A VARIATIONAL PROOF OF EXISTENCE OF TRANSIT ORBITS IN THE RESTRICTED THREE-BODY PROBLEM

RICHARD MOECKEL

ABSTRACT. Because of the Jacobi integral, solutions of the planar, circular restricted three-body problem are confined to certain subsets of the plane called Hill's regions. For certain values of the integral, one component of the Hill's region consists of disklike regions around each of the two primary masses, connected by a tunnel near the collinear Lagrange point, L_2. A *transit orbit* is a solution which crosses the tunnel, in a sense which can be made precise using Conley's isolating block construction.

For values of the Jacobi integral sufficiently close to its value at L_2, Conley found transit orbits by linearizing near the equilibrium point. The goal of this paper is to develop a method for proving existence of transit orbits for values of the Jacobi constant far from equilbrium. The method is based on the Maupertuis variational principle but isolating blocks turn out to play an important role.

1. The Restricted Three-Body Problem

The paper is divided into four sections. This section contains the equations of motion for the planar, circular restricted three-body problem or PCR3BP. In addition, the isolating blocks used by Conley [3] are described and used to define the concept of *transit orbit.* In section 2 the classical Maupertuis variational principle is explained and the idea of finding transit orbits as minimizers of this principle for curves crossing a rectangle is introduced. Section 3 shows how isolating blocks can be used to show that the variational minimizers are really classical solutions of the problem. Finally, in section 4, the hypotheses of the existence theorem are verified (with some numerical assistance) for two different choices of the mass ratio and Jacobi constant.

1.1. **Equations of Motion.** The restricted three-body problem models the motion of a small mass under the influence of the gravitational forces of two larger masses called the *primaries.* The primaries, which have masses $m_1 = 1 - \mu$ and $m_2 = \mu$, are assumed to move on circular orbits of the two-body problem, unaffected by the third mass. In a uniformly rotating coordinate system, the primaries will remain on the x-axis at positions $q_1 = (-\mu, 0)$ and $q_2 = (1 - \mu, 0)$. Let $(x, y) \in \mathbb{R}^2$ and $(u, v) \in \mathbb{R}^2$ denote the position and velocity vectors of the third mass with respect to the rotating coordinate system. Then if the primaries are rotating counterclockwise

Date: November 13, 2003, revised August 4, 2004.
2000 *Mathematics Subject Classification.* 70F10, 70F15, 37N05.
Key words and phrases. Celestial mechanics, three-body problem, calculus of variations.
Research supported by NSF grant DMS 0200992 and by the IMCCE, Paris.

with unit angular velocity, the equations of motion will be:

$$\dot{x} = u \qquad \dot{u} = V_x + 2v$$
$$\dot{y} = v \qquad \dot{v} = V_y - 2u \tag{1}$$

where

$$V(x,y) = \tfrac{1}{2}(x^2+y^2) + \frac{1-\mu}{r_{13}} + \frac{\mu}{r_{23}}$$

and $r_{13} = \sqrt{(x+\mu)^2+y^2}$, $r_{23} = \sqrt{(x+\mu-1)^2+y^2}$ are the distances to the primaries. V is just the Newtonian gravitational potential with an extra term representing the centrifugal force due to the rotating coordinates.

Equations (1) are the Euler-Lagrange equations of the action functional:

$$I(\gamma) = \int_{t_0}^{t_1} L(x,y,\dot{x},\dot{y})\,dt \tag{2}$$
$$L(x,y,u,v) = \tfrac{1}{2}(u^2+v^2) + V(x,y) + (xv - yu)$$

where $\gamma(t) = (x(t), y(t))$ is a parametrized curve in the plane.

1.2. **Energy Manifolds and Hill's Regions.** The Jacobi integral

$$H(x,y,u,v) = \tfrac{1}{2}(u^2+v^2) - V(x,y)$$

is a constant of motion for (1). For any constant $h < 0$ let

$$\mathcal{M}(h) = \{(x,y,u,v) \in \mathbb{R}^4 : H(x,y,u,v) = h\}$$
$$\mathcal{H}(h) = \{(x,y) \in \mathbb{R}^2 : \exists (u,v) \in \mathbb{R}^2, H(x,y,u,v) = h\}.$$

The constant h will be called the energy and $\mathcal{M}(h)$ will be called an *energy manifold*. Its projection onto the configuration space, $\mathcal{H}(h)$, is the corresponding *Hill's region*.

For points in $\mathcal{M}(h)$ one has $V(x,y)+h = \frac{1}{2}(u^2+v^2) \geq 0$. In fact, this inequality characterizes the Hill's region and so

$$\mathcal{H}(h) = \{(x,y) \in \mathbb{R}^2 : V(x,y) \geq -h\}.$$

The Hill's regions can be visualized by making a contour plot of $V(x,y)$ (see figure 1). The shaded region represents a value of h typical of those considered below.

1.3. **The Lagrange Point L_2.** The critical points of $V(x,y)$ are called the *Lagrange points*. They determine equilibrium points of the differential equations in rotating coordinates. In non-rotating coordinates, they determine periodic solutions for which the configuration rotates rigidly at constant speed. There are five well-known critical points, all of which are visible in figure 1. The points L_4 and L_5 correspond to equilateral triangle configurations. The other three critical points lie on the x-axis and so represent collinear configurations. These are critical points of the collinear potential $V(x,0)$.

