Classical billiards in magnetic fields

1985 J. Phys. A: Math. Gen. 18 1361

(http://iopscience.iop.org/0305-4470/18/9/019)

J. Phys. A: Math. Gen. **18** (1985) 1361-1378. Printed in Great Britain

Classical billiards in magnetic fields

M Robnik and M V Berry

H H Wills Physics Laboratory, Royal Fort, Tyndall Avenue, Bristol BS8 1TL, UK

Received 29 November 1984

Abstract. A particle moves in circular arcs with Larmor radius R between specular reflections at the smooth convex boundary of a planar region. The dynamics depends on the value of R in relation to the extreme curvature radii $\rho_{\min}$ and $\rho_{\max}$ and the radius R^* of the largest circle that can be inscribed in the boundary. For $R < R^*$ some orbits are complete Larmor circles and constitute an integrable component of the motion; all other orbits bounce repeatedly. For $\rho_{\min} < R < \rho_{\max}$ there are 'flyaway intervals' on the boundary for which glancing orbits are a powerful source of chaos in the map (on the phase cylinder) relating successive bounces; this type of chaos is a characteristic feature of magnetic billiards. For sufficiently large R the simplest closed orbits consist of two arcs associated with diameters of the boundary; their existence and stability can be determined. In several regimes where motion consists of short skips between nearby boundary points (including the strong-field case $R \to 0$), an explicit adiabatic invariant can be found which gives an excellent approximation to the exact invariant curves in these regimes. Computations for a magnetic billiard with elliptic boundary illustrate the theory.

1. Introduction

A classical particle ('billiard ball') with mass m and charge q moves with speed v in a plane region with perfectly reflecting smooth convex boundary; a uniform constant magnetic field with strength B is directed perpendicular to the plane. The resulting orbits consist of a series of arcs of circles with the Larmor radius $R = mv/qB$, connected by specular reflection at the boundary; if R is small enough, some orbits form complete circles entirely inside the boundary. Our purpose here is to investigate the geometry of these orbits and in particular to discover how their regularity or chaos depends on R for a given boundary. This work is thus complementary to conventional classical billiards, where there is no field, so that $R = \infty$ (straight line segments), but where the dependence of orbits on the shape of the boundary is considered (for an elementary review see Berry (1981)).

There are several reasons for studying classical magnetic billiards. Ours is as a prelude to an investigation of the quantum mechanics of the corresponding systems, motivated by the fact that magnetic fields break time-reversal symmetry (see appendix 2) and should therefore give rise to new phenomena in the distribution of level spacings (Berry 1984). Another reason is that magnetic billiards are primitive models for more general sorts of bound systems (such as atoms) in magnetic fields (Ruderman 1975), whose orbits might be easier to understand following our analysis. Furthermore, ergodic properties of magnetic billiards might help in understanding the statistical mechanics of diamagnetism (Peierls 1979). Finally, the motion of charged particles

0305-4470/85/091361+18$02.25

magnetically skipping along a wall (boundary between two magnetic domains) is a problem frequently occurring in astrophysical contexts (Drury 1983).

The central feature of magnetic billiards is that the motion depends on the value of R in relation to the radii of curvature of the boundary. This gives rise to a series of curvature regimes which will be described in § 2; we will emphasise the emergence of chaos associated with glancing trajectories. For a given R the phase space is dominated by the shortest closed orbits; these are particularly important if they consist of just two bounces and the existence and stability of such orbits are studied in § 3. In several limiting regimes, including the strong-field limit $R \to 0$, motion consists of 'short skips', meaning that the distance between successive bounces is small. For such cases, a surprisingly accurate adiabatic invariant will be derived in § 4; this shows that motion becomes more nearly integrable as $R \to 0$. When applied to billiards without magnetic field ($R \to \infty$) it yields as a by-product an analytic description of the boundary-hugging orbits whose existence was proved by Lazutkin (1973) (see also Sinai 1976). Finally, in § 5, we illustrate the preceding theory with numerical calculations for elliptic boundaries, for which the zero-field motion is integrable and the magnetic motion not.

We wish to emphasise that this paper is restricted to magnetic billiards with boundaries that are sufficiently *smooth* (e.g. analytic). For boundaries with discontinuities of slope (e.g. polygons) or curvature (e.g. the stadium), our arguments do not apply, and indeed exploratory computations reveal rather different behaviour, on which we plan to report in a separate paper.

2. Curvature regimes

We use the notation illustrated in figure 1. Orbits are arcs of circles with radius R, and will be considered without loss of generality to gyrate clockwise. Points where orbits hit the boundary are labelled by arc length s, reckoned anticlockwise ('forwards') thus $0 \leqslant s < \mathscr{L}$ where $\mathscr{L}$ is the perimeter of the boundary. Directions of emergence

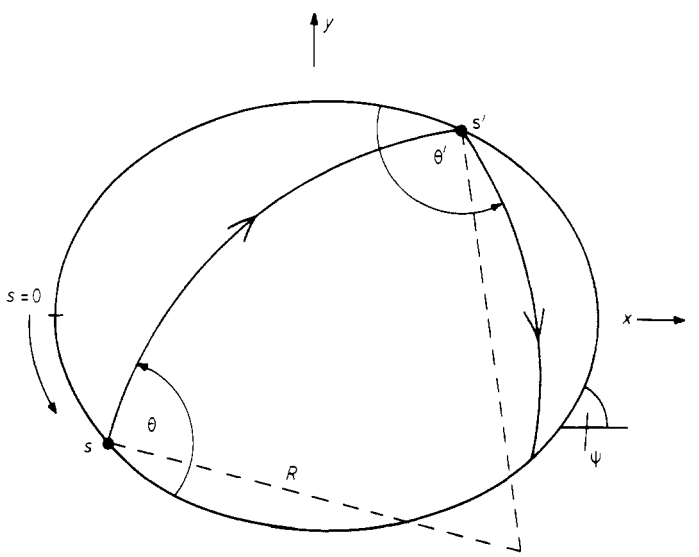


Figure 1. Geometry and coordinates for Larmor orbits in convex billiard.

from a bounce are labelled by the angle θ measured from the forward tangent to the boundary; thus $0 \leq \theta \leq \pi$. Tangential momentum is $p \equiv \cos\theta$, and is a useful quantity because the mapping relating successive bounces on the phase cylinder s, p is area-preserving. The global symmetries of this map are discussed in appendix 2. The boundary has radius of curvature $\rho(s)$ and curvature $\kappa(s) \equiv 1/\rho(s)$.

For $R > R^*$, where R^* is the radius of the largest circle that can be inscribed in the boundary, all orbits bounce repeatedly. For $R < R^*$ some orbits can form complete circles lying entirely within the boundary. The set of all these circles forms an integrable component of the motion for these values of R. R^* must lie between the largest and smallest radii ρ_{max} and ρ_{min} of the boundary, because the largest inscribed circle must touch the boundary at at least two points (figure 2(a)), whereas circles with $R < \rho_{min}$ and $R > \rho_{max}$ may touch at most once. (Generically, the largest inscribed circle touches the boundary at two points (figure 2(a)) or three points (figure 2(b)).)

