

Integrable elliptic billiards and ballyards

Peter Lynch

School of Mathematics and Statistics, University College, Dublin

E-mail: `Peter.Lynch@ucd.ie`

July 2019

Submitted to: *Eur. J. Phys.*

Abstract. The billiard problem concerns a point particle moving freely in a region of the horizontal plane bounded by a closed curve Γ , and reflected at each impact with Γ . The region is called a ‘billiard’, and the reflections are specular: the angle of reflection equals the angle of incidence. We review the dynamics in the case of an elliptical billiard. In addition to conservation of energy, the quantity L_1L_2 is an integral of the motion, where L_1 and L_2 are the angular momenta about the two foci. We can regularize the billiard problem by approximating the flat-bedded, hard-edged surface by a smooth function. We then obtain solutions that are everywhere continuous and differentiable. We call such a regularized potential a ‘ballyard’. A class of ballyard potentials will be defined that yield systems that are completely integrable. We find a new integral of the motion that corresponds, in the billiards limit $N \rightarrow \infty$, to L_1L_2 . Just as for the billiard problem, there is a separation of the orbits into boxes and loops. The discriminant that determines the character of the solution is the sign of L_1L_2 on the major axis.

Keywords: Billiards. Integrable systems. Particle dynamics.

1. Introduction

In his *Lectures on Theoretical Physics*, Arnold Sommerfeld [6] wrote “The beautiful game of billiards opens up a rich field for applications of the dynamics of rigid bodies. One of the illustrious names in the history of mechanics, that of Coriolis, is connected with it.” Sommerfeld was referring to a book by Gaspard-Gustave de Coriolis, *Théorie mathématique des effets du jeu de billiard*, published in Paris in 1835 [5].

Billiards has been used to examine questions of ergodic theory [4]. In ergodic systems, all configurations and momenta compatible with the total energy are eventually explored. Such questions lie at the foundation of statistical mechanics. We know that the dynamics on an elliptic billiard is integrable: its caustics are confocal ellipses and hyperbolas. George Birkhoff conjectured that if the neighbourhood of a strictly convex smooth boundary curve is foliated by caustics, then the curve must be an ellipse. So

far, this conjecture — that ellipses are characterized by their integrability — remains an open problem.

The simplest billiard is circular. The dynamics are easily described: every trajectory makes a constant angle with the boundary and is tangent to a concentric circle within it. Berger [1, p. 713] observed that “Circular billiards is indeed a banal subject, but nonetheless we visualize it rapidly.” Every trajectory is either a polygon (perhaps star-shaped) or is everywhere dense in an annular region. Moreover, the elliptical billiard problem is completely resolved, thanks to Poncelet’s theorem and the known geometry of confocal conics. We have periodic trajectories, or ones that are dense in regions of two distinct topological types. More generally, Tabachnikov [7] discusses the theory of convex smooth billiards — the elliptic case. Berry [2] considered various deformations of the circular billiard, and showed how they support a variety of different orbits.

In §2 we review the well-known theory of billiards on an elliptical domain. In this case, the velocity has a jump discontinuity at each impact. In §3, we introduce *ballyard* domains, regularizing the problem by approximating the flat-bedded, hard-edged surface by a smooth function. This ensures that the solutions are everywhere smooth. We define a countably infinite class of ballyard potentials depending on a parameter N , each of which is completely integrable. For this class, we find a new integral of the motion, \mathbb{L} , that corresponds, in the billiards limit $N \rightarrow \infty$, to $L_1 L_2$. It follows from this integral and the nature of the potential surface that the orbits split into boxes and loops. The discriminant that determines the character of the solution is the sign of $L_1 L_2$ on the major axis.

2. Elliptical Billiards

We idealize the game of billiards, assuming the ball is a point mass moving at constant velocity between elastic impacts with the boundary, or cushion, of the billiard table. The energy is taken to be constant. The path traced out by the moving ball may form a closed periodic loop or, more generically, it may cover the table or part of it densely, never returning to the starting conditions.

The billiard table may be of any shape, but we generally assume that it is either rectangular, like a normal table, or elliptical. In this section, we examine the orbits for an elliptical table. We assume for simplicity that the centre of the ellipse is at the origin, the major axis coincides with the x -axis and the semi-axes are a and b . Thus, the boundary is described by the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

It is also useful to recall the parametric form

$$x = a \cos \theta, \quad y = b \sin \theta$$

where θ is called the eccentric anomaly (note that θ is not identical to the polar angle,

which we denote by ϑ). The foci are at $(f, 0)$ and $(-f, 0)$ where $f^2 = a^2 - b^2$. Defining the eccentricity e by $e^2 = 1 - (b/a)^2$, it follows that the foci are at $(ea, 0)$ and $(-ea, 0)$.