It turns out that $V(x,0)$ is a convex function with a unique critical point in each of the intervals $(-\infty,-\mu), (\mu, 1-\mu), (1-\mu,\infty)$. This paper is mainly concerned with a neighborhood of L_2, the collinear Lagrange point which lies between the primaries. Let $\bar{x} \in (-\mu, 1-\mu)$ be the x-coordinate of L_2 and let its energy be $H(\bar{x},0,0,0) = -V(\bar{x},0) = \bar{h}$. The critical energy "manifold" $\mathcal{M}(\bar{h})$ is actually not a manifold near the critical point. But the nearby manifolds $\mathcal{M}(h)$, $h > \bar{h}$, are non-critical. The shaded Hill's region in figure 1 corresponds to such a value of h. Of course these nearby manifolds do not contain the equilibrium point. Instead they may contain orbits which remain for all time near L_2 but are not equilibria.

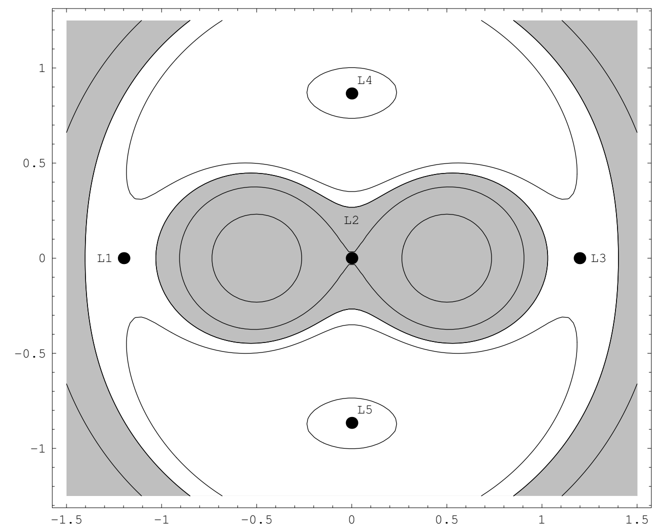

FIGURE 1. Hill's regions and Lagrange points for the planar restricted three-body problems with $\mu = \frac{1}{2}$. The shaded region is the Hill's region for energy $h = -1.8$.

For $h - \bar{h} > 0$ and small, one can investigate this question by local methods. First, consider the linearization of the differential equations near the equilibrium. The 4×4 linearized system has a matrix A which always has two real eigenvalues and two imaginary eigenvalues. In other words, the equilibrium is of saddle-center type. It follows from the center manifold theorem that the manifolds $\mathcal{M}(h)$ each contain a hyperbolic periodic orbit near L_2 for $h - \bar{h} > 0$ sufficiently small. When $h - \bar{h} > 0$ gets larger, it is not clear what will happen to this periodic orbit.

1.4. **Isolating Blocks and Transit Orbits.** In [3], Conley investigated the local dynamics near L_2 using isolating blocks. He considered a neighborhood of L_2 in the Hill's regions, $\mathcal{H}(h)$, of the form $a \leq x \leq b$ where $-\mu < a < \bar{x} < b < 1 - \mu$. Such a neighborhood, $T_h(a, b)$, can be viewed as a "tunnel" between the two lobes of the Hill's region as illustrated in figure 2.

The preimage of $T_h(a, b)$ in $\mathcal{M}(h)$ will be called $B_h(a, b)$. Conley showed that, at least for $h - \bar{h}$ sufficiently small, the sets B_h are manifolds-with-boundary which are *convex to the flow*. This means that any orbit which meets one of the boundary manifolds where $x = a$ or $x = b$ tangentially must lie outside of B_h both before and after the encounter. In later work with Easton, such submanifolds were called *isolating blocks* and the maximal invariant sets inside them were called *isolated invariant sets* [5, 6, 4]. The convexity condition ensures that the boundary points of the block are not part of the maximal invariant set in the block so the invariant set is in the interior of the block.

It is easy to formulate the isolating block condition for B_h using second derivatives. One requires that any orbit with $x = a$ and $\dot{x} = 0$ must also have $\ddot{x} < 0$ and

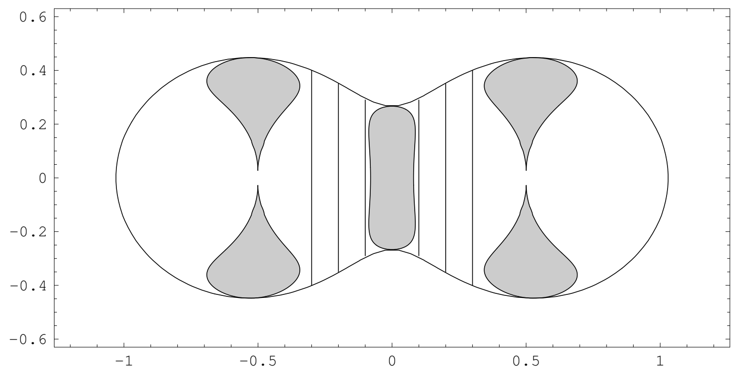

FIGURE 2. Tunnels $T_h(a,b)$ for $\mu = \frac{1}{2}$ and energy, $h = -1.8$. Any of the indicated line segments $x = a$ on the left of the origin could be paired with any segment $x = b$ on the right to construct an isolating block.

similarly that whenever $x = b$ and $\dot{x} = 0$ one must have $\ddot{x} > 0$. Now

$$\dot{x} = u$$
$$\ddot{x} = \dot{u} = V_x + 2v.$$

So the isolating block condition on the right boundary amounts to showing that $V_x > -2v$ for every point $(b, y, 0, v) \in \mathcal{M}(h)$ (note: Conley assumed a clockwise rotation of the primaries so there are some sign differences in his formulas). From the convexity of $V(x, 0)$ it follows that $V_x(b, 0) > 0$. Also, the equation $H = h$ implies that $v^2 \leq 2(V + h)$. Thus it suffices to show that

$$V_x^2(x,y) > 8(V(x,y) + h) \tag{3}$$

holds for all points $(x, y) \in \mathcal{H}_h$ with $x = b$. A similar analysis at $x = a$ yields the same inequality.

