J. DIFFERENTIAL GEOMETRY
30 (1989) 191-205

PERIODIC POINTS OF THE BILLIARD BALL MAP IN A CONVEX DOMAIN

MAREK RYSZARD RYCHLIK

Abstract

We study periodic orbits of the billiard ball map in a strictly convex domain with a C^∞ boundary. We conjecture that the Lebesgue measure of all periodic points is 0. We are able to prove the following partial result: the measure of period three periodic orbits is 0.

The question of whether the mentioned measure is 0 appeared in the study of spectral invariants of a planar region. The author learned about it from R. Melrose.

0. Introduction

In this paper we study the Lebesgue measure of the set of periodic points of a billiard ball map in a strictly convex region with a smooth boundary. The question of whether this measure is 0 has some significance in the theory of spectral invariants of a planar region (cf. [3]). We learned about the problem from R. Melrose. We notice that if the measure is 0 then the length spectrum of a billiard in a convex domain has measure 0.

In spite of our attempts to give a complete solution, we have only been able to settle the simplest case of period three. (The reader will notice that the case of period two is trivial since in this case the reflection has to happen under the right angle with the boundary.) The case of higher periods seems to be much more difficult for more or less the same reasons as verifying that a critical point of a multi-variable function is an extreme point in the case when the second derivative test fails. A solution along the same lines as in this paper seems likely to be related to the singularity theory of functions.

In our solution we use the length function $\mathscr{L}_n$ (see §1). It is the same function that Birkhoff used in his proof of the existence of many periodic orbits (cf. [3]).

Received June 26, 1987 and, in revised form, August 23, 1988. The author was supported in part by National Science Foundation Grant DMS-8701789.

Question. Do almost all critical points of $\mathscr{L}_n$ satisfy the Milnor condition (cf. [2])?

In our opinion the problem and our result are interesting because they distinguish a billiard ball map from a general twist map. From this point of view our result is an exercise in using properties of the billiard ball map other than the twist condition and symplecticity alone. (The reader is encouraged to find his own example of a twist map which has a disk of periodic points of period three.)

In the final stage of our argument we ran into a lengthy computation. The author chose to apply MACSYMA[1] to do the job of differentiating and simplifying the expressions to a form which is used to draw the final conclusions. We took a reasonable amount of trouble to verify that the expressions after the simplification were equivalent to the ones we started with. The computations can be done by hand, but we consider the computer a more reliable tool for carrying them out. Besides, the computations do not seem very enlightening.

Probably it is more important to say that the final product of our manipulations has a surprisingly simple form. It could mean that our approach is not the best and the whole proof could be done differently.

Perhaps the main idea which led us towards the final computation has been lost in technicalities. The idea is that the class of billiard ball maps possesses some amount of transversality amongst all symplectic twist maps of a cylinder. A careful reader will also notice some similarity of conditions (2.22) to the Frobenius condition of integrability of a two-dimensional distribution in $\mathbb{R}^3$. In fact, it is possible to formulate (2.22) as a Frobenius condition.

We would also like to point out that our result generalizes to billiards which are not strictly convex, with piecewise C^3 boundaries, where obvious care has to be taken of the occurring singularities.

The author would like to thank A. Katok for his interest in the result. This was an invaluable incentive to complete the work on this version of our paper.

1. Preliminaries

Suppose that D is a strictly convex region in the plane $\mathbb{R}^2$. We assume that the boundary ∂D is a C^3 curve with positive curvature.

[1]MACSYMA is the largest symbolic computation system developed at MIT and is a trademark of Symbolics, Inc.

Throughout the paper we use the notion of a *directed angle* between two nonzero vectors $\mathbf{v}_1$ and $\mathbf{v}_2$. The directed angle is a class of numbers modulo 2π. A number θ belongs to the class $\angle(\mathbf{v}_1, \mathbf{v}_2)$ iff there is $\lambda > 0$ such that $\mathbf{v}_2 = \lambda e(\theta)\mathbf{v}_1$, where $e(\theta)$ is the rotation of the plane by θ given by the matrix:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}. \tag{1.1}$$

We pick the counterclockwise orientation of ∂D and by $\mathbf{T}(P)$ we denote the unit tangent vector to ∂D at $P \in \partial D$, pointing in the counterclockwise direction.

The *billiard ball map* β is defined on the set of pairs $(P, \mathbf{v})$, where $P \in \partial D$ and $\mathbf{v}$ is a unit vector on the plane such that $\angle(\mathbf{T}(P), \mathbf{v}) \cap [0, \pi] \neq \emptyset$. This subset of $\mathbb{R}^4 = \mathbb{R}^2 \times \mathbb{R}^2$ is often denoted by $S_+(\partial D)$, as it may be considered a part of the unit circle bundle of ∂D, often denoted $S(\partial D)$.