Initial Conditions.

Initial conditions for motion in a plane normally require four numbers, two for the position and two for the momentum. Since we can assume without loss of generality that the ball moves at unit speed, only the direction of the initial momentum is required. Likewise, if the motion starts from the boundary, a single value determines the position. We suppose a trajectory starts at a boundary point with eccentric anomaly θ_0 and moves at an angle ψ_0 to the x -axis. The pair of values $\{\theta_0, \psi_0\}$ determines the entire motion.

Each bounce is a specular reflection: the angle between the normal to the cushion and the incoming trajectory is equal to the corresponding angle for the outgoing trajectory. The tangential component of velocity is unchanged after impact, while the normal component reverses sign. Thus, the speed remains constant.

The first linear segment of the trajectory is tangent to another conic confocal with the boundary. Because of the specular reflection at the boundary, all subsequent segments are also tangent to this conic [1]. The conic is called the caustic of the orbit, and may be an ellipse or a hyperbola.

Generic Motion: Box Orbits & Loop Orbits.

Since the speed is a constant, taken to be unity, the orbit is determined by the initial position and angle of motion. Alternatively, we can simply give the first two points of impact with the cushion. There are two generic types of orbit. The first arises when the ball crosses the major axis between the foci, the second when it crosses outside.

In the first case, once the trajectory brings the ball across the interval $-f < x < +f$ on the x -axis, it is constrained by the geometry of the ellipse to remain within a region bounded by a hyperbola having the same foci as the original ellipse. Each segment of the orbit between impacts with the ellipse is tangent to this hyperbola (see figure 1, left panel). Generically, the orbit is dense in this region, coming arbitrarily close to any given point within it.

If the ball crosses the major axis outside the foci, that is for $x \in (-a, -f)$ or $x \in (f, a)$, it will continue to do so. It will never pass between the foci. Each segment is tangent to a smaller ellipse, confocal with the original boundary (see figure 1, right panel). Generically, the orbit is dense in the annular region between the two ellipses.

The Homoclinic Orbit: Motion through the Foci.

The two generic families of orbits are separated by the homoclinic orbits, for which the ball passes through a focus. Due to specular reflection at the boundary, it will pass through a focus on each subsequent segment of its path. Moreover, the orbit rapidly approaches horizontal motion back and forth along the major axis (see figure 2).

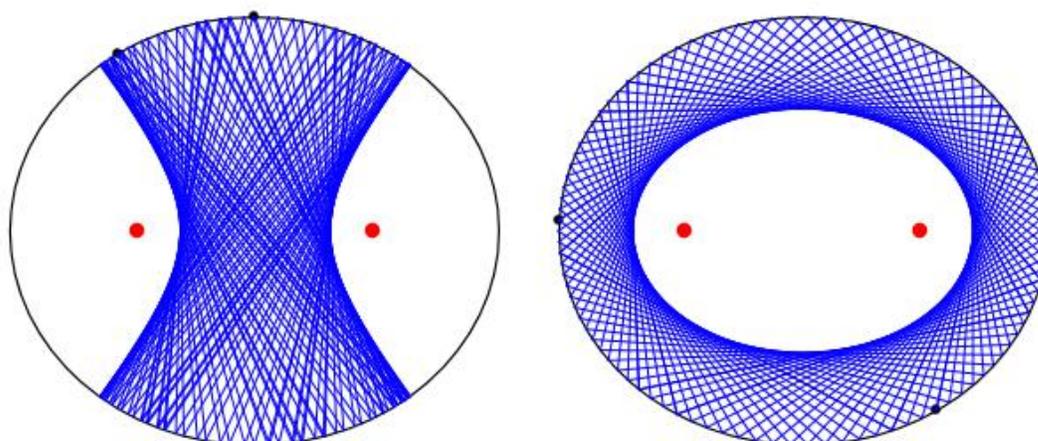


Figure 1. Generic orbits. Left: Box orbit, crossing the major axis between the foci. Right: Loop orbit, crossing the major axis outside the foci.