Inequality (3) can be verified for $h - \bar{h} > 0$ small by using Taylor expansions near L_2. This is one of Conley's results in [3]. One can also try to analyze inequality (3) for more general values of a, b, h. This can be done in some cases but the inequality is complicated enough to require some computer-assisted computations. The shaded region in figure 2 shows the part of the Hill's region where (3) is violated. Any vertical line segment in the complement will satisfy the convexity condition and so could be used to define the wall of an isolating block.

The existence of an isolating block, $B = B_h(a, b)$, makes it possible to give a clean classification of the possible behaviors of orbits in the tunnel, $T_h(a, b)$. Let ϕ_t be the flow and define subsets

$$S = \{p : \phi_t(p) \in B, t \in \mathbb{R}\}$$
$$S^+ = \{p : \phi_t(p) \in B, t \geq 0\}$$
$$S^- = \{p : \phi_t(p) \in B, t \geq 0\}$$

$S = S^+ \cap S^-$ is the maximal invariant set in the block and $S^\pm$ are its stable and unstable sets. Orbits in the complement $B \setminus (S^+ \cup S^-)$ must leave the block in both forward and backward time. They can be classified according to which wall, $x = a$ or $x = b$, they cross when entering and leaving. Let U_{ab} be the orbits (if any)

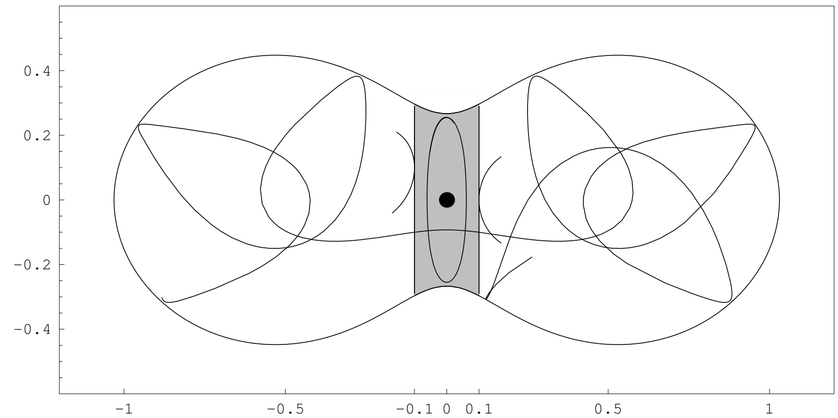

FIGURE 3. Isolating block for $\mu = \frac{1}{2}$ and $h = -1.8$. The orbits shown include a probable period orbit from the invariant set S, two orbits illustrating the convexity condition on the boundary segments, and a transit orbit.

which enter through the wall $x = a$ and leave through $x = b$ and define U_{aa}, U_{ba} and U_{bb} in a similar way. It follows easily from the convexity to the flow that each of these sets is open, while the sets $S, S^{\pm}$ are closed. Moreover, they are invariant under continuous changes in the positions of the walls, so long as the isolating block condition is preserved.

The open sets U_{aa} and U_{bb} are always nonempty because near any orbit which meets the wall $x = a$ or $x = b$ tangentially there is an orbit which crosses the wall and then immediately leaves through the same wall. It is not so easy to show that the sets of transit orbits U_{ab} and U_{ba} are nonempty though such orbits can be found numerically (see figure 3). The goal of the present paper is to develop a method for proving that they exist.

Although this paper is about transit orbits, the sets S, S^+, S^- are also of interest. For $h - \bar{h} > 0$ small, S consists of a single orbit, the hyperbolic periodic orbit mentioned above. The sets S^+, S^- are its stable and unstable manifolds. These are two-dimensional submanifolds of the three-dimensional energy manifold $\mathcal{M}(h)$. For larger energies, the nature of these sets in not clear. However, using the Conley index theory, one can show that S is always a nonempty compact invariant set and that S^+, S^- have topological dimension two and separate the energy manifold.

One of the motivations for this paper is a theorem of Easton which uses the existence of transit orbits to deduce some topological information about S. He shows that if all four of the sets U_{ij} are nonempty then the Cech cohomology group $\check{H}^1(S) \neq 0$, so at least the invariant set has something in common with the circle [6]. For the three-dimensional restricted three-body problem, a similar isolating block can be constructed. For $h - \bar{h} > 0$ sufficiently small, the invariant set is now a normally hyperbolic invariant three-sphere. In this case, Easton's theorem shows that for larger energies, the existence of the isolating blocks and transit orbits imply that $\check{H}^3(S) \neq 0$. Since a transit orbit for the planar problem is also a transit orbit for the three-dimensional problem, this will follow from the existence proof below.

2. Some Variational Existence Theorems

This section describes a classical variational existence theorem which applies, in particular, to the restricted three-body problem

2.1. Maupertuis' Variational Principle.

As noted above, the differential equations of the restricted three-body problem are the Euler-Lagrange equations of the action functional I of (2). For the construction of transit orbits, it is necessary to restrict attention to a fixed energy level $H(x, y, u, v) = h$. Solutions on this energy level are stationary curves of a different action functional:

$$J(\gamma) = \int_{t_0}^{t_1} F(x, y, \dot{x}, \dot{y})\, dt \tag{4}$$

$$F(x, y, u, v) = \sqrt{2(V(x, y) + h)}\,\sqrt{u^2 + v^2} + (xv - yu)$$

where this time the curve $\gamma(t) = (x(t), y(t))$ is required to lie in the Hill's region $\mathcal{H}(h) = \{(x, y) \in \mathbb{R}^2 : V(x, y) + h \geq 0\}$. The value of the functional J is invariant under reparametrizations of γ. Every stationary curve of I is also stationary for J but a stationary curve of J is stationary for I only when it is parametrized so that $H = h$. In studying J, however, any convenient convention about parametrizations may be used. For example, one may assume that the time interval is $[t_0, t_1] = [0, 1]$.