The value of β on the pair $(P, \mathbf{v})$ is another pair $(Q, \mathbf{w})$, where Q and $\mathbf{w}$ are uniquely determined from the following conditions (see Figure 1):

(i) $Q = P + t\mathbf{v}$, where $t \geq 0$;

(ii) $\mathbf{w}$ is the reflection of $\mathbf{v}$ through the direction of $\mathbf{T}(Q)$, or

$$\angle(\mathbf{T}(Q), \mathbf{v}) = -\angle(\mathbf{T}(Q), \mathbf{w}). \tag{1.2}$$

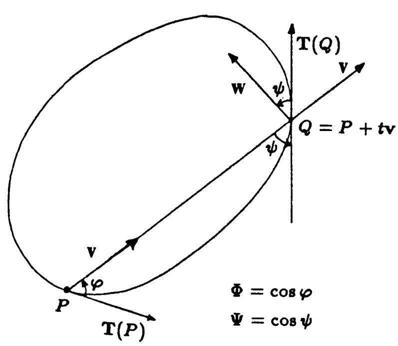

FIGURE 1

It is customary to use the following description of β employing coordinates. Let $h\colon \mathbb{R} \to \mathbb{R}^2$ be a parametrization of ∂D by length, i.e., $\|h'(t)\| = 1$

for every $t \in \mathbb{R}$, and $h'(t) = \mathbf{T}(h(t))$. Clearly, h is L-periodic, where L is the length of ∂D. The *curvature function* $k: \mathbb{R} \to \mathbb{R}_+$ is defined through the following equation:

$$h''(t) = k(t) \cdot J \cdot h'(t), \tag{1.3}$$

where $J = e(\pi/2)$, i.e. $J = \left[\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$. We will always assume that $k(t) > 0$ for all t.

The *length function* $l: \mathbb{R}^2 \to \mathbb{R}_+$ is defined as

$$l(x,y) = \|h(y) - h(x)\|. \tag{1.4}$$

We introduce a map $p: \mathbb{R} \times [-1,1] \to S_+(\partial D)$, which maps the pair (x, Φ) to $(h(x), \mathbf{v})$, where $\mathbf{v}$ is determined from the condition

$$\cos \angle(h'(x), \mathbf{v}) = \Phi. \tag{1.5}$$

Obviously, p is a covering of $S_+(\partial D)$ and $p(x+L, \Phi) = p(x, \Phi)$. This yields a natural identification of $S_+(\partial D)$ with a cylinder. (The reader will notice that the natural orientation is reversed, though.)

It is well known that the lifting of β to $\mathbb{R} \times [-1,1]$ can be described purely in terms of the length function l. We denote the lifting also by β. This new β maps (x, Φ) to (y, Ψ), where y and Ψ should be determined from the equations

$$-\Phi = \frac{\partial l(x,y)}{\partial x}, \qquad \Psi = \frac{\partial l(x,y)}{\partial y}. \tag{1.6}$$

One can show that the first equation can always be solved for y in terms of x and $\Phi \in [-1,1]$. Then we determine Ψ from the second equation. A reader familiar with hamiltonian mechanics (in the scope of [1], for instance) will recognize l as a *generating function* of β. The map β is symplectic, i.e., it preserves the 2-form $d\Phi \wedge dx$. Indeed, $dl = -\Phi\, dx + \Psi\, dy$, so $0 = d(dl) = -d\Phi \wedge dx + d\Psi \wedge dy$. This is the desired invariance. In particular, β is area preserving. We notice that β is not differentiable at the boundary $\mathbb{R} \times \{-1, 1\}$.

Let us fix $a \in \mathbb{R}$. One can easily verify that h' and h can be expressed in terms of k. In fact, we have

$$h'(t) = e\left(\int_a^t k(\tau)\, d\tau\right) \cdot h'(a), \tag{1.7a}$$

$$h(x) = h(a) + \int_a^x e\left(\int_a^t k(\tau)\, d\tau\right) \cdot h'(a). \tag{1.7b}$$

From (1.7b) we can easily find that

$$l(x,y) = \left(\int_x^y \int_x^y \cos\left(\int_s^t k(\tau)\, d\tau\right) ds\, dt\right)^{1/2}. \tag{1.8}$$

This formula allows one to compute l, given the curvature function. However, later we will derive a differential equation connecting l and k, which seems to be of greater value for our purposes.

Let $\mathscr{L}_n\colon \mathbb{R}^n \to \mathbb{R}_+$ be defined by

$$\mathscr{L}_n(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} l(x_1, x_{i+1}). \tag{1.9}$$

(We tacitly assume that $n+1=1$.)