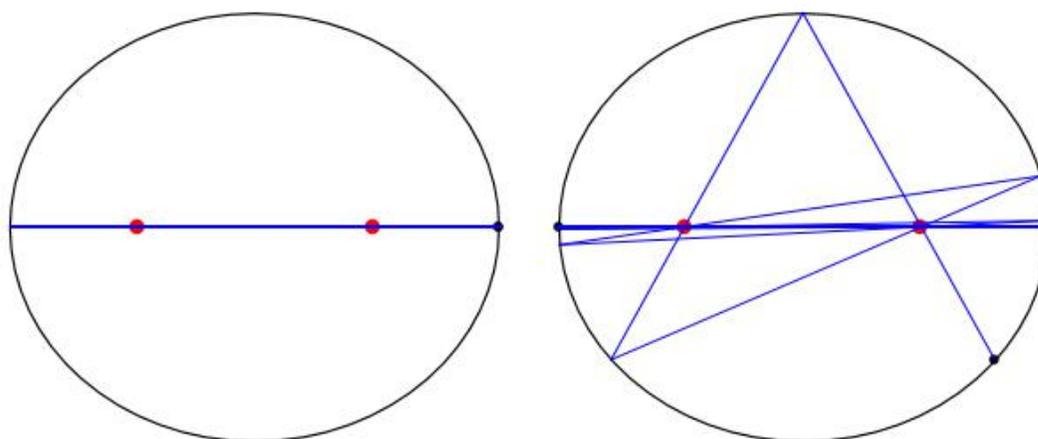


Figure 2. Homoclinic orbits. Left: Singular orbit passing through the two foci. Right: Any orbit through a focus rapidly approaches motion along the major axis.

There is an interesting paradox here: the dynamical behaviour is reversible in time. However, all orbits through the foci tend to the same limiting orbit, ultimately bouncing back and forth between the ends $(-a, 0)$ and $(+a, 0)$ of the ellipse. What if we start on the same trajectory but reverse the time? We will find that, once again, the orbit will approach horizontal oscillations. So, starting from motion close to the end state, two things can happen: either this back-and-forth motion will continue indefinitely or the y -component will gradually increase, reach a maximum and then die away rapidly. The graph of y against time is reminiscent of a wave-packet or soliton.

Periodic Orbits.

For special choices of the initial position and direction of movement, closed orbits ensue. The simplest are period-2 orbits, where there are just two points of impact. These are along the major or minor axes. We have seen that horizontal oscillations are unstable to some small perturbations: the motion may develop a large component along the minor axis, and then return to a quasi-horizontal orbit.

There is a simple period-4 orbit touching the ellipse at the four stationary points, that is the points where the distance from the origin is maximal or minimal. Now move the initial point $(a, 0)$ slightly by varying $\theta_0 = \epsilon$, leaving the direction ψ_0 unchanged. The first impact will move from $(0, b)$ to a nearby point and, after four segments, the ball will be close to $(a, 0)$. It is clear that, by tweaking the initial angle ψ_0 , the orbit can be made to return to the initial point after four segments, thereafter repeating in a period-4 orbit. In fact, this is guaranteed by Poncelet's closure theorem, discussed below. It follows from the theorem that there are period-4 orbits from every starting point. Indeed, Poncelet's theorem implies that there is a period- n orbit for any starting point and any $n \geq 3$.

Poncelet's closure theorem, also called Poncelet's porism, is discussed at length in [1]. A porism is a problem that has either an infinite number of solutions or none at all. Poncelet's porism states that, given any two conics, if there is a polygon with vertices on one conic (the outer conic) and all sides tangent to the other one (the inner conic), then an arbitrary point on the outer conic is a vertex of such a polygon. For elliptic billiards, this means that if there is a periodic orbit, then there exist orbits of the same period for any value of the initial position θ_0 . Berger [1] describes Poncelet's result — which is simply stated but difficult to prove — as “the most beautiful theorem on conics”.

Phase Portrait.

The billiard problem may be described as a Hamiltonian dynamical system. Between impacts with the boundary, the momentum is constant, so the equations of motion are

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{\mathbf{p}} = \mathbf{0},$$

representing motion in a straight line at constant speed. At each impact, the normal component of momentum is reversed while the tangential component is unchanged.

