are essentially the same phase space. That is, there is an obvious bijection between trajectories in either phase space (just scale the trajectory by the ratios of η). These equivalent paths are traversed at the same rate with respect to θ, but not with respect to time.

Now let us compare the real-space trajectories of equivalent paths. The following result is a direct consequence of the definitions of v and w:

$$\ell_n\left(\frac{r}{r_0}\right) = \int_{\theta_0}^{\theta} \frac{w}{v}\, \mathrm{d}\theta \tag{13}$$

This determines the change in r when moving along a path in v-w space, with r_0 and θ_0 denoting the initial values of r and θ, respectively. This result shows that equivalent paths from two different phase spaces correspond to exactly the same trajectory in real space.

Therefore, the classification of the trajectories in Sec. V is valid for any value of η, as long as $\xi = 0.2$. The v–w paths will be scaled by $-\eta$, but the corresponding real-space trajectory will be unchanged. It can be shown that the time to move around on these trajectories is scaled by $1/\sqrt{-\eta}$.

We now consider the effects of changing ξ. The results from Sec. IV are valid as long as the corresponding objects still exist (e.g., the notes about equilibrium points are valid, as long as equilibrium points still exist). As ξ changes, the shape of the figures change in a continuous and smooth way. However, there are distinct values of ξ in which the nature of the phase space changes significantly. The following subsections investigate each of these transitions.

A. Transition 1: Kepler Motion

The first transition corresponds to $\xi = 0$ (i.e., no transverse force). Such situations have to be treated completely differently from those in which $\xi \neq 0$. Assuming that $\eta \neq 0$, these systems are essentially Keplerian with an altered gravitational parameter of $-\eta\mu$. The angular momentum is conserved, there are no logarithmic spiral solutions, and the only equilibrium point (at $v = -\eta\mu$ and $w = 0$) corresponds to circular motion. Any point on the $v = 0$ axis is an attainable singularity and corresponds to solar system escape. The phase space when $\eta = -1$ and $\xi = 0$ is shown in Fig. 2.

Figure 19 shows the v–w space when $\eta = -0.95$ and $\xi = 1/19$ (which is 0.05263, calculated to four significant figures). In these figures and in all subsequent v–w figures in this section, the special trajectory of region 3 (the third trajectory from region 3 in Sec. V) and numerous example trajectories are plotted to better show the flow of the phase space.

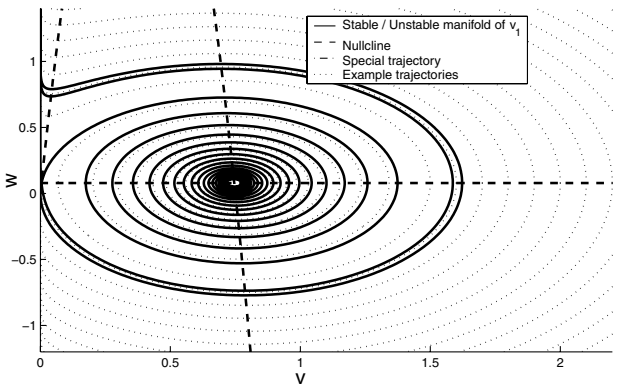

Fig. 19 A v–w plot when $\eta = -0.75$ and $\xi = 1/19$.

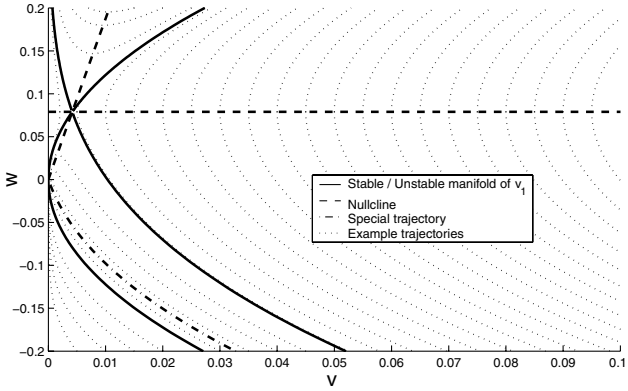

Fig. 20 Close-up of Fig. 19 about the origin.