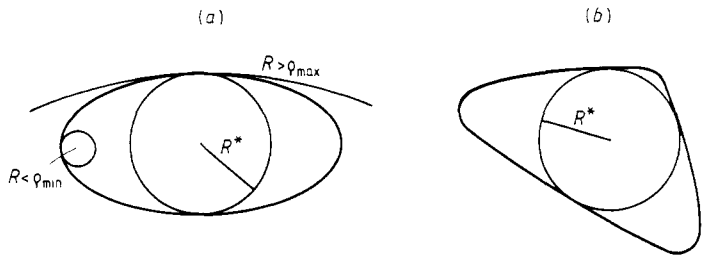


Figure 2. (a) Largest inscribed circle R^* (touching at two points) and circles touching from inside and outside. (b) Largest inscribed circle touching at three points.

To obtain the measure of the integrable component for $R < R^*$, and also to understand the other curvature regimes between $R = \rho_{min}$ and $R = \rho_{max}$, it is useful to introduce the *locus of centres of rolled Larmor circles for radius R*, which we henceforth call the 'R locus'. The different R loci are the involutes of the evolute E of the boundary (i.e. its envelope of normals). (In an optical analogy, the R loci are the successive wavefronts evolving by inwards normal propagation of boundary points, and E is the caustic curve of the corresponding rays. In differential geometry, R loci are called 'parallel curves'. As illustrated in figure 3, the curve E is smooth except for cusp points on the normals from boundary 'vertices' where $\rho(s)$ takes an extreme value; in appendix 1 we show that there must be at least two vertices where ρ is locally minimal and two where it is locally maximal, so that E has at least four cusps. (In degenerate cases, cusps corresponding to different vertices may coincide. An extreme is when E is a deltoid curve, each of whose three cusps is covered twice as the boundary is traversed once; however, the slightest perturbation will recover the generic situation in which E has six cusps and is covered once.)

The R loci are smooth curves for $R < \rho_{min}$ and $R > \rho_{max}$. Otherwise, the R loci have cusps where they meet E at right angles (figure 3). These cusps divide each R locus into a series of segments corresponding to boundary regions where alternately $\rho(s) < R$ and $\rho(s) > R$. The endpoints of these segments correspond to *osculation points* on the boundary, for which $\rho(s) = R$.

To see the significance of the alternating segments of the R locus, consider an orbit starting out from s with $p \approx -1$ ($\theta = \pi$). If $R > \rho(s)$ (figure 4), this is part of a circle

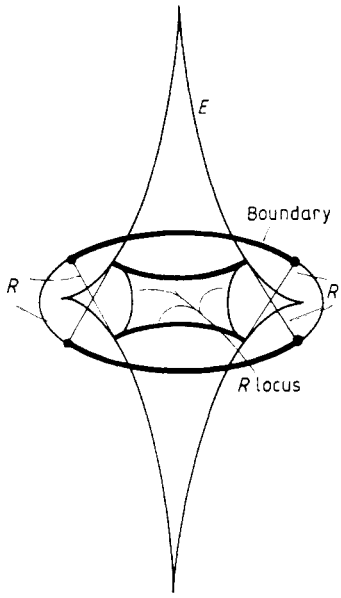


Figure 3. The boundary and its evolute E, together with an R locus for $R > R^*$ (no complete Larmor circles). The bold lines correspond to flyaway intervals $\rho(s) > R$, and the filled circles are the osculation points where $\rho(s) = R$.

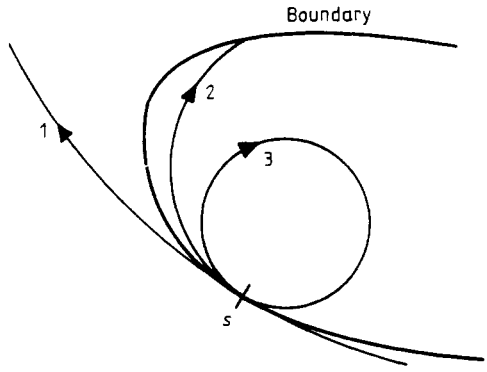


Figure 4. Three grazing orbits with $p = -1$ at s: 1, $R > \rho(s)$; 2, $R < \rho(s)$, giving flyaway; 3, $R < \rho(s)$ giving return to s.

touching the boundary from outside, and the orbit consists of small bounces grazing the boundary, with s decreasing at each bounce. If $R < \rho(s)$, however, the circle touches the boundary from inside, corresponding to a *glancing collision.* There are then two possibilities: the grazing circle may be part of a complete circle which therefore returns near s, so that the orbit again consists of small bounces, with s now increasing at each bounce; or the grazing circle may next intersect the boundary at a point far from s. We call the latter case *flyaway*, and the corresponding boundary segments *flyaway intervals.* Flyaway occurs wherever there are osculation points, i.e. for $\rho_{\min} < R < \rho_{\max}$; it corresponds to the non-existence of an invariant curve with $p = -1$, and hence to a chaotic region associated with these glancing orbits. For field-free billiards

($R = \infty$), osculation points can occur only if the curvature vanishes somewhere (e.g. at an inflection of a non-convex boundary). In this case the non-existence of invariant curves with $p = -1$ has been rigorously proved by Mather (1982) and is illustrated by the heart-shaped billiard of Robnik (1983).

Because we are considering only clockwise-gyrating particles, the grazing orbits with $p = +1$ never display flyaway, so that $p = +1$ is always an invariant curve and the nearby orbits are locally integrable (as shown by Lazutkin (1973) for field-free smooth convex billiards).

When $R^* < R < \rho_{max}$, the flyaway intervals are the boundary segments with $\rho(s) > R$, illustrated in figure 3. When $\rho_{min} < R < R^*$ some sections of these segments are not flyaway intervals because they correspond to complete R circles. In these cases the R locus has self-intersections (figure 5), and each flyaway interval corresponds to a section of the R locus between a cusp and a self-intersection.

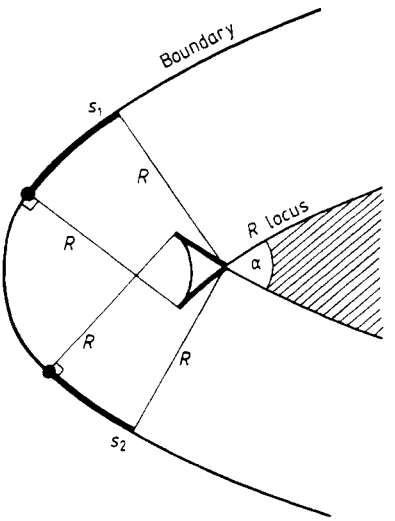


Figure 5. Part of R locus for $\rho_{min} < R < R^*$, showing flyaway intervals (bold) and zone of complete circles (shaded).