2.2. Classical Results on Parametric Problems.

There is an extensive classical theory of "parametric" variational problems like (4). Let $U \subset \mathbb{R}^2$ be an open set. $F(x, y, u, v)$ will be called a *parametric integrand on U* if it is C^3 function on $U \times (\mathbb{R}^2 \setminus 0)$ which is positively homogeneous of degree one with respect to (u, v), i.e.,

$$F(x, y, ku, kv) = kF(x, y, u, v)$$

for every real number $k > 0$. This condition assures that the corresponding functional is invariant under sense-preserving reparametrizations of the curve. Differentiating the homogeneity relation with respect to k at $k = 1$ gives $uF_u + vF_v = F$. Then differentiating again with respect to u and with respect to v gives:

$$uF_{uu} + vF_{uv} = uF_{uv} + vF_{vv} = 0.$$

It follows that the following three ratios are equal (where defined):

$$\frac{F_{uu}}{v^2} = \frac{F_{vv}}{u^2} = -\frac{F_{uv}}{uv}.$$

For $u^2 + v^2 \neq 0$, let $F_1(x, y, u, v)$ denote this common value. For example, if F is the integrand of (4) then

$$F_1(x, y, u, v) = \frac{\sqrt{2(V(x, y) + h)}}{(u^2 + v^2)^{\frac{3}{2}}}. \tag{5}$$

The following result can be found in [2, ch.VII]

Theorem 1. *Suppose F is parametric integrand on U and that $R \subset U$ is a compact, convex subset such that*

$$F(x, y, u, v) > 0 \qquad F_1(x, y, u, v) > 0$$

in $R \times (\mathbb{R}^2 \setminus 0)$. Then given two points $P, Q \in R$, the functional

$$J(\gamma) = \int_0^1 F(x, y, \dot{x}, \dot{y})\, dt \tag{6}$$

attains an absolute minimum on the set $\mathcal{K}$ of rectifiable curves in R from P to Q.

A few remarks are needed about rectifiable curves and about the interpretation of the integral; see [2, 7] for details. Let $\gamma(t) = (x(t), y(t))$ denote a continuous curve. The distance $d(\gamma, \gamma')$ between continuous parametrized curves defined on a common interval, $[t_0, t_1]$, will be the usual uniform distance in $\mathbb{R}^2$. If $h : [t_0', t_1'] \to [t_0, t_1]$ is a sense-preserving homeomorphism then the curve $\gamma'(t) = \gamma(h(t))$ will be considered equivalent to γ. More generally, two continuous curves γ, γ' are called equivalent when there is a sequence of such reparametrizations, $h_n(t)$, such that $d(\gamma(h_n(t)), \gamma'(t)) \to 0$. An equivalence class can be viewed as an unparametrized curve. The length of a parametrized curve is defined in the familiar way by taking the supremum of the lengths of approximating polygons over all partitions of $[t_0, t_1]$. The lengths of equivalent curves are equal so the length of an unparametrized curve is well-defined. A curve is rectifiable if it has finite length. The distance between unparametrized curves is defined as the infimum of the $d(\gamma, \gamma')$ over parametrized representatives of the two curves. If $\gamma(t) = (x(t), y(t))$ is any parametrization of a rectifiable curve then the functions $x(t), y(t)$ are of bounded variation and hence are differentiable for almost all $t \in [0, 1]$. If these functions have no higher regularity then the integral (6) has to be interpreted as a "Weierstrass integral". Because of the positive homogeneity of F, the value of this integral is independent of the choice of parametrization. But any rectifiable curve can be reparametrized by arclength (or by a constant multiple of arclength to arrange that the time interval be $[0, 1]$). If this is done, then the functions $x(t), y(t)$ are Lipschitz and so their derivatives $\dot{x}(t), \dot{y}(t)$ will be in $L_\infty([0, 1])$. Then (6) can be interpreted as a Lebesgue integral.

Here is an outline of the proof of theorem 1. Since $F \geq \alpha > 0$ for some constant $\alpha > 0$, the functional satisfies $J(\gamma) \geq \alpha L(\gamma)$ where L denotes the length of the rectifiable curve. Let $\mu = \inf_{\gamma \in \mathcal{K}} J(\gamma) > 0$. Then one can restrict attention to the set $\mathcal{K}_\lambda$ of rectifiable curves from P to Q whose length does not exceed $\lambda = (\mu + 1)/\alpha$. By a theorem of Hilbert, this set of curves is compact with respect to the metric described above. The hypothesis $F_1 \geq 0$ can be used to show that J defines a lower semicontinuous function on this space of curves. The existence of a minimizer follows immediately.

The regularity of the minimizing curve, $\bar{\gamma} \in \mathcal{K}$, is also addressed in the classical literature. It is based on the locally minimizing properties of the classical solutions. The following result is due to Weierstrass [2, sec.28e]:

Theorem 2. *Under the hypotheses of theorem 1, there are constants $\delta, \delta_0 > 0$ such that any two points $P, Q \in R$ with $|P - Q| < \delta_0$ are joined by a unique classical solution of length less than δ. Moreover, this solution is the minimizer of J over all rectifiable curves from P to Q which remain in R_δ, the δ neighborhood of R.*

This leads to the following result about the minimizer $\bar{\gamma}$ [2, ch.VII]:

Theorem 3. *Let $\bar{\gamma}(t), 0 \leq t \leq 1$ be a minimizing rectifiable curve as in theorem 1. Then $\bar{\gamma}$ has no self-intersections. Moreover, if $I \subset [0, 1]$ is an open interval such that $\bar{\gamma}(t)$ is in the interior of R for all $t \in I$, then $\bar{\gamma}|_I$ is a C^2 solution of the Euler-Lagrange equations.*

The positivity of F rules out self-intersections since the part of the curve between the intersections could be deleted to give a lower value of J. The proof of the second statement uses theorem 2. Any two sufficiently near points on $\bar{\gamma}$ which are in the interior of R will be connected by a unique minimizing classical solution curve in the interior of R and this curve must agree with the segment of $\bar{\gamma}$ between these points since $\bar{\gamma}$ is minimizing.