The effects of $\tilde{v}_2$ dominate Fig. 19, with $\tilde{v}_1$ only noticeable in a close-up of the figure near the origin (Fig. 20). Despite looking distinctly different from the v–w plots in which $\eta = -0.75$ and $\xi = 0.2$, the general nature is the same. Figure 19 is divided into three distinct regions: region 1 outside all the manifolds, region 3 in the thin annulus containing the special curve, and region 2 in the remaining center area. There are two logarithmic spiral solutions, with a heteroclinic path joining the two, and the two singularities SP_0 and SP_∞ still exist and have the same nature and interpretation.

Real-space trajectories of curves in regions 1 and 3 still look as they did when $\xi = 0.2$; however, this is not true for region 2. The sign of $\dot{r}$ undergoes 24 or 26 changes during the entire history of any trajectory in region 2. This is distinctly different from the case examined in Sec. V, and so we shall explore this further.

Paths in v–w space that are within region 2 originate from $\tilde{v}_2$, the second equilibrium point. The number of times that $\dot{r}$ changes sign depends upon the number of times that the heteroclinic path joining the two equilibrium points dips under the $w = 0$ axis (it always starts above the axis). If the heteroclinic paths dips under n times, then the sign of $\dot{r}$ changes $2n$ or $2n + 2$ times during the history of a path in region 2. For $\xi = 0.2$, $n = 0$, and so the region-2 paths undergo 0 or 2 changes in the sign of $\dot{r}$. For $\xi = 1/19$, n is 12, and so region-2 paths undergo 24 or 26 changes in the sign of $\dot{r}$.

For small values of ξ, the integer n is large as the second equilibrium point gets close to the $w = 0$ axis and motion (within the v–w space) around $\tilde{v}_2$ becomes near circular. As ξ gets bigger, $\tilde{v}_2$ moves away from the $w = 0$ axis and n becomes small, eventually becoming zero. As ξ continues to increase, the second transition of ξ is eventually encountered (discussed subsequently), beyond which the sign of $\dot{r}$ never changes for all region-2 paths.

B. Transition 2: Sign of $\dot{r}$ Changing Three Times

One of the key features of region-3 trajectories was that throughout their history, $\dot{r}$ undergoes three changes of sign (as opposed to just the one change for region-1 trajectories). However, this is only the case if the unstable manifold leaving $\tilde{v}_1$ goes under the $w = 0$ axis. The second transition value of ξ is when this stops happening (or, equivalently, when the unstable manifold of $\tilde{v}_1$ just kisses the $w = 0$ axis). This second transition occurs when ξ is approximately 0.3014 (this number is found numerically and is calculated to four significant figures). To obtain such a value of ξ requires a lightness factor of over 0.6, and so all near-term missions will be beneath the second transition of ξ.

Figures 21 and 22 show two region-3 curves in the v–w phase space when $\eta = -0.75$ and $\xi = 0.35$. The v–w representation of the first curve dips under the $w = 0$ axis a second time, and hence the trajectory experiences three change in the sign of $\dot{r}$. In contrast, the v–w representation of the second curve does not dip under the $w = 0$ axis a second time, and hence the trajectory only experiences one change in the sign of $\dot{r}$.

When ξ is greater than the second transition but less than the fourth (see Sec. VI.D), there is still a distinct difference between v–w curves

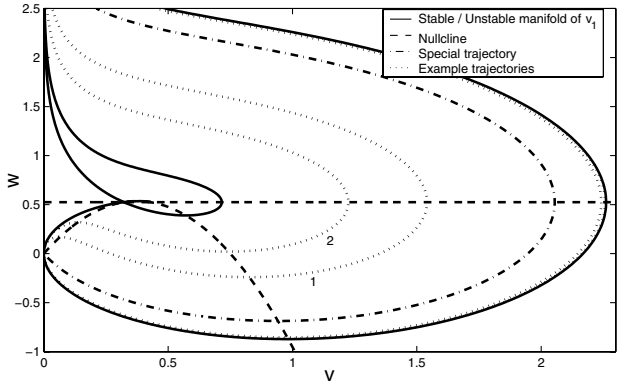

Fig. 21 A v–w plot when $\eta = -0.75$ and $\xi = 0.35$, with two example trajectories labeled 1 and 2.

in regions 1 and 3. As the curves leave SP_0 and head off to SP_∞, region-1 curves are bounded to the left of region 2. In contrast, region-3 curves must go around and approach SP_∞ to the right of region 2. Because the ratio w/v gives the flight angle, these restrictions to the v–w path tell us about the profile of the flight angle. For region-1 curves, the flight angle swings away from purely radial only slightly and then tends back toward radial. For region-3 trajectories, the flight angle moves away from radial, swings through the spiral angles of both logarithmic spirals, peaks, and then swings back down to purely radial. The stable manifold acts as a boundary between the two, joining SP_0 to $\tilde{v}_1$, for which the trajectory ends up on a logarithmic spiral.