There is a well-known connection between periodic points of β of period n and the critical points of $\mathscr{L}_n$. We introduce the following definitions (see (1.6)):

$$-\Phi_i = \frac{\partial l(x_i, x_{i+1})}{\partial x_i}, \qquad \Psi_i = \frac{\partial l(x_{i-1}, x_i)}{\partial x_i}. \tag{1.10}$$

Using the above notation we can write

$$\frac{\partial \mathscr{L}_n}{\partial x_i} = -\Phi_i + \Psi_i. \tag{1.11}$$

Now it is clear that $\{(x_i, \Phi_i)\}_{i=1}^n$ is a periodic trajectory of β (or, more precisely, it projects down to a periodic trajectory) iff $\Phi_i = \Psi_i$ for $i = 1, 2, \cdots, n$. Now we are able to state our main results. For the sake of clarity of our presentation, we will spend most of the time proving

Theorem 1.1. *For $n=3$ the set of periodic orbits of β is nowhere dense, i.e., it has empty interior.*

In §3 we will indicate how one can enhance the above result to obtain

Theorem 1.2. *For $n=3$ the Lebesgue measure of the set of periodic orbits of β has Lebesgue measure* 0.

In the sequel $\mathrm{Fix}_n \subset \mathbb{R}^2$ means the set of all periodic points of the billiard ball map of period n, and $\mathrm{Crit}_n \subset \mathbb{R}^n$ denotes the set of all critical points of $\mathscr{L}_n$. We adopted the notation $d^2\mathscr{L}_n(\mathbf{x})$ for the Hessian matrix of the function $\mathscr{L}_n$ evaluated at $\mathbf{x} \in \mathbb{R}^n$.

2. Some computations with billiards

Let Φ and Ψ be as in (1.6) and let us define $\hat{\Phi} = \sqrt{1-\Phi^2}$ and $\hat{\Psi} = \sqrt{1-\Psi^2}$.

Lemma 2.1. *If $l = l(x,y)$, then*

$$d^2 l(x,y) = \begin{bmatrix} \hat{\Phi}^2/l - k(x)\hat{\Phi} & \hat{\Phi}\hat{\Psi}/l \\ \hat{\Psi}\hat{\Phi}/l & \hat{\Psi}^2/l - k(y)\hat{\Psi} \end{bmatrix}. \tag{2.1}$$

Proof. This formula can be derived by geometric differentiation, and we recommend it as an exercise (see Figure 2). Here we present a different proof based on (1.8). From (1.7b) it follows that

$$h(y) - h(x) = \int_x^y e\left(\int_x^t k(\tau)\,d\tau\right) dt \cdot h'(x). \tag{2.2}$$

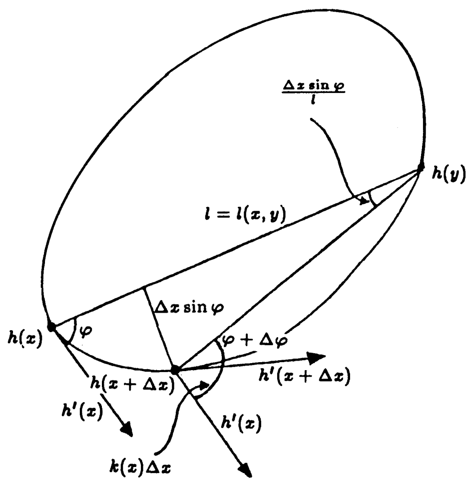

FIGURE 2. We can easily verify that the following formulas hold:

$$\Delta\varphi = \frac{\Delta x \cdot \sin\varphi}{l} - k(x)\Delta x, \quad \frac{\partial\varphi}{\partial x} = \frac{\sin\varphi}{l} - k(x), \quad \Phi = \cos\varphi,$$
$$\frac{\partial\Phi}{\partial x} = -\sin\varphi\frac{\partial\varphi}{\partial x} = -\frac{\sin^2\varphi}{l} + k(x)\sin\varphi = -\frac{\hat{\Phi}^2}{l} + k(x)\hat{\Phi}.$$

Hence the matrix of the rotation that carries $h'(x)$ to $h(y) - h(x)$ is

$$\begin{bmatrix} \Phi & -\hat{\Phi} \\ \hat{\Phi} & \Phi \end{bmatrix} = \frac{1}{l}\int_x^y e\left(\int_x^t k(\tau)\,d\tau\right) dt. \tag{2.3}$$