The dynamics are completely specified by considering the discrete mapping from one bounce to the next. Let (x_n, y_n) be the n -th point of impact, and (u_n, v_n) the velocity between points (x_n, y_n) and (x_{n+1}, y_{n+1}) . The slope of this segment is $m_n = v_n/u_n$. Given the values $(x, y; m)_n$ we can calculate the position of the next bounce:

$$x_{n+1} = -x_n - \frac{2a^2 m_n (y_n - m_n x_n)}{m_n^2 a^2 + b^2}, \quad y_{n+1} = y_n + m_n (x_{n+1} - x_n).$$

Then, defining $\nu_{n+1} = (a^2 y_{n+1}) / (b^2 x_{n+1})$, we get

$$m_{n+1} = \frac{2\nu_{n+1} - (1 - \nu_{n+1}^2)m_n}{(1 - \nu_{n+1}^2) + 2\nu_{n+1}m_n}.$$

We now have $(x, y; m)_{n+1}$. This discrete map can be iterated to generate the entire orbit.

We can determine the motion once the initial position $\theta_0 \in [-\pi, \pi]$ and *opening angle* $\phi_0 \in [-\pi, \pi]$ are known. Indeed, the continuous system may be represented by a discrete mapping from one impact to the next:

$$\theta_{n+1} = f(\theta_n, \phi_n), \quad \phi_{n+1} = g(\theta_n, \phi_n).$$

Plotting the representative points on a (θ, ϕ) diagram, we obtain a phase-portrait of the motion. Such portraits are shown in figure 3 for ten choices of initial conditions. Three represent box orbits; these fall within the separatrix, which includes the point $(\theta, \phi) = (0, 0)$. The remaining six are for clockwise and counter-clockwise loop motions.

There is a striking similarity between these phase plots and the phase portrait for a simple pendulum, which has librational and rotational motions separated by a homoclinic orbit that asymptotes to the unstable equilibrium point.

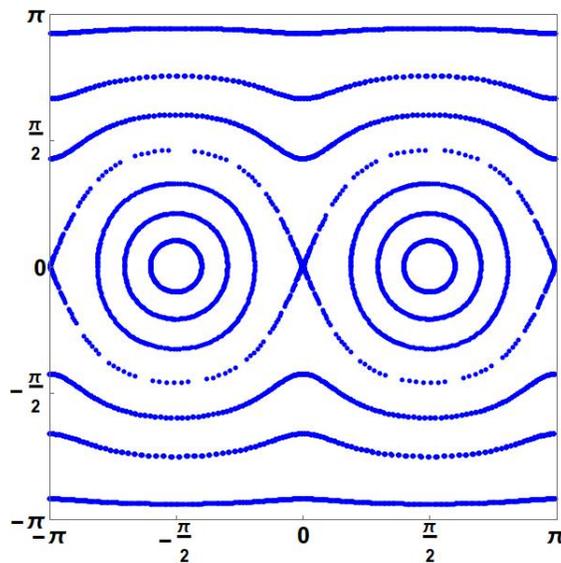


Figure 3. Phase plots of ten orbits. Horizontal axis is $\theta \in [-\pi, \pi]$ and vertical axis is $\phi \in [-\pi, \pi]$. The separatrix contains the point $(\theta, \phi) = (0, 0)$.

Constants of Motion.

Since the system has no dissipation and since energy is conserved at boundary impacts, the kinetic energy

$$T = \frac{1}{2}(\dot{x}^2 + \dot{y}^2)$$

is a constant of the motion. For a circular table it is convenient to use polar coordinates (r, ϑ) . Then the Lagrangian may be written

$$\mathcal{L} = \frac{1}{2}(\dot{r}^2 + r^2\dot{\vartheta}^2) - V(r).$$

Since ϑ is an ignorable coordinate (\mathcal{L} is independent of ϑ), the conjugate variable $p_\vartheta = \partial\mathcal{L}/\partial\dot{\vartheta} = r^2\dot{\vartheta}$, the angular momentum about the centre, is conserved.