C. Transition 3: Eigenvalues of $\tilde{v}_2$ Real

The eigenvalues of $\tilde{v}_2$ always have positive real parts, but for most (currently attainable) values of ξ, the imaginary parts are nonzero. However, when $\xi \geq 6/17$ (which is 0.3529, calculated to four significant figures), the eigenvalues become purely real. Figures 23 and 24 show the v–w space just before the third transition, and Figs. 25 and 26 show the v–w phase space at the transition.

Comparing Figs. 24 and 26, one can clearly see the changes associated with the third transition. The linear nature of the heteroclinic path approaching $\tilde{v}_2$ can be clearly seen in Fig. 26. After the third transition, points near $\tilde{v}_2$ move directly away from the equilibrium point, rather than spiraling outward. This means that the flight angle no longer oscillates about the spiral angle corresponding to $\tilde{v}_2$.

Comparing Figs. 23 and 25, one can see that the effects associated with the third transition are only noticeable near $\tilde{v}_2$. The rest of the phase space shows very little change.

D. Transition 4: Equilibrium Points Disappear

So far, this paper has only considered situations in which two distinct equilibrium points exist. When $\xi = 1/2\sqrt{2}$ (which is 0.3536,

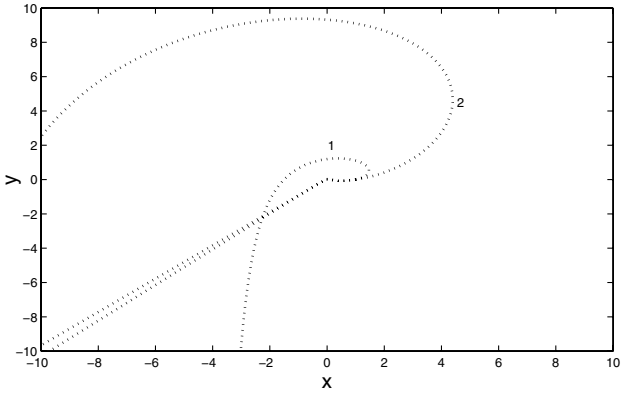

Fig. 22 Real-space plot of the example trajectories from Fig. 21.

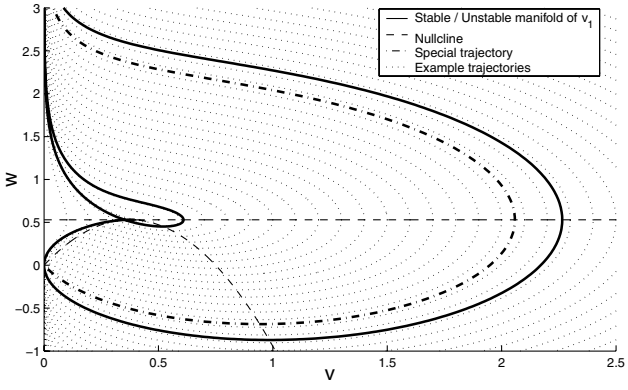

Fig. 25 A v–w plot when $\eta = -0.75$ and $\xi = 6/17$.

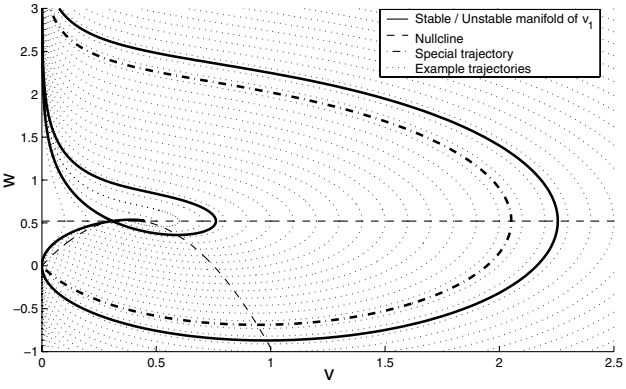

Fig. 23 A v–w plot when $\eta = -0.75$ and ξ is slightly beneath $6/17$.

