Pergamon

www.elsevier.com/locate/jappmathmech

PII: S0021-8928(00)00121-0

J. Appl. Maths Mechs, Vol. 64, No. 6, pp. 903–907, 2000

0021-8928/00/$—see front matter

TWO-LINK BILLIARD TRAJECTORIES: EXTREMAL PROPERTIES AND STABILITY†

V. V. KOZLOV

Moscow

(*Received* 10 *March* 2000)

Two-link periodic trajectories of a plane convex billiard, when a point mass moves along a segment which is orthogonal to the boundary of the billiard at its end points, are considered. It is established that, if the caustic of the boundary lies within the billiard, then, in a typical situation, there is an even number of two-link trajectories and half of them are hyperbolic (and, consequently, unstable) and the other half are of elliptic type. An example is given of a billiard for which the caustic intersects the boundary and all of the two-link trajectories are hyperbolic. The analysis of the stability is based on an analysis of the extremum of a function of the length of a segment of a convex billiard which is orthogonal to the boundary at one of its ends.

1. STABILITY CONDITIONS

Suppose there is a smooth regular closed curve Γ – which is the boundary of a plane billiard. We shall assume that its curvature is positive at all points. Suppose there is a segment of length l which intersects Γ orthogonally at the end points γ_1 and γ_2. Then, this segment is a two-link periodic trajectory of the convex billiard. We recall that, according to Birkhoff [1], a dynamical system within Γ with elastic collisions is called a billiard. Suppose R_1 and R_2 are the radii of curvature of Γ at the points γ_1, γ_2; to be specific, we shall assume that $R_1 \leqslant R_2$.

It has been established in [2] that a two-link periodic trajectory will be elliptic (its complex multipliers λ_1, λ_2 lie on the unit circle) if

$$0 < l < R_1 \quad \text{or} \quad R_2 < l < R_1 + R_2 \tag{1.1}$$

As a rule, elliptic trajectories are stable in the Lyapunov sense [3]. If, however

$$R_1 < l < R_2 \quad \text{or} \quad R_1 + R_2 < l \tag{1.2}$$

the trajectory will be hyperbolic (λ_1, λ_2 are real and $|\lambda_1| > 1$, and $\lambda_2| < 1$). Hyperbolic trajectories are, of course, unstable.

When $l = R_1 + R_2$, we have the degenerate case: $\lambda_1 = \lambda_2 = 1$. If $l = R_1$ or $l = R_2$, then $\lambda_1 = \lambda_2 = -1$. An alternative derivation of the stability conditions (1.1) and (1.2) can be found in [4].

Example. Suppose the boundary Γ is an ellipse with semi-axes $a \geqslant b > 0$. If $a = b$, then we have a whole family of degenerate two-link trajectories. We shall fix a and decrease b. Then, the major axis of the ellipse becomes a hyperbolic trajectory and the minor axis becomes an elliptic trajectory. It is true that the property of ellipticity is lost when $a^2 = 2b^2$: $\lambda_1 = \lambda_2 = -1$. When b is reduced further, the minor axis again becomes elliptic trajectory. Using the pattern of the trajectories of an elliptic billiard [1], it is also possible to show that, when $a^2 = 2b^2$, the minor axis becomes stable in the Lyapunov sense.

Among the segments with ends in Γ, there is a segment of maximum length. It will obviously be a two-link periodic trajectory. In the neighbourhood of the end points, the boundary Γ is described by the equations

$$y_1 = a_1 x_1^2 + o(x_1^2), \quad y_2 = l - a_2 x_2^2 + o(x_2^2)$$

The radii of curvature are respectively equal to $R_1 = 1/(2a_1)$ and $R_2 = 1/(2a_2)$. A two-link trajectory corresponds to the values $x_1 = x_2 = 0$. The square of the distance between the points with coordinates (x_1, y_1) and (x_2, y_2) is equal to

†*Prikl. Mat. Mekh.*, Vol. 64, No. 6, pp. 942–946, 2000.

$$d^2 = l^2 + f_2 + \ldots, \quad f_2 = (1-2a_1l)^2x_1^2 - 2x_1x_2 + (1-2a_2l)^2x_2^2$$

The dots denote terms of order $\geqslant 3$.