This implies that

$$\Phi = \frac{1}{l}\int_x^y \cos\left(\int_x^t k(\tau)\,d\tau\right) dt, \tag{2.3a}$$

and similarly

$$\hat{\Phi} = \frac{1}{l}\int_x^y \sin\left(\int_x^t k(\tau\, d\tau)\right) dt. \tag{2.3b}$$

In a similar way we obtain

$$\Psi = \frac{1}{l}\int_x^y \cos\left(\int_t^y k(\tau)\, d\tau\right) dt, \tag{2.4a}$$

$$\hat{\Psi} = \frac{1}{l}\int_t^y \sin\left(\int_t^y k(\tau)\, d\tau\right) dt. \tag{2.4b}$$

By differentiating (2.3a) we get

$$\begin{aligned}\frac{\partial \Phi}{\partial x} &= -\frac{1}{l^2}\frac{\partial l}{\partial x}\int_x^y \cos\left(\int_x^t k(\tau)\, d\tau\right) dt \\ &\quad + \frac{1}{l}\left(-1 + \int_x^y \sin\left(\int_x^t k(\tau)\, d\tau\right) dt \cdot k(x)\right) \\ &= -\frac{1-\Phi^2}{l} + k(x)\hat{\Phi} = -\frac{\hat{\Phi}^2}{l} + k(x)\hat{\Phi}.\end{aligned} \tag{2.5}$$

In a similar fashion we derive the following formulas:

$$\frac{\partial \Psi}{\partial y} = \frac{\hat{\Psi}^2}{l} - k(y)\hat{\Psi}, \quad \frac{\partial \Phi}{\partial y} = -\frac{\hat{\Phi}\hat{\Psi}}{l}, \quad \frac{\partial \Psi}{\partial x} = \frac{\hat{\Phi}\hat{\Psi}}{l}. \tag{2.6}$$

Now a part of the lemma follows from the observation that

$$\frac{\partial \Phi}{\partial x} = \frac{\partial}{\partial x}\left(-\frac{\partial l}{\partial x}\right) = -\frac{\partial^2 l}{\partial x^2}.$$

Similar relations are true for other partials of l.

Lemma 2.2. *In the notation of Lemma* 1.1 *we have*

$$\Phi\Psi - \hat{\Phi}\hat{\Psi} = \cos\left(\int_x^y k(\tau)\, d\tau\right), \qquad \Phi\hat{\Psi} + \hat{\Phi}\Psi = \sin\left(\int_x^y k(\tau)\, d\tau\right). \tag{2.7}$$

Remark 2.2. Geometrically (2.7) means that

$$\angle(\mathbf{T}(P), \mathbf{T}(Q)) = \angle(\mathbf{T}(P), \overrightarrow{PQ}) + \angle(\overrightarrow{PQ}, \mathbf{T}(Q)). \tag{2.8}$$

(Also see Figure 3.)

Proof. We will only supply a proof of the first of the two equalities, since the second one is very similar. It can be obtained by the following

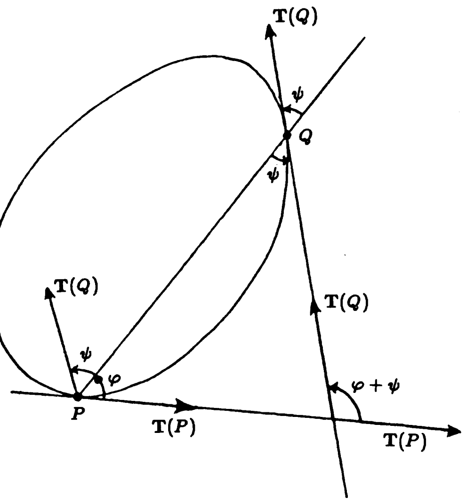

FIGURE 3. It is easy to verify that $\varphi + \psi = \int_x^y k(\tau)\,d\tau$.

computation:
(2.9)

$$\begin{aligned} l^2(\Phi\Psi - \hat{\Phi}\hat{\Psi}) &= \int_x^y \int_x^y \cos\left(\int_x^t k(\tau)\,d\tau + \int_s^y k(\tau)\,d\tau\right) dt\,ds \\ &= \int_x^y \int_x^y \cos\left(\int_x^y k(\tau)\,d\tau + \int_s^t k(\tau)\,d\tau\right) dt\,ds \\ &= \int_x^y \int_x^y \cos\left(\int_x^y h(\tau)\,d\tau\right) \cos\left(\int_s^t k(\tau)\,d\tau\right) ds\,dt \\ &\quad - \int_x^y \int_x^y \sin\left(\int_x^y k(\tau)\,d\tau\right) \sin\left(\int_s^t k(\tau)\,d\tau\right) ds\,dt. \end{aligned}$$

The second term is 0 by symmetry, and the first one is

$$\cos\left(\int_x^y k(\tau)\,d\tau\right) \cdot \int_x^y \int_x^y k \cos\left(\int_s^t k(\tau)\,d\tau\right) ds\,dt. \tag{2.10}$$

By comparison of formulas (2.10), (1.8), and (2.9) we get the lemma.