Flyaway implies *lines of discontinuity* in the bounce map on the s, p phase cylinder. The lines are the pre-images (figure 6) of the flyaway s intervals with $p = -1$. If A and B are two nearby points separated by a discontinuity line, their iterates $\Phi(A)$ and $\Phi(B)$ may be far apart. Nevertheless, the next iterate of one of the points, $\Phi^2(B)$, say, may lie close to $\Phi(A)$, so that the trajectories in the phase cylinder do lie arbitrarily close together as $A \to B$ except that the orbit of B, say, contains extra steps.

Returning now to the complete circles within the boundary when $R < R^*$, we see that the measure of this integrable component of the motion is proportional to the *area inside the R locus.* If $R < \rho_{min}$ the R locus is a smooth curve, but if $\rho_{min} < R < R^*$ the R locus has cusps and the area inside it has corners at the self-intersections, with interior angles (figure 5) given by

$$\alpha = \pi - \int_{s_1}^{s_2} \kappa(s)\, ds \tag{1}$$

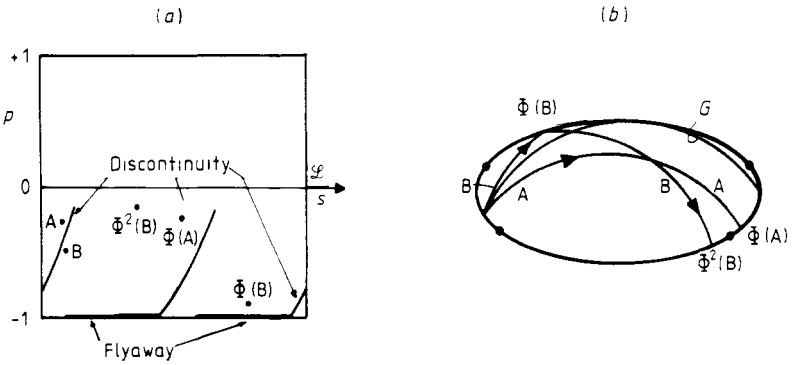


Figure 6. (*a*) Lines of discontinuity of the bounce map on the phase cylinder, emanating from flyaway intervals. (*b*) Trajectories A and B on different sides of a glancing circle *G*.

where s_1 and s_2 are non-osculatory endpoints of the two flyaway intervals associated with the corner.

To summarise this section we predict the following behaviour as the field increases and R decreases from infinity (figure 7). If $\rho_{max} < R < \infty$ (figure 7(*b*)), motion is qualitatively similar to the field-free case; in particular there are invariant curves near $p = -1$, with bounce points rotating slowly clockwise (decreasing s) round the boundary. For $\rho_{min} < R < \rho_{max}$ (figures 7(*c*)(*d*)(*e*)), flyaway implies chaotic motion near

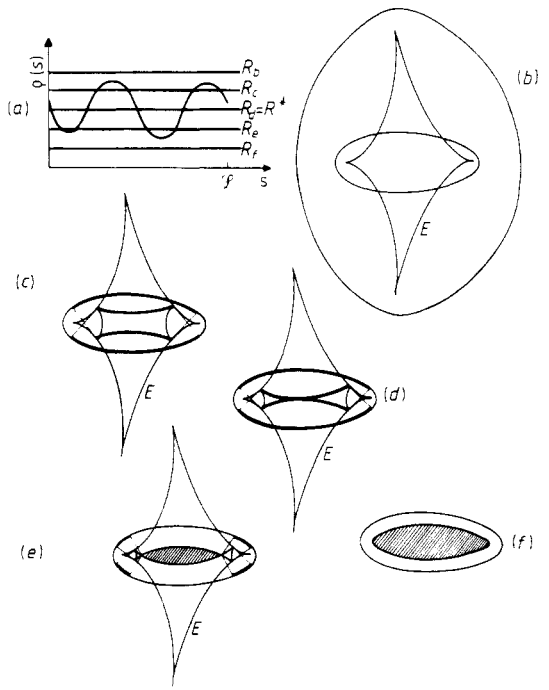


Figure 7. (*a*) Radius-of-curvature function $\rho(s)$; (*b*)-(*f*) boundary, evolute (E) and R loci for R values in (*a*), with flyaway intervals shown bold and zones of complete Larmor circles shaded.

$p = -1$. Within the flyaway range, at $R = R^*$ (figure 7(*d*)), complete Larmor circles appear and the measure of this integrable component increases as R decreases. For $R < \rho_{\min}$ (figure 7(*f*)) there are again invariant curves near $p = -1$ but now bounce points rotate slowly anticlockwise. One can say that the chaos for $\rho_{\min} < R < \rho_{\max}$ enables the 'turning of the tide' from clockwise to anticlockwise at $p = -1$.

3. Diametral closed orbits

Away from the boundaries $p = \pm 1$, the phase cylinder s, p will be dominated by the fixed points associated with the simplest closed orbits; stable fixed points are surrounded by invariant curves, and unstable fixed points generically lie in a chaotic area. If the magnetic field is not too large (R not too small) these simplest orbits will consist of two arcs bouncing between the ends s_1 and s_2 of *diameters* (figure 8), which are

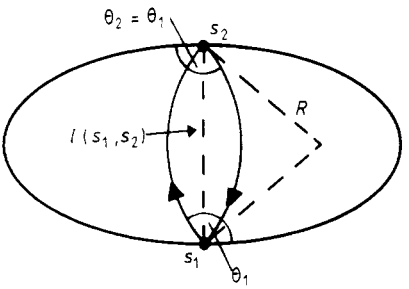


Figure 8. Diametral closed orbit.

defined as chords whose ends both meet the boundary at right angles. If $l(s_1, s_2)$ is the length of this diameter, it follows from elementary geometry (figure 8) that the fixed points corresponding to the closed orbit lie at

$$s = s_1, \qquad s = s_2 \qquad \text{and} \qquad p = \cos\theta_1 = \cos\theta_2 = -l(s_1, s_2)/2R. \tag{2}$$

As R decreases from infinity, the orbit based on a given diameter must eventually cease to exist. It certainly cannot exist when R is smaller than $l(s_1, s_2)/2$, and it may cease to exist for larger R as a result of hitting the boundary between s_1 and s_2.

Every boundary has at least two diameters and hence at least two diametral orbits for large enough R. To see this it is necessary to make use of the fact that diameters are critical points of the arc length function, i.e. s_1 and s_2 satisfy

$$\frac{\partial l(s_1, s_2)}{\partial s_1} = \frac{\partial l(s_1, s_2)}{\partial s_2} = 0. \tag{3}$$

Now l has the following obvious properties:

$$\begin{aligned} &l(s_1 + m\mathscr{L}, s_2 + n\mathscr{L}) = l(s_1, s_2) \quad \text{for all integers } m \text{ and } n \\ &l(s_1, s_2) = l(s_2, s_1) \\ &l(s, s) = 0 \\ &l(s_1, s_2) > 0 \qquad \text{if } s_1 \neq s_2. \end{aligned} \tag{4}$$

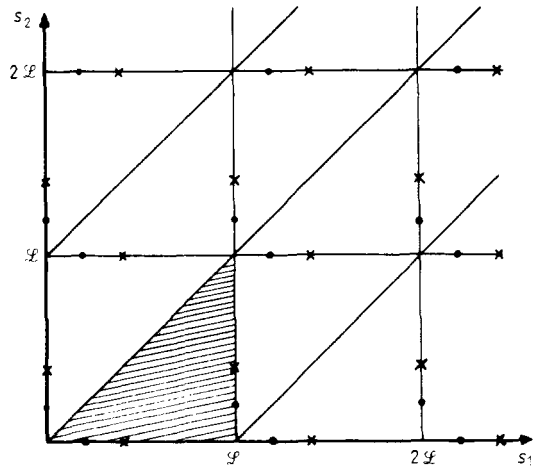


Figure 9. The plane s_1, s_2 of the chord length function l, with identifications generated by (4) illustrated by dots and crosses and implying that the fundamental domain of l is a Möbius strip (shaded).