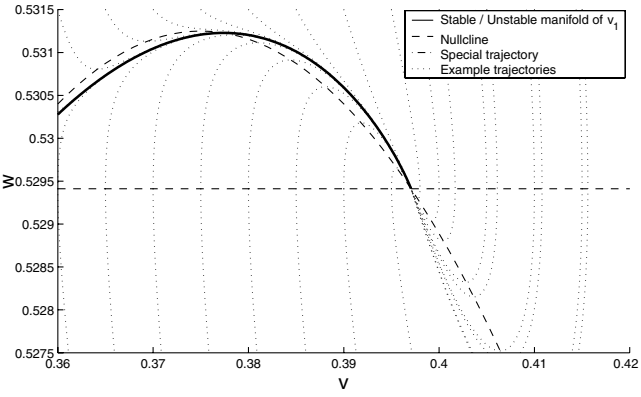

Fig. 26 Close-up of Fig. 25 about $\tilde{v}_2$.

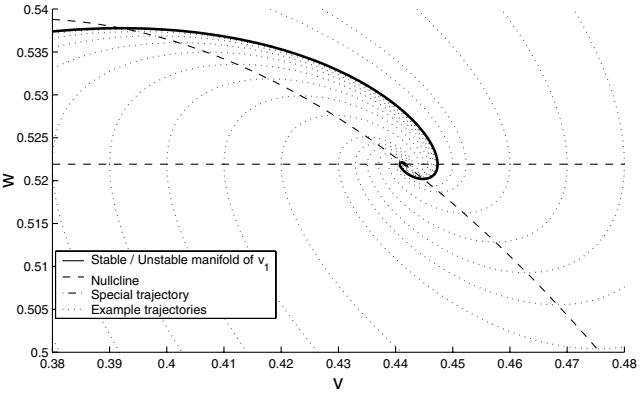

Fig. 24 Close-up of Fig. 23 about $\tilde{v}_2$.

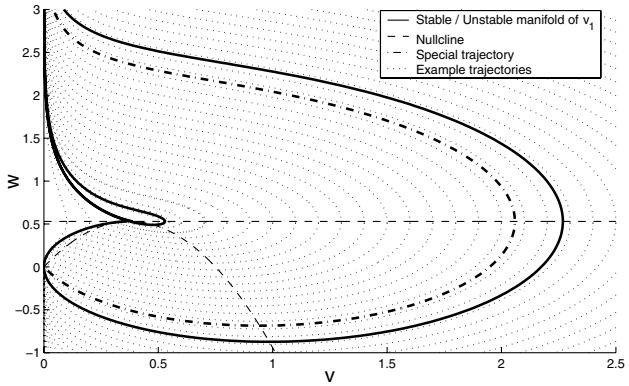

Fig. 27 A v–w plot when $\eta = -0.75$ and $\xi = 1/2\sqrt{2}$.

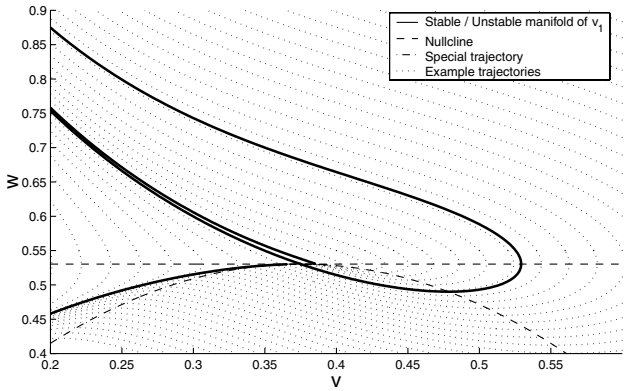

Fig. 28 Close-up of Fig. 27 about $\tilde{v}_2$.

calculated to four significant figures), the two equilibrium points merge into one; when ξ is greater than this limit, then no equilibrium points exist at all.