Since the two-link trajectory has maximum length, the quadratic form $(-f_2)$ is non-negative. In particular.

$$l \geqslant \frac{1}{2a_1} + \frac{1}{2a_2} = R_1 + R_2 \tag{1.3}$$

Hence, it follows, in particular, that the caustic curve of the boundary Γ (the curve of the centres of curvature) cannot lie wholly within Γ.

Example. Suppose Γ is an ellipse with semi-axes $a \geqslant b$. If $a^2 < 2b^2$, then the caustic curve (which is an asteroid) lies within the ellipse. When $a^2 = 2b^2$, it rests at the ends of the minor axis, and, when $a^2 > 2b^2$, the caustic curve lies outside the limits of the ellipse (Fig. 1).

If the critical point $x_1 = x_2 = 0$ of the function d is non-degenerate (the form f_2 is non-degenerate), inequality (1.3) becomes a strict inequality. Hence, in the typical case, the segment of maximum length is a hyperbolic trajectory with positive multipliers. This result was mentioned in [5] as a consequence of a more general construction.

2. EXTREMA AND STABILITY

Suppose γ is a point on the boundary Γ and that $I(\gamma)$ is a segment with its ends on Γ, orthogonal to Γ at the point γ (Fig. 2). We denote the length of the segment $I(\gamma)$ by $L(\gamma)$. By virtue of the convexity of the curve Γ, the function L: $\Gamma \to \mathbb{R}$ is smooth. Suppose $\gamma' \in \Gamma$ is the other end of the segment $I(\gamma)$.

Assertion 1. Suppose γ is a stationary point of the function L and the value of $L(\gamma)$ is not equal to the radius of curvature of the curve Γ at the point γ (γ' does not belong to the caustic of Γ). Then, the segment $I(\gamma)$ is a two-link periodic trajectory.

Proof. The equations of the curve Γ in the neighbourhood of the points γ and γ' have the form

$$y = a_1x^2 + o(x^2), \quad y = l + \mu x + o(x)$$

In this notation,

$$L^2 = l^2 + 2l\mu(1 - 2a_1l)x + o(x)$$

By assumption, $x = 0$ is the critical point of the function L. Consequently, $\mu(1 + l/R_1) = 0$, where R_1 is the radius of curvature of the curve Γ at the point γ. Since, $l = L(\gamma) \neq R_1$, $\mu = 0$. This means that the segment $I(\gamma)$ is orthogonal to Γ at the point γ, which it was required to prove.

Suppose $I(\gamma)$ is a two-link trajectory. Then, the function L takes stationary values at the points γ and γ'. Suppose the curve Γ is given by the equations

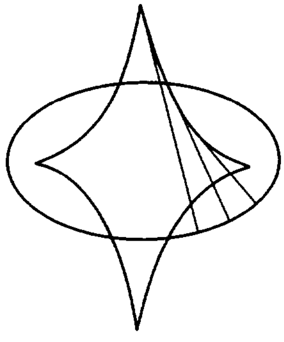

Fig. 1

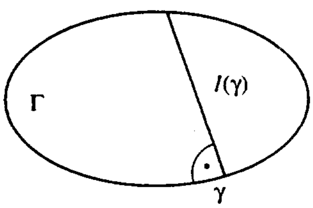

Fig. 2

$$y = a_1x^2 + o(x^2), \quad y = l - a_2x^2 + o(x^2)$$

in the neighbourhood of these points, respectively. The expansion of the function L^2 in a Taylor series in the neighbourhood of the point γ has the form

$$L^2 = l^2 + (2a_1l - 1)[2a_1l - 2a_2l(2a_1l - 1)]x^2 + o(x^2) \tag{2.1}$$

If the stationary point $x = 0$ is non-degenerate, then $2a_1l - 1 \neq 0$ and, consequently, γ' does not belong to the caustic of the boundary Γ (compare with Assertion 1).

Assertion 2. Suppose γ and γ' are non-degenerate stationary points of the function L. Then,

1. if γ and γ' are points of a local minimum of L, the trajectory $I(\gamma) = I(\gamma')$ is elliptic;
2. if γ and γ' are points of a local maximum and minimum of L, the trajectory $I(\gamma)$ is hyperbolic with negative multipliers;
3. if γ and γ' are points of a local maximum of L and if the caustic of the boundary Γ does not intersect Γ, the trajectory $I(\gamma)$ is hyperbolic with positive multipliers.