2.3. **Application to the Restricted Three-Body Problem.** To use this theorem for the restricted three-body problem, it is necessary to avoid the boundary of the Hill's region. Let U be the interior of $\mathcal{H}(h)$ and let $R \subset U$ be a rectangular region (see figure 4). Clearly F is C^3 and positively homogeneous with respect to (u, v) on $U \times (\mathbb{R}^2 \setminus 0)$ and (5) shows that $F_1 > 0$.

After a slight modification of the functional, the positivity condition $F > 0$ will hold provided the rectangle R is not too large. To see this, first choose constants $c(R), C(R)$ such that

$$0 < c(R) \leq \sqrt{2(V(x,y)+h)} \leq C(R) \tag{7}$$

for all $(x, y) \in R$. So the first term of F is at least $c(R)\sqrt{u^2+v^2}$. Next, let (α, β) denote the center of R and k_1, k_2 the half-width and half-height, respectively. If the second term in F, $xv - yu$, is replaced by $(x-\alpha)v - (y-\beta)u$, then $J(\gamma)$ is only changed by a constant, independent of γ and the value of F_1 is not affected. So after a translation of coordinates, one may assume without loss of generality that $(\alpha, \beta) = (0, 0)$. Then since $xv - yu = (x, y) \cdot (v, -u)$, the Cauchy inequality gives

$$F(x,y,u,v) \geq \left(c(R) - \sqrt{x^2+y^2}\right)\sqrt{u^2+v^2}.$$

If the rectangle is chosen so that

$$c(R) > \sqrt{k_1^2 + k_2^2} \tag{8}$$

then the required condition $F > 0$ will hold in $R \times (\mathbb{R}^2 \setminus 0)$.

Using theorem 1 and theorem 3 for the case of the restricted three-body problem gives:

Theorem 4. *Let R be a rectangle in the interior of the Hill's region $\mathcal{H}(h)$ such that inequality (8) holds. Let $P, Q \in R$ and let $\mathcal{K}$ be the set of rectifiable curves in R from P to Q. Then there is a curve in $\mathcal{K}$, $\bar{\gamma}(t), 0 \leq t \leq 1$, which minimizes J on $\mathcal{K}$. Moreover, on any open interval $I \subset [0, 1]$ such that $\bar{\gamma}(t)$ is in the interior of R, $\bar{\gamma}(t)$ is a classical solution of the restricted three-body problem.*

3. Existence of Transit Orbits

Fix a value of h and constants a, b so that $B_h(a, b)$ is an isolating block. Theorem 4 will be applied to a rectangle of the form $R = [a, b] \times [-c, c]$ with the points P, Q chosen on the left and right edges. Then any classical solution connecting P and Q in R will be a transit orbit. By theorem 4, the minimizer $\bar{\gamma}(t)$ will be such a solution provided it lies in the interior of the rectangle R (except for its endpoints P and Q). To guarantee this, one needs to impose additional conditions on R.

3.1. **Directional Convexity to the Flow.** The constant $c > 0$ will be chosen so that the bottom edge of R is convex to the flow for solutions moving from left to right while the top edge is convex for solutions moving from right to left. More precisely, the required condition on the bottom edge is that for any solution with

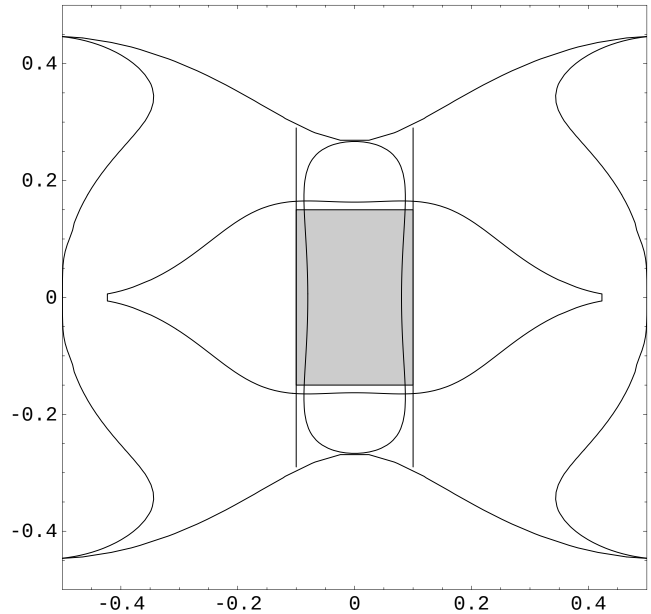

FIGURE 4. Rectangle convex to the flow for $\mu = \frac{1}{2}, h = -1.8$. The left and right edges are taken from an isolating block; the top and bottom are only directionally convex to the flow.

energy h such that $a \leq x \leq b$, $y = -c$, $\dot{y} = 0$ and $\dot{x} > 0$, one has $\ddot{y} < 0$. Now since $\ddot{y} = V_y - 2u$ and $u = \dot{x} > 0$, it suffices to check that the inequality

$$V_y^2(x, y) < 4u^2 = 8(V(x, y) + h) \tag{9}$$

holds along the bottom edge of R. A similar discussion for the top boundary leads to the same inequality. The lemon-shaped region in figure 4 shows the part of Hill's region where the required inequality holds. A rectangle R satisfying all of the required convexity conditions is also indicated.