For an elliptical table, the angular momentum about the centre is no longer conserved. However, there is another conserved quantity, as we shall now show. Since the boundary is an ellipse, it is convenient to use elliptic coordinates (ξ, η) :

$$x = f \cosh \xi \cos \eta, \quad y = f \sinh \xi \sin \eta. \quad (1)$$

The components of the velocity $\mathbf{v} = (u, v)$ are then

$$\begin{aligned} \dot{x} = u &= f \sinh \xi \cos \eta \dot{\xi} - f \cosh \xi \sin \eta \dot{\eta} \\ \dot{y} = v &= f \cosh \xi \sin \eta \dot{\xi} + f \sinh \xi \cos \eta \dot{\eta} \end{aligned}$$

and the radii from the center and foci are

$$\begin{aligned} \mathbf{r}_0 = (x, y) &= f(\cosh \xi \cos \eta, \sinh \xi \sin \eta) \\ \mathbf{r}_1 = (x - f, y) &= f(\cosh \xi \cos \eta - 1, \sinh \xi \sin \eta) \\ \mathbf{r}_2 = (x + f, y) &= f(\cosh \xi \cos \eta + 1, \sinh \xi \sin \eta) \end{aligned}$$

We can now compute the angular momenta $\mathbf{L}_1 = \mathbf{r}_1 \times \mathbf{v}$ and $\mathbf{L}_2 = \mathbf{r}_2 \times \mathbf{v}$ about the foci, from which it follows that

$$\mathbf{L}_1 \cdot \mathbf{L}_2 = L_1 L_2 = f^4 (\cosh^2 \xi - \cos^2 \eta) [(-\sin^2 \eta) \dot{\xi}^2 + (\sinh^2 \xi) \dot{\eta}^2]. \quad (2)$$

Since there is no cross-term containing $\dot{\xi}\dot{\eta}$, the quantity $L_1 L_2$ does not change its value at an impact, where $(\dot{\xi}, \dot{\eta}) \rightarrow (-\dot{\xi}, \dot{\eta})$. Clearly, since segments between bounces are linear, $\mathbf{L}_1 = \mathbf{r}_1 \times \mathbf{v}$ and $\mathbf{L}_2 = \mathbf{r}_2 \times \mathbf{v}$ are constant along each segment. Thus, $L_1 L_2$ is a constant of the motion [7]

The transformation to elliptic coordinates (1) can be inverted:

$$\xi = \operatorname{arccosh} \left(\frac{r_1 + r_2}{2f} \right), \quad \eta = \arccos \left(\frac{r_2 - r_1}{2f} \right).$$

Constant ξ corresponds to $r_1 + r_2$ constant, yielding an ellipse, while constant η corresponds to $r_1 - r_2$ constant, yielding a hyperbola. We note that

$$r_1 = f(\cosh \xi - \cos \eta), \quad r_2 = f(\cosh \xi + \cos \eta).$$

so that

$$r_1 r_2 = f^2 (\cosh^2 \xi - \cos^2 \eta) = f^2 (\sinh^2 \xi + \sin^2 \eta).$$

Since a constant product of distances from two points yields a Cassinian oval, we see that the contours of the function $(\cosh^2 \xi - \cos^2 \eta)$ are such ovals. The particular case $r_1 r_2 = f^2$ is the lemniscate of Bernoulli.

It is clear that for loop orbits, L_1 and L_2 are either both positive or both negative, so $L_1 L_2$ is positive. For box orbits, which pass between the foci, L_1 and L_2 are of

opposite signs. For the homoclinic orbit, passing through the foci, one or other of these components vanishes. Thus, L_1L_2 acts as a discriminant for the motion:

$$\text{Orbit is } \begin{cases} \text{Box type} & \text{if } L_1L_2 < 0 \\ \text{Homoclinic} & \text{if } L_1L_2 = 0 \\ \text{Loop type} & \text{if } L_1L_2 > 0. \end{cases}$$

Confocal Conics.

We consider the set of confocal conics

$$\frac{x^2}{a^2 + \lambda} + \frac{y^2}{b^2 + \lambda} = 1. \quad (3)$$

The case $\lambda = 0$ corresponds to the boundary of the billiard table. The range $\lambda \in (-b^2, +\infty)$ gives a family of ellipses all having the same foci, while the range $\lambda \in (-a^2, -b^2)$ gives a family of confocal hyperbolas orthogonal to the ellipses [1]. The two families cover the points of the plane twice, and provide an orthogonal coordinate system (see figure 4).

Suppose the initial segment lies on the line $y = mx + c$. The condition that this line is tangent to a confocal conic (3) leads to a quadratic equation whose discriminant must vanish, yielding

$$\lambda = \frac{c^2 - (m^2a^2 + b^2)}{1 + m^2}. \quad (4)$$

But since this is true for all segments of the trajectory, λ is an invariant of the motion.