Figures 27 and 28 shows the v–w phase space at this transition. The single equilibrium point has one positive eigenvalue and one zero eigenvalue. The nature of the motion along the zero eigenaxis is that of a shunt, with trajectories coming in from $v < \tilde{v}$ and going out toward $v > \tilde{v}$. This means that the equilibrium point is only approachable along one path, whereas it could be approached on two different paths before.

It is perhaps surprising that region 2 still exists at the point at which the equilibrium points merge. We can see from Fig. 27 that the general shape of the rest of the phase space has not changed significantly. The flow moves out from the source at SP_0 and moves toward the sink at SP_∞.

If ξ is greater than $1/2\sqrt{2}$, then no equilibrium points exist, and the stable and unstable manifolds that separated the distinct regions of the phase space vanish. Figure 29 shows the phase space in such a situation, with $\eta = -0.75$ and $\xi = 0.4$. Of most importance is the disappearance of region 2. With the equilibrium points no longer existing, the backward θ propagation of any point in the v–w space passes through SP_0. Therefore, in this situation, all trajectories change their angular momentum and both forward and backward propagation leads to solar system escape (SP_∞).

The special curve still exists, because this only relied upon the nature of the system about the singular point SP_0. We can see from Fig. 29 that, as before, nearby paths wrap around this special curve. The fact that this curve's existence is totally invariant with respect to η and ξ further demonstrates its importance.

There are no more transitions in the phase space as ξ continues to increase beyond the fourth transition. Instead, the phase space continues to change in a smooth way. This paper does not investigate these changes any further. To justify this choice, recall from Fig. 1 that to attain the fourth transition of ξ requires a lightness factor of about 0.7, which is substantially beyond the current capabilities of solar sails.

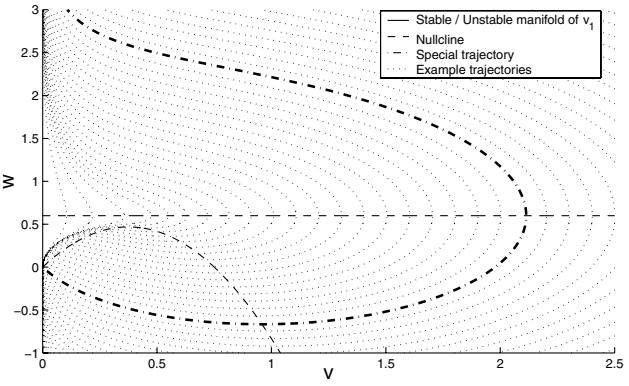

Fig. 29 A v–w plot when $\eta = -0.75$ and $\xi = 0.4$.

VII. Conclusions

This paper used a new phase-space approach to examine and better understand all two-dimensional trajectories for heliocentric solar sails with a fixed-sun-angle. As well as giving insight into the dynamics of the system, the phase space naturally partitioned the possible motion into three groups: pseudohyperbolic trajectories, angular momentum reversal trajectories, and spiral trajectories. Using the phase space, one can clearly see the connection between the different groups and the important boundaries that separate them. By showing how the phase space changes with the quality of the sail and sun angle, this paper has shown how the phase space could be used to understand, and perhaps to help design, variable-sun-angle solar sail trajectories.

References

[1] Nishimura, Y., Tsuda, Y., Mori, O., and Kawaguchi, J., "The Deployment Experiment Of Solar Sail With A Sounding Rocket," 55th International Astronautical Congress, Vancouver, Canada, Vol. International Astronautical Congress Paper 04-A.5.10, Oct. 2004.

[2] Lichodziejewski, D., Derbes, B., West, J., Rienert, R., Belvin, K., and Pappa, R., "Bringing an Eective Solar Sail Design to TRL6," 39th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, Huntsville, AL, Vol. AIAA Paper 2003-4659, July 2003.

[3] Wie, B., "Dynamic Modeling and Attitude Control of Solar Sail Spacecraft —Part I," AIAA Guidance, Navigation, and Control Conference and Exhibit, Monterey, CA, AIAA Paper 2002-4572, Aug. 2002.

[4] Murphy, D., and Wie, B., "Robust Attitude Control Systems Design for Solar Sails (Part 1): Propellantless Primary ACS," AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, RI, AIAA Paper 2004-5010, Aug. 2004.