Lemma 2.3. *If* $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ *is a critical point of* $\mathscr{L}_n$, *then*

$$d^2\mathscr{L}_n(\mathbf{x}) = \begin{bmatrix} A_1 & B_1 & 0 & \cdots & 0 & B_n \\ B_1 & A_2 & B_2 & \cdots & 0 & 0 \\ 0 & B_2 & A_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{n-1} & B_{n-1} \\ B_n & 0 & 0 & \cdots & B_{n-1} & A_n \end{bmatrix}, \tag{2.11}$$

where

$$A_i = \hat{\Phi}_i^2 \left(\frac{1}{l_{i-1,i}} + \frac{1}{l_{i,i+1}}\right) - 2k(x_i)\hat{\Phi}_i, \quad B_i = \frac{\hat{\Phi}_i \hat{\Phi}_{i+1}}{l_{i,i+1}}, \tag{2.12}$$
$$l_{i,i+1} = l(x_i, x_{i+1}).$$

Proof. The proof follows immediately from Lemma 2.2 and the definitions of §1.

Proposition 2.1. *If* $\mathbf{x} = (x_1, x_2, x_3)$ *is a critical point of* $\mathscr{L}_3$ *and* $\operatorname{rank} d^2\mathscr{L}_3(\mathbf{x}) = 1$, *then*

$$\begin{aligned} k(x_1) &= \frac{1}{2}\left(\frac{1}{l_3} + \frac{1}{l_2} - \frac{l_1}{l_2 l_3}\right)\hat{\Phi}_1, \\ k(x_2) &= \frac{1}{2}\left(\frac{1}{l_1} + \frac{1}{l_3} - \frac{l_2}{l_1 l_3}\right)\hat{\Phi}_2, \\ k(x_3) &= \frac{1}{2}\left(\frac{1}{l_1} + \frac{1}{l_2} - \frac{l_3}{l_1 l_2}\right)\hat{\Phi}_3, \end{aligned} \tag{2.13}$$

where we have used the following abbreviations:

$$l_1 = l(x_2, x_3), \quad l_2 = l(x_3, x_1), \quad l_3 = l(x_1, x_2).$$

Proof. Indeed, by Lemma 2.3 we have

$$d^2\mathscr{L}_3(\mathbf{x}) = \begin{bmatrix} A_1 & B_1 & B_3 \\ B_1 & A_2 & B_2 \\ B_3 & B_2 & A_3 \end{bmatrix}. \tag{2.14}$$

If $\operatorname{rank} d^2\mathscr{L}_3(\mathbf{x}) = 1$, then all 2×2 minors of $d^2\mathscr{L}_3(\mathbf{x})$ must vanish. For instance $A_1 B_2 = B_1 B_3$, or

$$\left(\hat{\Phi}_1^2 \left(\frac{1}{l_2} + \frac{1}{l_3}\right) - 2k(x_1)\hat{\Phi}_1\right) \frac{\hat{\Phi}_2 \hat{\Phi}_3}{l_1} = \frac{\hat{\Phi}_1 \hat{\Phi}_2}{l_3} \cdot \frac{\hat{\Phi}_3 \hat{\Phi}_1}{l_2}. \tag{2.15}$$

This implies the first of the formulas (2.13). We can prove the remaining two formulas by considering other minors of $d^2\mathscr{L}_3(\mathbf{x})$. It is easy to see that equations (2.13) imply that the rank of $d^2\mathscr{L}_3(\mathbf{x})$ is equal to one.

Proposition 2.2. *If equations* (2.13) *are satisfied, then the nullspace of* $d^2\mathscr{L}_3(\mathbf{x})$ *is given by the equation*

$$\hat{\Phi}_1\,dx_1 + \hat{\Phi}_2\,dx_2 + \hat{\Phi}_3\,dx_3 = 0. \tag{2.16}$$

Proof. We have $A_1 = B_1B_3/B_2$ from the previous proof. This implies that the first column of $d^2\mathscr{L}_3(\mathbf{x})$ is orthogonal to the vector $(\hat{\Phi}_1, \hat{\Phi}_2, \hat{\Phi}_3)$.