It follows that the domain of $l(s_1, s_2)$ is a *Möbius strip* (figure 9), on whose edge, $s_1 = s_2$, l vanishes. Within the strip l is positive and moreover has no minima (a diameter of a convex boundary can always be made smaller by parallel displacement, i.e. $\delta s_1 = -\delta s_2$). Therefore the only critical points can be maxima or saddles, and the index theorem guarantees that every function on this Möbius strip has at least one of each, corresponding to at least two diameters.

The *stability* of the diametral orbits is described by the matrix M giving the state after two bounces of an orbit that starts out infinitesimally displaced from one of the fixed points (2). This is the product of the deviation matrices for individual bounces, i.e.

$$M = M_{21}M_{12} \qquad \text{where } M_{12} = \begin{pmatrix} \dfrac{\partial s_2}{\partial s_1} & \dfrac{\partial s_2}{\partial p_1} \\ \dfrac{\partial p_2}{\partial s_1} & \dfrac{\partial p_2}{\partial p_1} \end{pmatrix}. \tag{5}$$

A substantial simplification results from the fact that each deviated arc is associated with a deviated chord whose length l differs only quadratically from that of the diameter, whereas M relates linear deviations of points on the phase cylinder. Therefore the stability matrix for a diametral orbit is *independent of the magnetic field*, and elementary calculations (e.g. in Robnik (1983)) give

$$M = \begin{pmatrix} 1-2l(\kappa_1+\kappa_2)+2l^2\kappa_1\kappa_2 & 2l(1-l\kappa_2) \\ 2(l\kappa_1-1)(\kappa_1+\kappa_2-l\kappa_1\kappa_2) & 1-2l(\kappa_1+\kappa_2)+2l^2\kappa_1\kappa_2 \end{pmatrix} \tag{6}$$

where $l \equiv l(s_1, s_2)$, $\kappa_1 \equiv \kappa(s_1)$ and $\kappa_2 \equiv \kappa(s_2)$. Of course, $\det M = 1$.

The eigenvalues of M either both lie on the unit circle, in which case the orbit is stable, or they are both real, in which case the orbit is unstable. Thus

$$|\mathrm{Tr}\, M| = 2|1-2l(\kappa_1+\kappa_2)+2l^2\kappa_1\kappa_2| \begin{cases} <2 \text{ (stability)} \\ >2 \text{ (instability)} \end{cases} \tag{7}$$

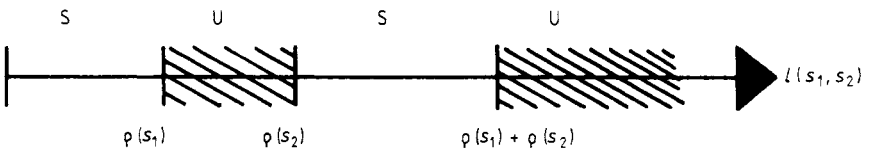


Figure 10. Stable (S) and unstable (U) regimes of chord length and curvature for diametral closed orbits.

It follows that the orbit is unstable if $l > \rho(s_1) + \rho(s_2)$ or if l lies between $\rho(s_1)$ and $\rho(s_2)$, and stable otherwise, as illustrated in figure 10.

In terms of the topography of the function $l(s_1, s_2)$ on its Möbius strip, maxima correspond to unstable orbits with $l > \rho(s_1) + \rho(s_2)$, and saddles correspond to unstable orbits with l between $\rho(s_1)$ and $\rho(s_2)$ (if contours issue into all four quadrants of s_1, s_2) or to stable orbits (if contours issue into only two quadrants). Therefore it does not follow from the necessary existence of at least one maximum and one saddle that one diametral closed orbit is stable, and indeed it is not hard to construct convex boundaries with two diameters which are both unstable (at least one of these diameters must have $\rho(s) \neq \rho(s_2)$).

4. Adiabatic skipping

There are certain regimes of R and p, illustrated in figure 11, for which the arc length interval Δs between successive bounces can be expected to be always much less than the local radius of curvature $\rho(s)$. For these short skips along the boundary it is natural to consider the curvature as changing adiabatically, and seek a corresponding conserved function $C(s, p)$ whose contours are invariant curves of the bounce map, implying local integrability in the sense of Kolmogorov, Arnold and Moser.

To determine the adiabatic invariant $C(s, p)$, we need to calculate $\Delta s \equiv s_2 - s_1$ and $\Delta\theta \equiv \theta_2 - \theta_1$ at the ends of an arc, to lowest order. Referring to figure 12 we have the

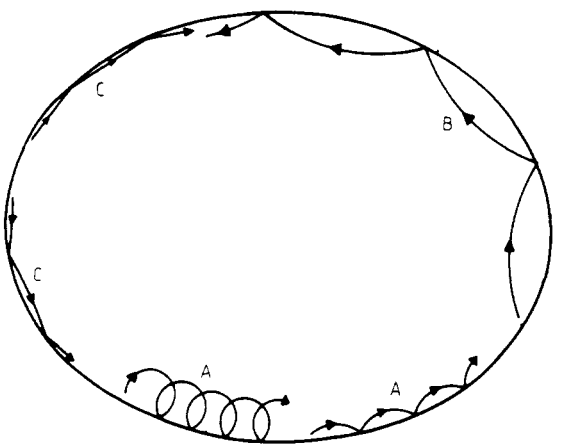


Figure 11. Adiabatic regimes of short skips along the boundary: A, $R \ll \rho_{min}$, p arbitrary; B, $\rho_{min} < R < \rho_{max}$, p near $+1$; C, $R \gg \rho_{max}$, p near ± 1.