[5] Hughes, G. W., and McInnes, C. R., "Solar-Sail Hybrid Trajectory Optimisation for Non-Keplerian Orbit Transfer," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 3, 2002, pp. 602–604.

[6] Dachwald, B., "Optimisation of Interplanetary Solar Sailcraft Trajectories Using Evolutionary Neurocontrol," *Journal of Guidance, Control, and Dynamics*, Vol. 27, No. 1, Jan.–Feb. 2004, pp. 66–72. doi:10.2514/1.9286

[7] McInnes, C. R., "Solar Sail Halo Orbits, Part 1: Heliocentric Case," *Journal of Spacecraft and Rockets*, Vol. 29, No. 4, July–Aug. 1992, pp. 466–479.

[8] McInnes, C. R., "Deflection of Near-Earth Asteroids by Kinetic Energy Impacts from Retrograde Orbits," *Planetary and Space Science*, Vol. 52, No. 7, July 2004, pp. 587–590. doi:10.1016/j.pss.2003.12.010

[9] Tsu, T. C., "Interplanetary Travel by Solar Sail," *Journal of the American Rocket Society*, Vol. 29, 1959, pp. 422–427.

[10] Bacon, R. H., "Logarithmic Spiral: An Ideal Trajectory for an Interplanetary Vehicle with Engines of Low Sustained Thrust," *American Journal of Physics*, Vol. 27, No. 3, 1959, pp. 164–165. doi:10.1119/1.1934788

[11] London, H. S., "Some Exact Solutions of the Equations of Motion of a Solar Sail With a Constant Setting," *Journal of the American Rocket Society*, Vol. 30, 1960, pp. 198–200.

[12] Kiefer, J. W., "Feasibility Considerations for a Sail-Powered Multi-Mission Probe," Proceedings of the 15th International Astronautical Congress, Vol. 1, International Astronautical Federation, Paris, 1965, pp. 383–416.

[13] Wesseling, P., "A Two-Variable Asymptotic Solution for Three-Dimensional Solar Powered Low-Thrust Trajectories in the Vicinity of the Ecliptic Plane," *Astronautica Acta*, Vol. 13, 1967, pp. 431–440.

[14] Van der Ha, J., and Modi, V. J., "Long-Term Evaluation of Three-Dimensional Heliocentric Solar Sail Trajectories with Arbitrary Fixed Sail Setting," *Celestial Mechanics and Dynamical Astronomy*, Vol. 19, No. 2, Feb. 1979, pp. 113–138.

[15] Vulpetti, G., "Sailcraft at High Speed by Orbital Angular Momentum Reversal," *Acta Astronautica*, Vol. 40, No. 10, May 1997, pp. 733–758. doi:10.1016/S0094-5765(97)00153-7

[16] Hughes, G. W., Macdonald, M., McInnes, C. R., Atzei, A., and Falkner, P., "Sample Return from Mercury and Other Terrestrial Planets Using Solar Sail Propulsion," *Journal of Spacecraft and Rockets*, Vol. 43, No. 4, 2006, pp. 828–835. doi:10.2514/1.15889

[17] Stevens, R., and Ross, I. M., "Preliminary Design of Earth-Mars Cyclers Using Solar Sails," *Journal of Spacecraft and Rockets*, Vol. 42, No. 1, Jan.–Feb. 2005, pp. 132–137. doi:10.2514/1.2947

[18] Sharma, D. N., and Scheeres, D. J., "Solar-System Escape Trajectories Using Solar Sails," *Journal of Spacecraft and Rockets*, Vol. 41, No. 4, July–Aug. 2004, pp. 684–687. doi:10.2514/1.2354

[19] Dachwald, B., "Optimal Solar Sail Trajectories for Missions to the Outer Solar System," AIAA/American Astronautical Society: Astrodynamics Specialist Conference and Exhibit, Providence, RI, AIAA Paper 2004-5406, Aug. 2004.

[20] Wright, J., *Space Sailing*, 1st ed., Gordon and Breach, Philadelphia, 1992, pp. 93–121.

[21] McInnes, C. R., *Solar Sailing: Technology, Dynamics and Mission Applications*, 1st ed., Springer Praxis, New York, 1999.

[22] Battin, R., *An Introduction to the Mathematics and Method of Astrodynamics*, AIAA Education Series, AIAA, Reston, VA, 1999, pp. 126–128.