Lemma 2.4. *If* $\mathbf{x} = (x_1, x_2, x_3)$ *is critical for* $\mathscr{L}_3$, *then* (*see Figure* 4)

$$\begin{aligned}
\hat{\Phi}_1 &= \frac{1}{2}\sqrt{\frac{(l_1+l_2+l_3)(l_2+l_3-l_1)}{l_2l_3}},\\
\hat{\Phi}_2 &= \frac{1}{2}\sqrt{\frac{(l_1+l_2+l_3)(l_1+l_3-l_2)}{l_1l_3}},\\
\hat{\Phi}_3 &= \frac{1}{2}\sqrt{\frac{(l_1+l_2+l_3)(l_1+l_2-l_3)}{l_1l_2}},\\
\Phi_1 &= \pm\frac{1}{2}\sqrt{\frac{(l_1+l_3-l_2)(l_1+l_2-l_3)}{l_2l_3}}\\
\Phi_2 &= \pm\frac{1}{2}\sqrt{\frac{(l_2+l_3-l_1)(l_2+l_1-l_3)}{l_1l_3}},\\
\Phi_3 &= \pm\frac{1}{2}\sqrt{\frac{(l_3+l_2-l_1)(l_1+l_3-l_2)}{l_1l_2}}.
\end{aligned} \tag{2.17}$$

(*The second and third pairs of these equations were obtained from the first by a cyclic permutation of the indices.*) *Moreover, in those equations where the choice of a sign is possible, the sign has to be the same, i.e., one has either three pluses or three minuses.*

Proof. Let $P_i = h(x_i)$ and let α_i be the angle of the triangle $P_1P_2P_3$ at the vertex P_i ($i = 1, 2, 3$). Let φ_i be the angle of reflection at P_i. The situation is drawn on Figure 2. It is easy to see that $2\varphi_i = \pi - \alpha_i$. Let us derive the third pair of our equations. By the Cosine Theorem

$$\cos\alpha_3 = \frac{l_1^2 + l_2^2 - l_3^2}{2l_1l_2}. \tag{2.18}$$

We also have the following trivial equalities:

$$\cos(\pi - \alpha_3) = -\cos\alpha_3 = \cos(2\varphi_3) = 1 - 2\sin^2\varphi_3 = 1 - 2\hat{\Phi}_3^2. \tag{2.19a}$$

Similarly, we can write

$$-\cos\alpha_3 = 2\cos^2\varphi_3 - 1 = 2\Phi_3^2 - 1. \tag{2.19b}$$

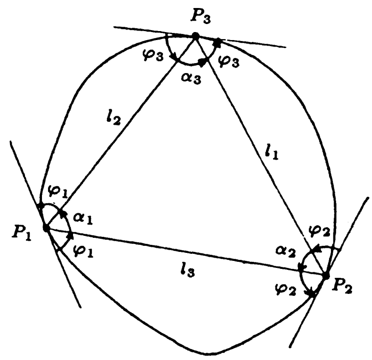

FIGURE 4. We have the following relations:

$$\begin{aligned} \Phi_1 &= \cos\varphi_1, & \hat{\Phi}_1 &= \sin\varphi_1, \\ \Phi_2 &= \cos\varphi_2, & \hat{\Phi}_2 &= \sin\varphi_2, \\ \Phi_3 &= \cos\varphi_3, & \hat{\Phi}_3 &= \sin\varphi_3. \end{aligned}$$

These equations solved for $\hat{\Phi}_3$ and Φ_3 are the desired formulas.

Proof of Theorem 1.1. Suppose that Fix_3 has a nonempty interior. Then there is a two-dimensional manifold $N \subseteq \mathbb{R}^3$ consisting of critical points of $\mathscr{L}_3$. It is easy to see that if $\mathbf{x} \in N$, then $\operatorname{rank} d^2\mathscr{L}_3(\mathbf{x}) = 1$. By Proposition 2.1 equations (2.13) are satisfied on that manifold. We will differentiate equations (2.13) to get a contradiction. Let us rewrite these equations in the following form:

$$k(x_i) = f_i(l_1, l_2, l_3) \qquad (i = 1, 2, 3), \tag{2.20}$$

where $f_i(l_1, l_2, l_3)$ is the right-hand side of the ith of formulas (2.13) with $\hat{\Phi}_i$ replaced by the ith of expressions (2.17). We can concentrate on the first equation only, due to symmetry. Clearly, the equation for the tangent space $T_{\mathbf{x}}N$ is given by (2.16). In that plane there is a direction satisfying $dx_1 = 0$, namely $(0, -\hat{\Phi}_3, \hat{\Phi}_2)$. Differentiating (2.20) in that direction we get

$$0 = \left(\frac{\partial}{\partial x_2} f_1(l_1, l_2, l_3)\right) \cdot (-\hat{\Phi}_3) + \left(\frac{\partial}{\partial x_3} f_1(l_1, l_2, l_3)\right) \cdot \hat{\Phi}_2. \tag{2.21}$$

Since the right-hand side of this expression allows a representation in terms of l_1, l_2 and l_3, we obtain an equation which must be satisfied by any vector (l_1, l_2, l_3) corresponding to a point $\mathbf{x} \in N$.