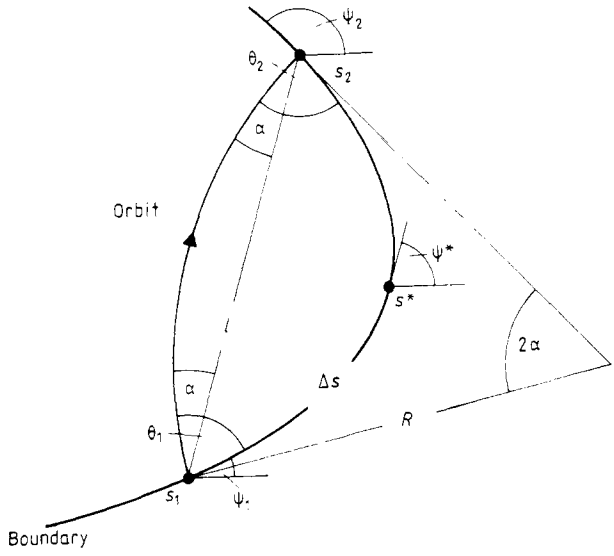


Figure 12. Geometry and notation for calculating adiabatic invariant.

following formulae for the direction ψ^* of the chord connecting 1 and 2:

$$\psi^* = \psi_1 + \theta_1 - \alpha = \psi_2 - \theta_2 + \alpha. \tag{8}$$

Thus

$$\Delta\theta = \psi_1 + \psi_2 - 2\psi^*. \tag{9}$$

The tangent directions ψ are related to the curvature by

$$\psi(s) = \psi^* + \int_{s^*}^{s} \mathrm{d}s'\, \kappa(s') \tag{10}$$

and the endpoints are related by the condition that the perpendicular distance from chord to boundary is zero, i.e.

$$\int_{s_1}^{s_2} \mathrm{d}s \sin(\psi^* - \psi) = 0. \tag{11}$$

So far all formulae are exact. Now we make a linear approximation for the curvature:

$$\kappa(s) \simeq \kappa_0 + \kappa_1(s - s^*). \tag{12}$$

Substituting this in (10) and then (11) and (9), retaining lowest-order terms gives, after elementary calculations,

$$\Delta\theta \simeq \kappa_1(\Delta s)^2/6. \tag{13}$$

Next we use (8) again, in the form

$$\psi_2 - \psi_1 \simeq \kappa_0 \Delta s = \theta_1 + \theta_2 - 2\alpha. \tag{14}$$

But

$$R \sin\alpha = l/2 \simeq \sin(\kappa_0 \Delta s/2)/\kappa_0 \sim \Delta s/2 \tag{15}$$

so that

$$\kappa_0 \Delta s = \theta_1 + \theta_2 - 2\sin^{-1}(\Delta s/2R). \tag{16}$$

It now follows from the smallness of $\kappa_0 \Delta s$ that

$$\Delta s \simeq \frac{2R\sin\theta}{1+\kappa R\cos\theta} = \frac{2R(1-p^2)^{1/2}}{1+\kappa R p} \tag{17}$$

where we now write κ for the local curvature κ_0 and θ for the local orbit direction $(\theta_1+\theta_2)/2$.

From (13) we now have, on taking the limit $\Delta s \to 0$,

$$\frac{d\theta}{ds} = \frac{\kappa_1 \Delta s}{6} = \frac{d\kappa}{ds}\frac{R(1-p^2)^{1/2}}{3(1+\kappa R p)} \tag{18}$$

so that

$$\frac{dp}{d\kappa} = -\sin\theta \frac{d\theta}{ds}\left(\frac{d\kappa}{ds}\right)^{-1} = \frac{-R(1-p^2)}{3(1+\kappa R p)}. \tag{19}$$

This is a linear equation for $d\kappa/dp$ which can be solved exactly to give the adiabatic invariant as

$$C(s,p) = \frac{R\kappa(s)+p(3-2p^2)}{(1-p^2)^{3/2}} = \frac{R\kappa(s)}{\sin^3\theta} + \cot^3\theta + 3\cot\theta. \tag{20}$$

We can expect the contours of this function to correspond to actual invariant curves of the bounce map whenever there are short skips, i.e. wherever (cf. (17))

$$\left|\frac{\kappa R(1-p^2)^{1/2}}{1+\kappa R p}\right| \ll 1. \tag{21}$$

This condition will hold for the three regimes in figure 11 which we now discuss individually.

In the *strong-field limit* $\kappa R \to 0$ (regime A), (20) gives, for the invariant curve passing through s_0, p_0,

$$p = p_0 + \frac{R(\kappa(s_0)-\kappa(s))}{3}(1-p_0^2) \tag{22}$$

and this formula holds for arbitrary p.

In the *forward-glancing limit* $p \to 1$ (regime B), (20) gives

$$2^{3/2} C(s,p) \to \frac{R\kappa(s)+1}{(1-p^2)^{3/2}} \tag{23}$$

for arbitrary R.

In the *weak-field glancing limit* $\kappa R \to \infty$, $p = \to \pm 1$ (regime C), (20) gives the constant as

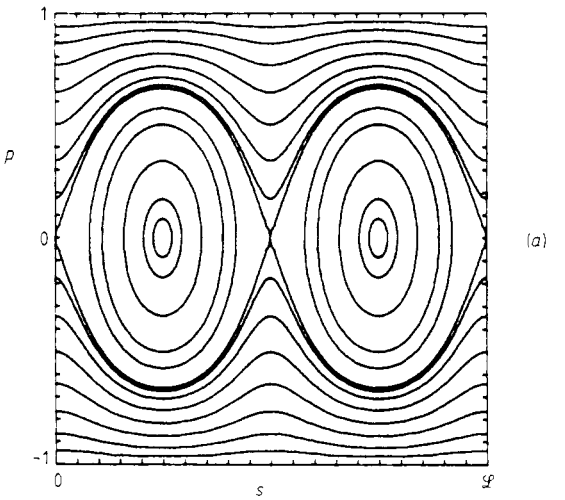
$$R^{-1} C(s,p) \to \frac{\kappa(s)}{(1-p^2)^{3/2}}. \tag{24}$$

This is the invariant associated with the caustics close to the boundary, whose existence was proved by Luzutkin (1973). Geometrically, it means that the difference between the arc length Δs and the chord length l is conserved when both are small.

The validity condition (21) is strongly violated at osculation points ($\kappa(s)R = 1$) for backward glancing orbits ($p = -1$), and this is consistent with our arguments in § 2 that because of flyaway such orbits are highly unstable and correspond to chaotic areas where the phase cylinder has no invariant curves.