The convexity conditions on the boundary of R will keep the minimizer from intersecting three out of the four boundary segments, as the following lemma shows.

Lemma 1. *Let R be chosen so that (8) holds and so that the left and right edges are convex to the flow and the bottom boundary is convex to the flow for orbits moving left to right. If P and Q are chosen on the left and right edges, respectively, then a miniming curve $\bar{\gamma}(t)$ as in theorem 4 does not intersect the left, right or bottom edges of R.*

Proof. Consider, for example, the right edge E where $x = b$. Suppose $P_1 = \bar{\gamma}(t_1)$ and $P_2 = \bar{\gamma}(t_2)$, $t_1 < t_2$, are intersection points of the minimizing curve with E. The convexity to the flow implies that if P_1, P_2 are sufficiently close, then they are connected by a solution of the Euler-Lagrange equations which lies strictly to the left of E. To see this, note that convexity to the flow is valid for nearby line segments $E' = \{x = b'\}$ with $b - b'$ sufficiently small. Now any pair of sufficiently near points in R are connected by a small solution curve. If the solution connecting the given points P_1 and P_2 did not lie strictly inside R, it would have to be tangent from the left side to some line segment E', a contradiction. By theorem 2, these small connecting curves give the minimum possible value of J for all curves in R from P_1 to P_2 (in fact, even in $R_\delta \supset R$) and so they must agree with the curve $\bar{\gamma}(t), t_1 \leq t \leq t_2$. Since $\bar{\gamma}$ has no self-intersections, it follows that there are no other intersection points of $\bar{\gamma}$ and E between P_1 and P_2. Hence the intersection of $\bar{\gamma}$ and E consists, at most, of isolated points. Now suppose $P_0 = \bar{\gamma}(t_0)$ is an isolated intersection with $0 < t_0 < 1$. Let E' denote a nearby vertical line segment $x = b'$ in R with $b' < b$. If $b - b'$ is sufficiently small then $\bar{\gamma}$ must contain two nearby intersection points P_1', P_2' with E'. But E' is also convex to the flow, so the same argument would imply that the minimizing curve must lie to the left of E'. Since P_0 lies to the right of E', this is a contradiction. It follows that $\bar{\gamma}(t), 0 < t < 1$ has no intersections with E at all. Similar arguments apply to the left edge.

Now let E denote the bottom edge. The convexity to the flow for left to right solutions means that for any two sufficiently close points of that edge, the Euler-Lagrange solution from the left one to the right one lies above E. Now if $P_1 = \bar{\gamma}(t_1), P_2 = \bar{\gamma}(t_2)$, $t_1 < t_2$, are intersections with the bottom edge, then P_1 must lie to the left of P_2. Otherwise, the curve segment $\bar{\gamma}(t)$, $0 \leq t \leq t_1$ would separate the point P_2 from the endpoint Q forcing $\bar{\gamma}$ to have self-intersections. As above, it follows that the intersections of $\bar{\gamma}$ with the bottom edge are isolated. But near any isolated intersection one could find intersections P_1', P_2' with a nearby horizontal line segment and these would also have to be ordered from left to right. Since this nearby line segment shares the convexity property, this would lead to a contradiction. Hence there can be no intersections of $\bar{\gamma}$ with E and the proof is complete. QED

A similar result holds for minimizers moving from right to left. If P is on the right and Q on the left and if the top boundary is convex to the flow for right to left solutions, then the minimizer from P to Q in R cannot intersect the left, right or top edges.

Convexity conditions on the boundary, including directional convexity, have been used in the calculus of variations for many years [11, 9, 1]. It is interesting that the same conditions arise for topological reasons in the theory of isolating blocks.

3.2. **An Existence Theorem for Transit Orbits.** It only remains to keep the minimizer $\bar{\gamma}(t)$ from intersecting the fourth edge of the rectangle R. Consider the case where P is on the left edge and Q is on the right. Fixing the points P and Q, let $\nu(P,Q) = \inf_{\mathcal{K}} J(\gamma)$ be the infimum of J on the set, $\mathcal{K}$, of rectifiable curves from P to Q in R and let $\nu_0(P,Q) = \inf_{\mathcal{K}_0} J(\gamma)$ be the infimum on the subset $\mathcal{K}_0 \subset \mathcal{K}$ of curves with no self-intersections which meet the top of the rectangle. Clearly the inequality

$$\nu(P,Q) < \nu_0(P,Q) \tag{10}$$

suffices to show that a minimizer $\bar{\gamma}(t)$ in $\mathcal{K}$ does not intersect the top edge. Combining this with theorem 4 and lemma 1 gives:

Theorem 5. *Let R be chosen such that (8) holds, such that the left and right edges are convex to the flow and such that the bottom boundary is convex to the flow for orbits moving left to right. If P and Q are chosen on the left and right edges, respectively, and if (10) holds then there exists a transit orbit from P to Q in R.*

Of course a similar result holds when P is on the right and Q is on the left. Transit orbits in the other direction can also be obtained from symmetry. In fact, it is easy to see that if a classical solution is reflected through the x-axis and time is reversed, another solution is obtained.

4. Examples

This section is devoted to demonstrating (with the aid of some numerical computations) the existence of transit orbits for two particular choices of the mass ratio, μ, and the energy, h. A rectangle R and points P, Q will be found which satisfy all the hypotheses of theorem 5.