Summarizing the last argument, we get a system of three equations for three "unknowns" l_1, l_2 and l_3:

$$\begin{aligned} \frac{\partial}{\partial x_2} f_1(l_1, l_2, l_3) \cdot \hat{\Phi}_3 &= \frac{\partial}{\partial x_3} f_1(l_1, l_2, l_3) \cdot \hat{\Phi}_2, \\ \frac{\partial}{\partial x_1} f_2(l_1, l_2, l_3) \cdot \hat{\Phi}_3 &= \frac{\partial}{\partial x_3} f_2(l_1, l_2, l_3) \cdot \hat{\Phi}_1, \\ \frac{\partial}{\partial x_1} f_3(l_1, l_2, l_3) \cdot \hat{\Phi}_2 &= \frac{\partial}{\partial x_2} f_3(l_1, l_2, l_3) \cdot \hat{\Phi}_1. \end{aligned} \tag{2.22}$$

We examined this system of equations with MACSYMA and showed that it has no solution (l_1, l_2, l_3) amongst triples of numbers representing sides of a triangle. The details of the computation are presented in the Appendix.

3. The proof of Theorem 1.2

The proof of Theorem 1.2 can be obtained by making a few adjustments in the proof of Theorem 1.1. One needs to verify that the differentiation argument we used to derive (2.22) can still be applied under the weaker assumption that Fix_3 has positive Lebesgue measure.

This is how it can be done. Let us introduce a map $q: \mathbb{R}^2 \to \mathbb{R} \times [-1, 1]$ via the rule $(x, y) \mapsto (x, \Phi)$, where Φ is determined from (1.6). It is easy to see that if Fix_3 has positive measure, then $q^{-1}(\mathrm{Fix}_3)$ also does.

We pick a point $\mathbf{x}_0 \in \mathrm{Crit}_3$ such that its projection onto the (x_1, x_2)-plane is a Lebesgue density point of $q^{-1}(\mathrm{Fix}_3)$. Then we show that any direction in the plane given by (2.16) at $\mathbf{x}_0$ can be approximated by a sequence of vectors of the form $(\mathbf{x}_m - \mathbf{x}_0)/\|\mathbf{x}_m - \mathbf{x}_0\|$, where $\mathbf{x}_m \in \mathrm{Crit}_3$, $\mathbf{x}_m \neq \mathbf{x}_0$ and $\mathbf{x}_m \to \mathbf{x}_0$ as $m \to \infty$. Indeed, if that was not the case, then one could find a cone at $\mathbf{x}_0$ disjoint with Crit_3 near $\mathbf{x}_0$. By definition, Crit_3 consists of the critical points of $\mathscr{L}_3$. Hence the sequence $\mathbf{x}_m$ must approach the plane (2.16) faster than it approaches $\mathbf{x}_0$. In fact, if $\mathscr{L}_3$ is of class C^3, then one can show that the distance of $\mathbf{x}_m$ from the plane is not greater than $\mathrm{const} \cdot \|\mathbf{x}_m - \mathbf{x}_0\|^{3/2}$. This is enough to find a cone in the (x_1, x_2)-plane disjoint with $q^{-1}(\mathrm{Fix}_3)$, which contradicts the definition of $\mathbf{x}_0$. The situation is sketched in Figure 5. The reader will easily fill in the details.

Appendix. Symbolic computation with MACSYMA

We used MACSYMA to bring the system of equations (2.22) to a manageable form. A glance at the expressions involved should convince the

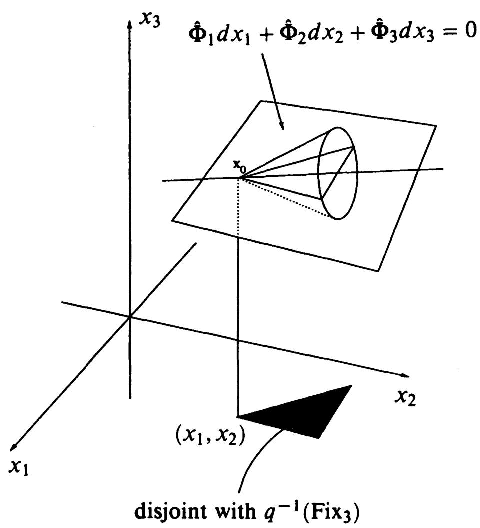

FIGURE 5

reader that the formulas obtained after performing the differentiations are a mess! As a matter of fact, they occupied about one and a half of a standard terminal screenful each. We applied MACSYMA's simplifier called RADCAN (prepared to deal with expressions containing lots of square roots) and after a computation, which lasted several minutes, we obtained a simplified version of the first of the equations. It clearly was a rational function, so we applied the FACTOR command to get the factorization over the integers. Here is the final equation:

$$\begin{aligned}&\frac{-3(l_3 - l_2 - l_1)(l_3 - l_2 + l_1)(l_3 + l_2 - l_1)(l_3 + l_2 + l_1)}{32 l_1 l_2^2 l_3^3} \\ &\quad = \frac{3(l_3 - l_2 - l_1)(l_3 - l_2 + l_1)(l_3 + l_2 - l_1)(l_3 + l_2 + l_1)}{32 l_1 l_2^3 l_3^2}.\end{aligned} \tag{A.1}$$

The other two equations can be obtained from the above one by a cyclic change of indices in variables l_1, l_2 and l_3.

This relatively simple form of (A.1) came as a surprise and, full of suspicion, the author started comparing numerical values of the initial complicated form of (A.1) with the final form that the reader is looking at. We plugged in several randomly chosen values of l's (sides of triangles). This finally convinced us of the correctness of the equation (A.1). (The author has to admit that he tried different methods of completing the computation and one based on the RATSIMP command produced results not necessarily equivalent to the ones we started with, apparently due to the abuse of the "identity" $\sqrt{a}\cdot\sqrt{b}=\sqrt{ab}$.)

A closer look at (A.1) leads to the following conclusions:

(i) If $l_1, l_2, l_3 > 0$ and (A.1) holds then both sides are equal to 0 (otherwise they would have opposite signs);

(ii) Therefore one of the numbers l_1, l_2, l_3 would have to be equal to the sum of the others, which cannot happen if they are sides of a triangle.

We provide the reader with the listing of the MACSYMA code which can be used to generate equation (A.1):

```
kill (all);
/*This reinitializes macsyma; any previous formulas will be
lost.*/
writefile (''session_record'');
/*The record of the session will be put in the named file.*/
/*The next six lines correspond exactly to formulas (2.17).*/
phi_1_roof:  (1/2) * sqrt((l1 + l2 + l3) * (l2 + l3 - l1)/(l2 * l3));
phi_2_roof:  (1/2) * sqrt((l1 + l2 + l3) * (l1 + l3 - l2)/(l1 * l3));
phi_3_roof:  (1/2) * sqrt((l1 + l2 + l3) * (l1 + l2 - l3)/(l1 * l2));
phi_1:  (1/2) * sqrt((l1 + l3 - l2) * (l1 + l2 - l3)/(l2 * l3));
phi_2:  (1/2) * sqrt((l2 + l3 - l1) * (l2 + l1 - l3)/(l1 * l3));
phi_3:  (1/2) * sqrt((l3 + l2 - l1) * (l1 + l3 - l2)/(l1 * l2));
/*Declaration of dependencies, see Proposition 2.1.*/
depends (l1, [x2, x3], l2, [x3, x1], 13, [x1, x2]);
/*In the next six lines we define the partials of l1, l2, l3*/
/*with respect to the named variables, using formulas (1.6).*/
gradef(l1, x2, -phi_2);
gradef(l1, x3, phi_3);
gradef(l2, x3, -phi_3);
gradef(l2, x1, phi_1);
gradef(l3, x1, -phi_1);
gradef(l3, x2, phi_2);
/*The next formula is the right-hand side of the first*/
/*of the formulas (2.13).*/
f1: (1/2) * (1/l3 + 1/l2 - l1)/(l2 * l3)) * phi_1_roof;
/*This is the first of equations (2.22).*/
equation:  diff(f1, x2) * phi_3_roof = diff(f1, x3) * phi_2_roof;
/*We apply radcan to the above equation.*/
answer:  radcan (equation);
/*We factor the result over the integers.*/
```

```
final_answer:  factor(answer);
/*We flush the output to the file named ''session_record''.*/
closefile (''session_record'');
```

Recently, this result of this computation has been confirmed using a different system for symbolic computation (Mathematica).

References

[1] V. I. Arnol'd, *Mathematical methods of classical mechanics*, Springer, New York, 1978.

[2] T. Bröcker & L. Lander, *Differential germs and catastrophes*, Cambridge University Press, Cambridge, 1975.

[3] S. Marvizi & R. Melrose, *Spectral invariants of convex planar regions*, J. Differential Geometry **17** (1982) 475–502.

UNIVERSITY OF WASHINGTON