It is interesting to ignore the validity condition and study the global structure of the invariant curves that (20) predicts. This is best understood in terms of the critical points of $C(s, p)$, which correspond to fixed points in the adiabatic approximation. By differentiation, these lie at s, p where

$$\frac{\mathrm{d}\rho(s)}{\mathrm{d}s} = 0 \qquad p = -\frac{\rho(s)}{R} \tag{25}$$



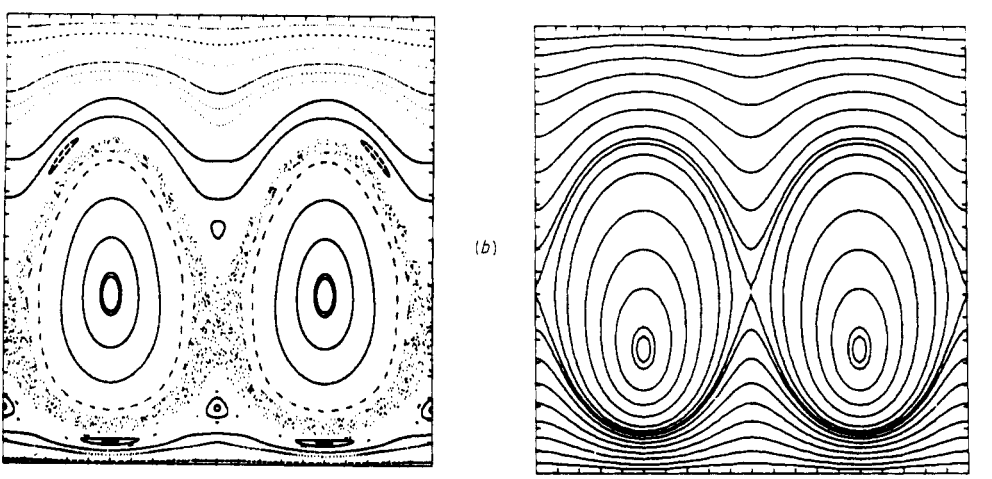


Figure 13. Bounce maps (left-hand pictures) and adiabatic invariant curves based on the approximation (20) (right-hand pictures) for magnetic billiards in an ellipse with semi-major axis $a = 1$ and semi-minor axis $b = \frac{3}{4}$, for Larmor radius R with the following values: (*a*) $R = \infty$ (integrable case); (*b*) $R = 3$; (*c*) $R = 0.75$; (*d*) $R = 0.435$; (*e*) $R = 0.2$. (The labelling of axes in (*b*)-(*e*) is the same as in (*a*).)

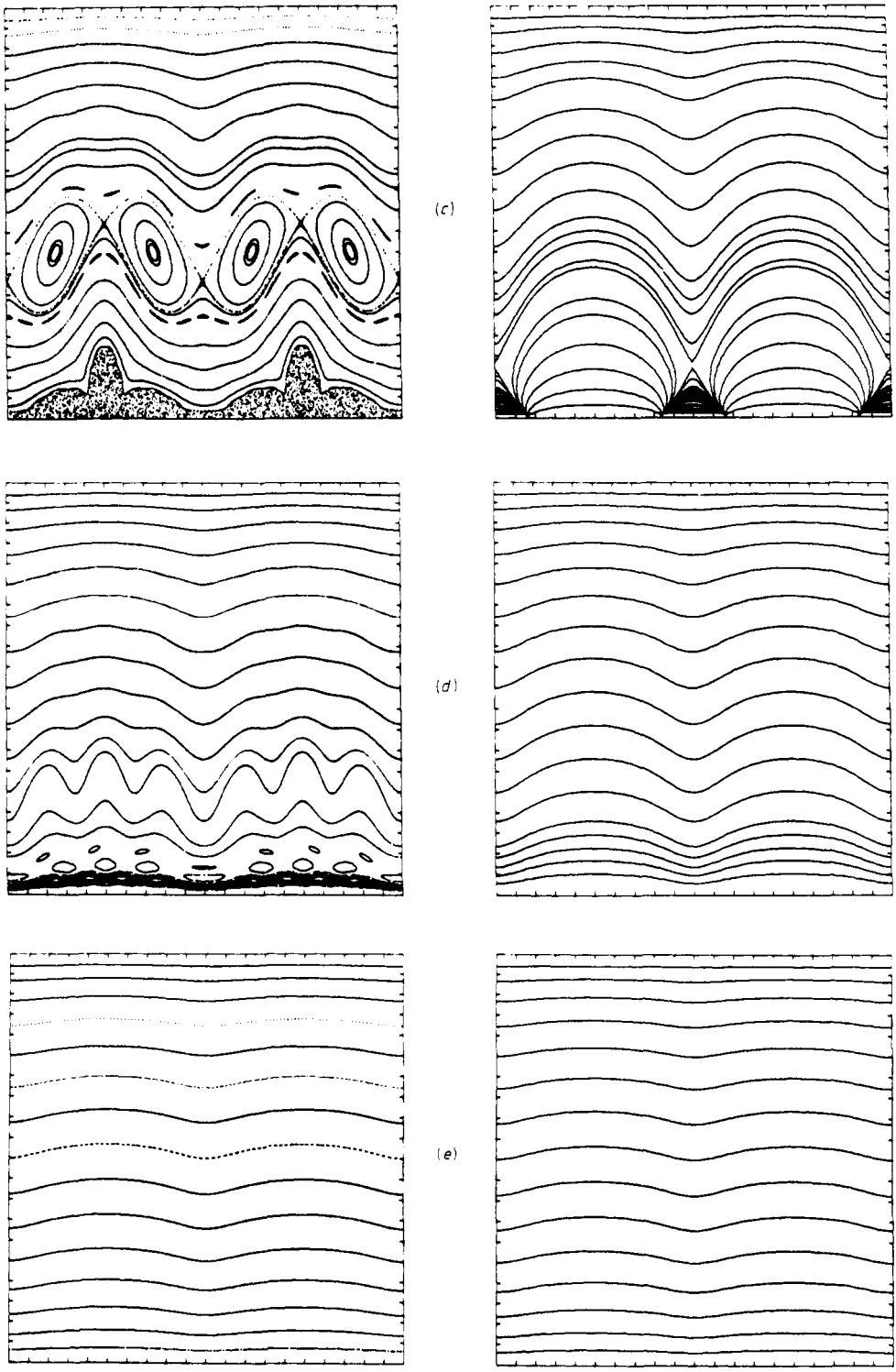


Figure 13.

and therefore correspond to vertices of the boundary. Maxima of ρ give minima of $C(s, p)$ (stable adiabatic fixed points), and minima of ρ give saddles of $C(s, p)$ (unstable adiabatic fixed points). There are no maxima of $C(s, p)$.

If $R < \rho_{\min}$ there are no critical points with $|p| < 1$, and the adiabatic invariant curves are all loops around the phase cylinder, corresponding to orbits skipping anticlockwise for all p. If $R > \rho_{\min}$ the phase cylinder has adiabatic fixed points in stable and unstable pairs (because the associated maximal and minimal vertices come in pairs as shown in appendix 1). These fixed points are adiabatic approximations to the true fixed points and by comparison with § 3 we see that they correspond to replacing the ends of diameters by vertices and (cf (2) and (25)) chord half-lengths by radii of curvature. Of course this represents the adiabatic approximation's attempt to describe the phase cylinder far outside its domain of validity, and any accurate prediction of the positions of the fixed points must be regarded as fortuitous.

We remark that higher-order adiabatic approximations can be constructed near periodic points provided the rotation number is small, again by replacing differences by differentials but this time in a power of the map rather than the map itself. These approximations have the advantage that they extend beyond the linear regime of the elementary stability analysis.