4.1. **Equal Primaries.** The simplest choice of mass ratio is $\mu = \frac{1}{2}$. Then the Newtonian potential

$$V(x,y) = \frac{1}{2\sqrt{(x+\frac{1}{2})^2+y^2}} + \frac{1}{2\sqrt{(x-\frac{1}{2})^2+y^2}}$$

is an even function of both x and y. The Lagrange point L_2 is at the origin and the critical energy level is $H(0,0,0,0) = -V(0,0) = -2$. When the energy is $h = -1.8$ the segments with $x = \pm 0.1$ determine an isolating block. The top and bottom boundaries of the rectangle, $y = \pm c$, must be chosen so that (9) holds. Since $V_y(x,0) = 0$ it clearly holds for all c sufficienly small, but in order to verify the inequality (10) it is necessary to use a rectangle whose aspect ratio k/h is not too small. The choice $c = 0.15$ is simple and is close to the maximum possible based on the numerical computation. So $R = [-0.1, 0.1] \times [-0.15, 0.15]$ will be used. This is the shaded rectangle in figure 4.

The maximum and minimum of $V(x,y)$ occur at the points $(\pm 0.1, 0)$ and $(0, \pm 0.15)$, respectively. So the constants $c(R)$ and $C(R)$ are given by

$$c(R) = \sqrt{2(V(0,0.15) - 1.8)} > 0.503 \qquad C(R) = \sqrt{2(V(0.1,0) - 1.8)} < 0.76.$$

So inequality (8) becomes

$$0.503 > \sqrt{0.0325} \approx 0.181.$$

It only remains to verify (10). For this it is convenient to modify the functional $J(\gamma)$ by changing the path of integration for the term $xv - yu$ so that it includes the given path γ from P to Q but then follows the right, bottom and left edges of the rectangle to complete a loop back to P. Since the added part of the path does not depend on γ, this just changes J by an additive constant. If J' denotes this new functional then

$$J'(\gamma) = \int_0^1 \sqrt{2(V(x,y)+h)}\,\sqrt{\dot{x}^2+\dot{y}^2}\,dt \;-\; 2A(\gamma)$$

where $A(\gamma) \geq 0$ is the area below γ in R (assuming the path has no self-intersections). It will be shown that $\nu' < \nu'_0$ where ν', ν'_0 are the infima of J' on $\mathcal{K}, \mathcal{K}_0$.

Let $\gamma_l(t)$ be the straight line path from P to Q and let $A_0(P,Q)$ be the area below this path in R. Then

$$\nu'(P,Q) \leq J'(\gamma_l) \leq C(R)|P-Q| - A_0(P,Q)$$

where $|P-Q|$ denotes the Euclidean distance. For example, if P and Q are the lower left and right corners of R then $\nu'(P,Q) \leq 2C(R)h < 0.152$. (This rather crude estimate will suffice, but a more accurate value is $\nu'(P,Q) \approx 0.109$.)

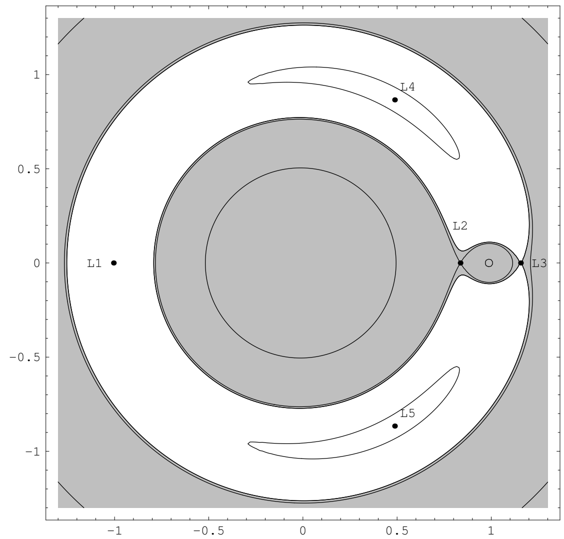

FIGURE 5. Hill's regions and Lagrange points for the planar restricted three-body problems with $\mu = \frac{1}{81}$. The shaded region is the Hill's region for energy $h = -1.587$.

To estimate $\nu_0(P, Q)$ from below, note that

$$J'(\gamma) \geq c(R)L(\gamma) - 8k_1k_2$$

since $4k_1k_2$ gives the area of the whole rectangle. The minimum value of $L(\gamma)$ for a path in $\mathcal{K}_0$ is the length of the path of a light ray from P to Q which bounces off the top edge of R. This can be written as $|P - Q'|$ where Q' is the reflection of Q through the top boundary line. Thus

$$\nu_0'(P, Q) \geq c(R)|P - Q'| - 8k_1k_2. \tag{11}$$

For example, if P and Q are the lower left and right corners of R then $\nu_0'(P, Q) \geq 2c(R)\sqrt{k_1^2 + 4k_2^2} - 8k_1k_2 > 0.198$. Thus, for the lower corners of the rectangle, one has

$$\nu'(P, Q) < 0.152 < 0.198 < \nu_0'(P, Q)$$

and hence also $\nu(P, Q) < \nu_0(P, Q)$. By continuity of the functions used as bounds, this inequality continues to hold for all points sufficiently close to the bottom of the rectangle. Thus there exist transit orbits connecting any such pair of points.

4.2. **The Lunar Case.** Next let $\mu = \frac{1}{81}$ which approximates the relative masses of the earth and the moon. This time the Hill's regions are not so symmetric (see figure 5). There is a large lobe near the earth and a small lobe near the moon. The Lagrange point L_2 is now located at $(\bar{x}, 0)$ where $\bar{x} \approx 0.836$ and the critical energy is $\bar{h} \approx -1.59507$. It will be shown that transit orbits exist for the energy $h = -1.587$.