This result uniquely relates the extremal properties of the length of a two-link trajectory to its stability.

Proof.

1. Suppose γ is the point of a local minimum of the function L. Then, by (2.1), we have the inequality

$$(2a_1l - 1)[2a_1l - 2a_2l(2a_1l - 1)] > 0 \tag{2.2}$$

Taking into account the fact that a_1, a_2 are positive, we obtain

$$l > 1/(2a_1) \text{ and } l < 1/(2a_1) + 1/(2a_2) \tag{2.3}$$

Since the function L also has a minimum at the point γ', the inequalities

$$l > 1/(2a_2) \text{ and } l < 1/(2a_1) + 1/(2a_2) \tag{2.4}$$

are derived in a similar manner. Since $R_1 = 1/(2a_1)$, $R_2 = 1/(2a_2)$ and $R_1 \leqslant R_2$, from (2.3) and (2.4) we obtain the second inequality of (1.1). Consequently, the trajectory I is elliptic.

2. Suppose γ is the point of a local maximum of the function L. We then have an inequality opposite to (2.2). The intervals

$$l < 1/(2a_1) \quad \text{or} \quad l > 1/(2a_1) + 1/(2a_2) \tag{2.5}$$

are the solution of this inequality. On the other hand, the function L reaches a minimum at the point γ'. Consequently, inequalities (2.4) hold. Inequalities (2.4) and (2.5) give the first inequality of (1.2). Therefore, I is a hyperbolic trajectory with negative multipliers.

3. If γ and γ' are local maxima of the function L, the inequalities

$$l < 1/(2a_2) \quad \text{or} \quad l > 1/(2a_1) + 1/(2a_2) \tag{2.6}$$

have to be added to the inequalities (2.5).

It follows from this that either $0 < l < R_1$ or $l > R_1 + R_2$. Since the caustic of the boundary Γ lies within Γ, then $l > R_1$ and the single inequality $l > R_1 + R_2$ therefore remains, which guarantees the property that I is hyperbolic with positive multipliers.

In a typical situation, all the critical points of the function $L:\Gamma \to \mathbb{R}$ are non-degenerate. The overall number of such points is then even and the maximum and minimum points alternate with one another.

Example. We will now again consider an elliptic billiard with semi-axes $a \geqslant b$. We fix the value of a and decrease b, starting with the value $b = a$. While $a^2 < 2b^2$, the function L will have four critical points at the vertices of the ellipse (at the ends of its axes) and, moreover, the ends of the major axis are the maxima of the function L, and the ends of the minor axis are the minima. When $a^2 = 2b^2$, the last two critical points are degenerate and, when b is reduced further, they will pass into local maxima. Since, when $a^2 > 2b^2$, the caustic intersects the ellipse at four points (Fig. 1), four non-degenerate local maxima of the function L appear between the vertices. The law of change of the extremum type still holds, but two-link periodic trajectories do not correspond to these new stationary points (see Assertion 1). This example also shows that, in conclusion 3 of Assertion 2 the condition that the caustic of the boundary Γ must not intersect Γ cannot be omitted.

3. THE EXISTENCE OF A STABLE TRAJECTORY

The main result is as follows.

Assertion 3. We shall assume that the caustic of the convex oval Γ does not intersect Γ and that the function $L: \Gamma \to \mathbb{R}$ only has non-degenerate stationary points. Then, the billiard within Γ has an even number $m \geqslant 2$ of two-link periodic trajectories and, moreover, half of them are hyperbolic and a half of them are of elliptic type. If the ends of these trajectories are numbered in sequence with the numbers $1, 2, \ldots, 2m$, the ends of the ith trajectory $(1 \leqslant i \leqslant m)$ are the points with numbers i and $m + i$, and hyperbolic (elliptic) trajectories correspond to even values of i and elliptic (hyperbolic) trajectories correspond to odd values of i, respectively.