5. Example: the ellipse

For an elliptic boundary with semi-major axis a and semi-minor axis b the eccentricity is $(1 - b^2/a^2)^{1/2}$ and the important radii of § 2 are

$$\rho_{\min} = b^2/a \qquad \rho_{\max} = a^2/b \qquad R^* = b. \tag{26}$$

There are four vertices, for which, because of symmetry, $\rho(s)$ has two equal minima (at $s = 0$ and $s = \mathscr{L}/2$, say) and two equal maxima (at $s = \mathscr{L}/4$ and $s = 3\mathscr{L}/4$). These vertices are also the endpoints of diameters. The long diametral orbit ($l = 2a$) is always unstable; the short diametral orbit is stable unless $b = a/\sqrt{2}$, when it is marginally stable (because the left-hand unstable interval in figure 10 degenerates to a point for equal endpoint curvatures).

In zero field the ellipse billiard is integrable (see, for example, Berry 1981) with the product of angular momenta about the foci being conserved. It is remarkable that this conserved quantity is given exactly by the adiabatic approximation in the limit $R \to \infty$ (equation (24)).

In the following computations of orbits of the bounce map, we used Newton's method to find intersections of the Larmor circle with the ellipse.

The curvature regimes are illustrated in figure 13 which shows bounce maps for the ellipse with $a = 1$, $b = \frac{3}{4}$, and thus $\rho_{\min} = \frac{9}{16}$, $\rho_{\max} = \frac{4}{3}$, $R^* = \frac{3}{4}$. In the integrable case $R = \infty$ (figure 13(a)), the phase cylinder is covered with invariant curves, surrounding it near $p = +1$ (iterates circulating anticlockwise) and near $p = -1$ (iterates circulating clockwise) and surrounding the two stable fixed points at $s = \mathscr{L}/4$ and $3\mathscr{L}/4$, $p = 0$. The separatrix passes through the unstable fixed points at $s = 0$, $\mathscr{L}/2$, $p = 0$.

In the weak magnetic field corresponding to $R = 3$ (i.e. $R > \rho_{\max}$), the bounce map (figure 13(b), left) shows that invariant curves survive near $p = \pm 1$ as predicted by the skipping theory of § 4, and are accurately represented by the contours (figure 13(b), right) of the adiabatic invariant (20). The diametral orbits survive and correspond to

the fixed points (2) as is usual, the stable one is surrounded by invariant curves and the unstable one is embedded in a chaotic area.

Increasing the field so that $R = R^* = \frac{3}{4}$ brings the system into the flyaway regime of § 2, and the bounce map (figure 13(*c*), left) shows the predicted absence of invariant curves near $p = -1$. The unstable diametral orbit no longer exists. The short diametral orbit is on the point of disappearing at $p = -1$ but it is not visible in the chaotic area of the map because its domain of stability has shrunk to zero. Higher-order (four bounce) stable and unstable closed orbits organise island chains for small negative p. For p near $+1$ invariant curves persist as expected, and as shown in figure 13(*c*) (right) they are well approximated by the invariant (20).

For the stronger field corresponding to $R = 0.435$ (i.e. $R < \rho_{\min}$) the system should no longer display flyaway, and there should be invariant curves near $p = \pm 1$. The bounce map (figure 13(*d*), left) shows the curves near $p = +1$ (iterates circulating

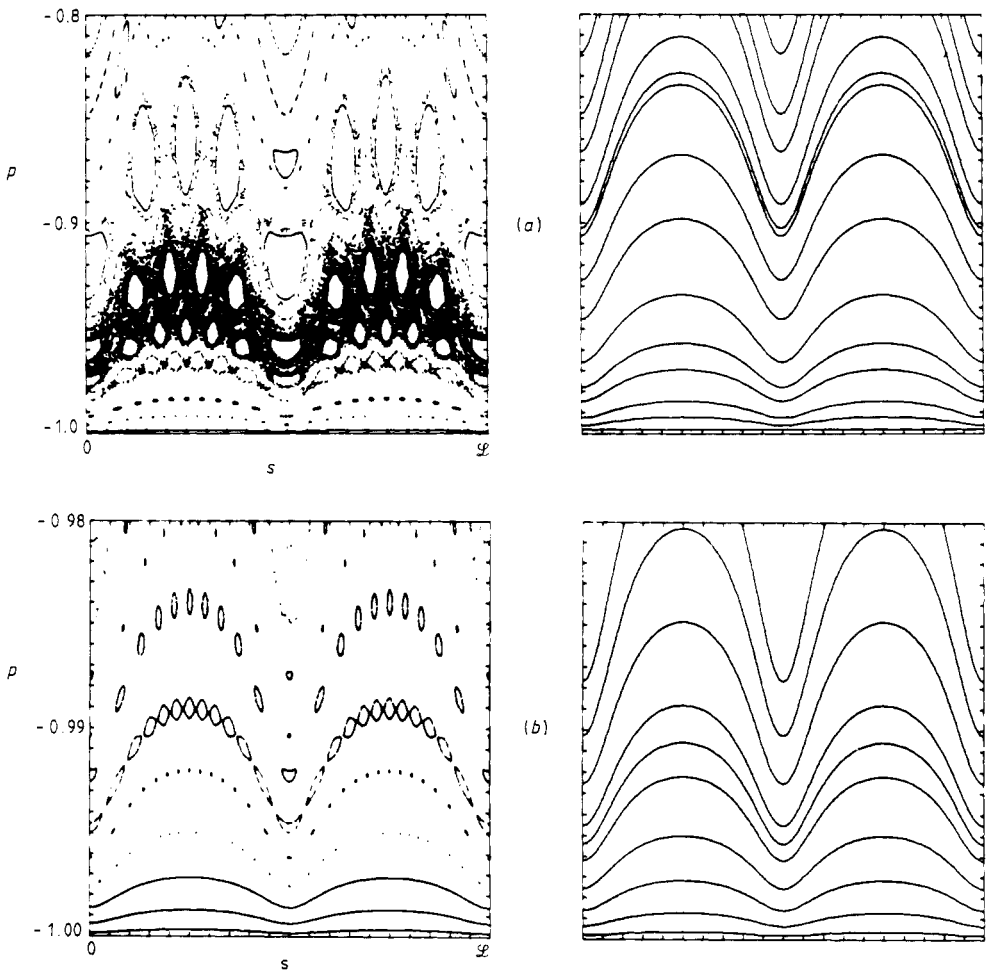


Figure 14. (*a*) Ten-fold magnification of figure 13(*d*) in p direction near $p = -1$; (*b*) ten-fold magnification of (*a*), showing predicted invariant curves near backward glancing orbits ($p \approx -1$). (The labelling of axes on the right-hand pictures is the same as on the left-hand ones.)

anticlockwise) very clearly, and moreover they are accurately described by the adiabatic invariant (figure 13(d), right) even down to $p \approx 0$. To see the invariant curves near $p = -1$ it is necessary to magnify the bounce map in this region. Two successive ten-fold magnifications are shown in figures 14(a) and (b) (left). These reveal a rich structure of islands and chaos, with the predicted invariant curves (also circulating anticlockwise) appearing only in the narrow range $-1 \leq p \leq -0.997$, and well approximated by the adiabatic invariant (figure 14(b), right).