Let $R = [0.798, 0.864] \times [-0.039, 0.039]$. The boundaries are chosen so that the convexity conditions are satisfied. Figure 6 shows the rectangle together with the boundary curves of the regions where the convexity inequalities hold. The left and right edges, $x = a, b$, must be chosen outside the kidney-shaped curve so that

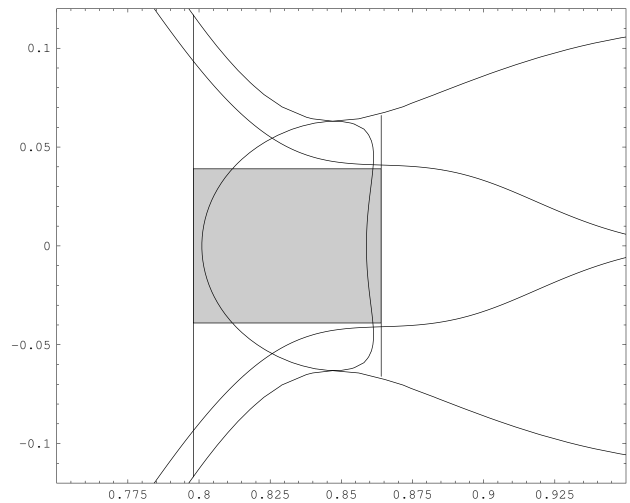

FIGURE 6. Rectangle convex to the flow for $\mu = \frac{1}{81}, h = -1.587$.

$B_h(a, b)$ is an isolating block. Evidently this is the case for $[a, b] = [0.798, 0.864]$. The top and bottom edges must be chosen inside the other curve in the figure to achieve the required directional convexity.

The maximum of $V(x, y)$ on R is achieved at $(0.798, 0)$ and the minimum is achieved on the top and bottom edges at a point whose position has been estimated numerically. In this way one obtains numerical estimates

$$0.0992 < c(R) < 0.0993 \qquad C(R) > 0.174.$$

The half-width and half-height of R are $k_1 = 0.033$ and $k_2 = 0.039$. Inequality (8) holds (after translating coordinates to the center of the rectangle) since $0.0992 > \sqrt{k_1^2 + k_2^2} \approx 0.052$.

Inequality (10) is more challenging this time and more careful estimates of the actions will be needed. As for the previous case, let P, Q be the bottom corners of the rectangle and replace the functional J by J'. Using the straightline path, one finds

$$\nu'(P, Q) \leq J'(\gamma_l) = \int_a^b \sqrt{2(V(x, -0.039) - 1.587)}\, dx < 0.0078$$

where numerical evaluation of the definite integral has been used to get the last inequality.

On the other hand, the simple estimate (11) gives approximately 0.0065 which is not good enough. The problem is that the minimum value $c(R)$ gives too small an estimate for $\sqrt{2(V(x, y) + h)}$. To get a better estimate, divide the rectangle into three horizontal strips $R_1 = [a, b] \times [-0.039, -0.02]$, $R_2 = [a, b] \times [-0.02, 0.02]$ and $R_3 = [a, b] \times [0.02, 0.039]$. On the middle strip the minimum of $\sqrt{2(V(x, y) + h)}$ is numerically found to satisfy $c_2 > 0.12$. Let $J''(\gamma) = \int_0^1 f(x, y)\sqrt{\dot{x}^2 + \dot{y}^2}\, dt - 8k_1k_2$, where $f(x, y)$ has the constant value $c(R)$ in rectangles R_1 and R_3 and c_2 in rectangle R_2. Now the minimum of the integral in $J''(\gamma)$ over paths from P to Q which meet the top of the rectangle is achieved along the path, γ_0, of a light ray which reflects off the top edge but also obeys Snell's law of refraction at the boundaries between the strips. It is not difficult to estimate this numerically. One finds $\nu_0'(P,Q) \geq J''(\gamma_0) > 0.0083$. This suffices to show that $\nu'(P,Q) < \nu_0'(P,Q)$ as required.

References

1. G.D. Birkhoff, *Dynamical systems with two degrees of freedom*, Trans. Amer. Math. Soc., **18** (1917) 199-300.
2. O. Bolza, *Lectures on the Calculus of Variations*, Haffner, New York (1946).
3. C.C. Conley, *Low energy transit orbits in the restricted three-body problem*, SIAM J. Appl. Math., **16, 4** (1968) 732–746.
4. C.C. Conley, *Isolated Invariant Sets and the Morse Index*, CBMS Regional Conference Series, 38, American Mathematical Society (1978).
5. C.C. Conley and R.W. Easton, *Isolated invariant sets and isolating blocks*, Trans. AMS, **158, 1** (1971) 35–60.
6. R.W. Easton, *Existence of invariant sets inside a submanifold convex to the flow*, JDE, **7** (1970) 54–68.
7. G.M. Ewing, *Calculus of Variations with Applications*, Dover Publishing, New York (1985).
8. W. S. Koon, et.al., *Constructing a low energy transfer between Jovian moons*, in Celestial Mechanics, Contemporary Mathematics, vol. 292, American Mathematical Society (2002) 129–146.
9. Signorini, Esistenza di un'estremale chiusa dentro un contorno di Whittaker, Rend. Circ. Mat. Palermo, **33** (1912), 187–193.
10. C. Simo,et.al.,*Dynamics and Mission Design Near Libration Points, vol I,III*, World Scientific (2001).
11. Whittaker, *On periodic orbits in the restricted problem of three bodies*, Monthly Not. Royal Astr. Soc., **62** (1901-02) 346–352.
12. A. Wintner, *The Analytical Foundations of Celestial Mechanics*, Princeton Math. Series 5, Princeton University Press, Princeton, NJ (1941).

School of Mathematics, University of Minnesota, Minneapolis MN 55455
E-mail address: rick@math.umn.edu