The proof makes use of Assertion 2 and some simple topological facts. We will first show that, if the curvature x of the closed oval Γ is positive everywhere, then any two segments of two-link trajectories necessarily intersect. Since $x > 0$, then (according to Frenet's formulae), during the motion of the point γ along Γ in a positive (negative) direction, the tangent to Γ at the point γ rotates in the same direction. Suppose γ and γ' are the ends of a two-link trajectory I; the tangents of Γ at these points are obviously parallel. Hence, if the two points γ_1 and $\gamma'_1 \in \Gamma$ lie on one side of I, the tangents at these points necessarily intersect and, consequently, they cannot be the ends of a two-link trajectory.

We now consider the trajectory I_1 of maximum length. Since the caustic of Γ lies within Γ, $2a_1 l - 1 > 0$ in formula (2.1). Consequently, the ends of I_1 are non-degenerate local maxima of the function of length L. By assumption, all the critical points of the function L are non-degenerate. This means that other critical points of L necessarily exist and, by Assertion 1, the additional two-link trajectories $I_2, \ldots, I_m$. All of them intersect I_1 and therefore, on each side of I_1 in Γ, there are $m - 1$ different critical points of the function L. By virtue of the alternation of the maxima and minima, their number must be odd. Consequently, m, the number of different two-link trajectories, is even.

In the oval Γ, we label the ends of the segments $I_1, I_2, \ldots, I_m$ in sequence with the numbers $1, 2, \ldots, 2m$. Since these segments intersect in a pairwise manner, the pairs of points with the numbers i and $m + i$ $(1 \leqslant i \leqslant m)$ turn out to be the ends of these segments. The ends of the longest segment I_1 are maxima of the function L. On taking account of the alternation of the type of extremum of the function L, we obtain that the ends of each of the segments I_j are either the two maxima or the two minima of L. By Assertion 2, the trajectory is hyperbolic in the first case and elliptic in the second, which it was required to show.

4. CONVEX BILLIARDS WITHOUT STABLE TWO-LINK TRAJECTORIES

We will now present an example of a convex curve Γ with a caustic which does not wholly lie within Γ and for which all of the two-link trajectories are hyperbolic. For this purpose, we take an equilateral triangle and consider three circles S_1, S_2, S_3 with their centres at the vertices of the triangle and with radii which (to be specific) are equal to one-third of the length of a side. Suppose Γ is the boundary of a convex shell drawn on the set $S_1 \cup S_2 \cup S_3$ (Fig. 3). This curve consists of arcs of circles and straight line segments. The billiard within Γ has exactly six two-link trajectories: three segments of maximum length and a further three segments on the ends of which the function of length L takes maximum and minimum values. All of these trajectories are hyperbolic. It is true that, in the case of the curve Γ, the curvature is not positive everywhere. However, this curve can be transformed into a strictly convex (even

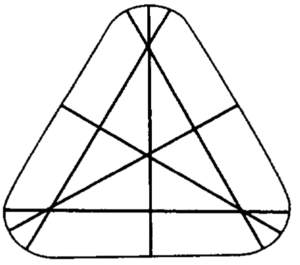

Fig. 3

analytic) oval, by means of a deformation which may be as small as desired, which again will have exactly six hyperbolic two-link trajectories.

This research was supported financially by the Russian Foundation for Basic Research (99-01-01096) and the "Universities of Russia" program (5581).

REFERENCES

1. BIRKHOFF, G. D. *Dynamical Systems*. American Mathematical Society, New York, 1927.
2. BABICH, V. M. and BULDYREV V. S., *Asymptotic Methods in Problems of the Diffraction of Short Waves*. Nauka, Moscow, 1972.
3. MARKEYEV, A. P., On mappings which conserve an area and their application in the dynamics of systems with collisions. *Izv. Ross. Akad. Nauk MTT*, 1996, 2, 37–54.
4. KOZLOV, V. V., A constructive method for substantiating the theory of systems with unilateral constraints. *Prikl. Mat. Mekh.*, 1998, **52**, 6, 883–894.
5. TRESHCHEV, D. V., The problem of the stability of the periodic trajectories of a Birkhoff billiard. *Vestnik Mosk. Gos. Univ. Series 1. Mathematics, Mechanics*, 1988, 2, 44–50.

Translated by E.L.S.