Finally, the very strong field with $R = 0.2$ shows a phase plane (figure 13(e), left) covered with invariant curves, corresponding to adiabatic skipping for all p. These curves are indistinguishable from those in figure 13(e) (right), generated from the invariant (20).

The general picture supported by these computations is that increasing field induces flyaway chaos near $p = -1$ when $\rho_{\min} < R < \rho_{\max}$, with this chaos being squeezed out of existence by invariant curves from above ($p \approx +1$) and below ($p \approx -1$) as the field increases further so that $R < \rho_{\min}$. We emphasise the point made in the introduction, that this picture holds only for sufficiently smooth boundaries, and does not apply for example to magnetic billiards in rectangles or stadiums.

Acknowledgments

This research was supported by an Assistantship (for MR) from the SERC, and not by any military agency.

Appendix 1. Every billiard has at least four curvature extrema

(A referee has pointed out that this is the 'four-vertex theorem' (Stoker 1969), and that an argument similar to that which follows has been given by Guckenheimer (1977).)

Referring to figure 1, let ψ be the direction between the clockwise tangent and the x axis. Then the radius of curvature $\rho(\psi)$ is a periodic function of ψ and so can be written

$$\rho(\psi) = \sum_{k=0}^{\infty} \rho_k \cos(k\psi + u_k). \tag{A1}$$

In terms of $\rho(\psi)$ the x and y coordinates of boundary points are

$$\begin{aligned} x(\psi) - x(0) &= \int_{s(0)}^{s(\psi)} \cos\psi \, \mathrm{d}s = \int_0^{\psi} \mathrm{d}\psi (\mathrm{d}s/\mathrm{d}\psi) \cos\psi = \int_0^{\psi} \mathrm{d}\psi \, \rho(\psi) \cos\psi \\ y(\psi) - y(0) &= \int_0^{\psi} \mathrm{d}\psi \, \rho(\psi) \sin\psi. \end{aligned} \tag{A2}$$

The boundary is a closed curve, so $x(2\pi) - x(0) = y(2\pi) - y(0) = 0$, and substituting the Fourier series (A1) immediately gives $\rho_1 = 0$. Curvature extrema satisfy

$$\frac{\mathrm{d}\rho}{\mathrm{d}\psi}(\psi) = -\sum_{k=2}^{\infty} k\rho_k \sin(k\psi + u_k) = 0. \tag{A3}$$

This is a periodic function and so generically has an even number of extrema (alternately maxima and minima) on the range $0 \leq \psi < 2\pi$. To show that this number is at least

four we assume the contrary, that there are only two. Then we can construct a function

$$g(\psi) = g_0 + g_1 \cos(\psi + v_1) \tag{A4}$$

whose zeros coincide with those of $d\rho/d\psi$ and which can therefore always have the same signs as $d\rho/d\psi$. Therefore we would have

$$\int_0^{2\pi} d\psi\, g(\psi)\, d\rho(\psi)/d\psi > 0. \tag{A5}$$

But direct substitution of (A3) and (A4) shows that this integral is identically zero and so we have generated a contradiction. Therefore ρ has at least four extrema.

Appendix 2. Fundamental discrete symmetries of the bounce map

Symmetries of a Hamiltonian H are important in understanding the organisation of classical phase space and particularly in finding periodic orbits. Continuous symmetries are related to additional integrals of motion and will not be discussed here. We shall consider discrete symmetries related to the following three operations in four-dimensional phase space (assuming a system with two freedoms such as a magnetic planar billiard):

Inversion P: $(x, y, p_x, p_y) \to (-x, -y, -p_x, -p_y)$
Time reversal T: $(x, y, p_x, p_y) \to (x, y, -p_x, -p_y)$
Reflection S_x: $(x, y, p_x, p_y) \to (-x, y, -p_x, p_y)$.

These maps of phase space onto itself satisfy $P^2 = T^2 = S_x^2 = I$ where I is the identity. For any symmetry U such that $U^2 = I$, we will say that H has the symmetry U if $HU = H$.

For a charged particle in a uniform constant magnetic field B and a scalar potential $V(x, y)$ (hard-walled in the case of billiards), there is a gauge in which

$$H = p_x^2 + (p_y - Bx)^2 + V(x, y).$$

It follows at once that

(i) H has P if V has P;
(ii) H has T if $B = 0$;
(iii) H has S_x if $B = 0$ and V has S_x;
(iv) H has TS_x if $B \neq 0$ and V has S_x.

In our case (billiards), $V(x, y)$ has the symmetry of the boundary curve, and the preceding assertions (i)-(iv) can be phrased in terms of the two-dimensional phase cylinder s, p of the bounce map, as follows.

(i) If the boundary has P, then corresponding to the orbit through (s, p) is a geometrically similar one through $(s + \mathscr{L}/2, p)$, i.e. the structure of the phase cylinder has period $\mathscr{L}/2$ rather than $\mathscr{L}$.

(ii) If $B = 0$, the phase cylinder has reflection symmetry about the (mirror) line $p = 0$.

(iii) If the boundary has S_x and $B = 0$, and if the arc length origin $s = 0$ is chosen on the symmetry line at the boundary, then corresponding to the orbit through (s, p) is a geometrically similar one through $(-s, -p)$, i.e. the bounce map has inversion symmetry as well as the T symmetry resulting from $B = 0$.

(iv) If the boundary has S_x and $B \neq 0$, then (with the same s origin as in (iii)) corresponding to the orbit through (s, p) is a geometrically similar one through $(-s, p)$, i.e. the bounce map has reflection symmetry about the (mirror) line $s = 0$.

One sees that (ii) and (iv) are geometrically equivalent in the sense that phase space has a mirror symmetry (although of course the cylinder topology prevents this equivalence from being complete for the bounce map). An important quantum mechanical consequence of this type of equivalence, which we will discuss further in another paper, is that the energy levels of a Hamiltonian operator without T but with TS_x may, in their statistical properties, mimic those of an operator with T—a sort of 'false time reversal symmetry'.

References

Berry M V 1981 *Eur. J. Phys.* **2** 91-102
—— 1984 *Proc. Como Conf. on Quantum Chaos* ed G Casati (New York: Plenum)
Drury L O'C 1983 *Rep. Prog. Phys.* **46** 973-1027
Guckenheimer J 1977 *Topology* **16** 177-80
Lazutkin Y F 1973 *Izv. Acad. Sci. Ser. Math.* **37** No 1, 186-216
Mather J N 1982 *Ergodic Theory and Dynamical Systems* **2** 3-4
Peierls R E 1979 *Surprises in Theoretical Physics* (Princeton, NJ: Princeton University Press)
Robnik M 1983 *J. Phys. A: Math. Gen.* **16** 3971-86
Ruderman M 1975 *Ann. NY Acad. Sci.* **257** 127-40
Sinai Ya G 1976 *Introduction to Ergodic Theory* (Princeton, NJ: Princeton University Press)
Stoker J J 1969 *Differential Geometry* (New York: Wiley-Interscience) pp 48-50