We will show now how this purely geometric result may be interpreted dynamically. The angular momentum about the centre is $L_0 = (xv - yu)$. The angular momenta about the foci are then $L_1 = L_0 - fv$ and $L_2 = L_0 + fv$. The slope m is related to the components of velocity, $m = v/u$. Using this in (4) and recalling that $u^2 + v^2 = 1$, we find that

$$\lambda = L_0^2 - f^2v^2 - b^2 = L_1L_2 - b^2$$

Since L_1L_2 is conserved, so is λ , and every segment of the trajectory is tangent to the same conic. For $\lambda \in (-b^2, +\infty)$ it is an ellipse, while for $\lambda \in (-a^2, -b^2)$ it is a hyperbola.

3. Regularizing the Motion: Ballyards

The perfectly-reflecting boundaries imply instantaneous changes in momentum at each bounce. This corresponds to a potential well that is constant in the interior and has a step discontinuity at the boundary. We can approximate this behaviour by a high-order polynomial. For a circular table of radius a we take the potential energy to be $V(r) = V_0(r/a)^N$ where N is a large positive integer. This corresponds to a radially attractive force that is negligible in the interior but large in a narrow boundary zone.

With polar coordinates (r, ϑ) , the conjugate momenta are $p_r = \dot{r}$ and $p_\vartheta = r^2 \dot{\vartheta}$ and the Hamiltonian may be written

$$H = \frac{1}{2}(p_r^2 + p_\vartheta^2/r^2) + V(r)$$

Since this is independent of ϑ , the azimuthal momentum p_ϑ is a constant of the motion.

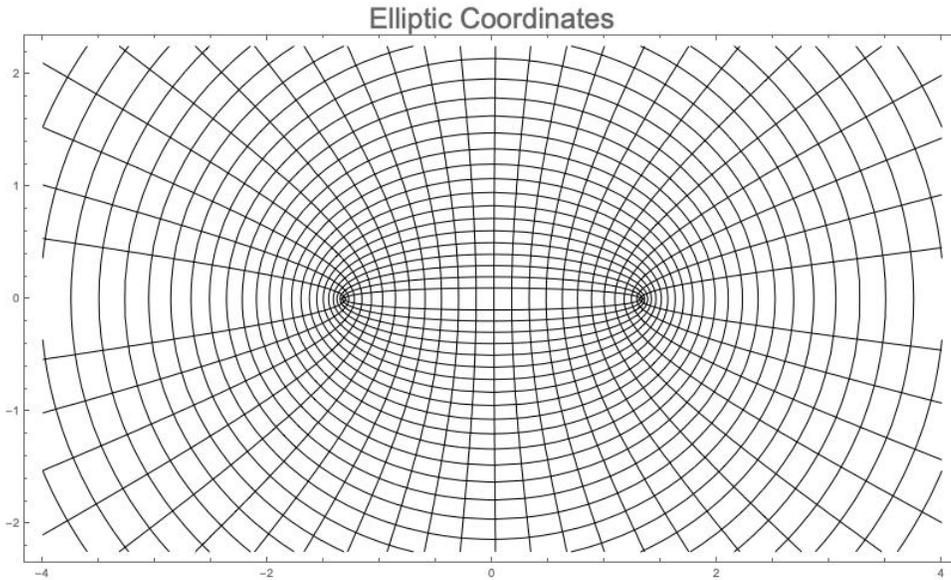


Figure 4. Elliptic coordinates (ξ, η) . The family of ellipses and the family of hyperbolas each cover the plane. All conics have the same foci.

For an elliptical table, we choose a potential function that rises sharply near the elliptical boundary. The kinetic energy may be written in elliptic coordinates:

$$T = \frac{1}{2}f^2(\cosh^2 \xi - \cos^2 \eta)(\dot{\xi}^2 + \dot{\eta}^2)$$

The Lagrangian then becomes

$$\mathcal{L} = T - V = \frac{1}{2}f^2(\cosh^2 \xi - \cos^2 \eta)(\dot{\xi}^2 + \dot{\eta}^2) - V(\xi, \eta). \quad (5)$$

The conjugate momenta in elliptic coordinates are

$$p_\xi = \frac{\partial T}{\partial \dot{\xi}} = f^2(\cosh^2 \xi - \cos^2 \eta)\dot{\xi}, \quad p_\eta = \frac{\partial T}{\partial \dot{\eta}} = f^2(\cosh^2 \xi - \cos^2 \eta)\dot{\eta}.$$

The Hamiltonian may now be written

$$H = \frac{p_\xi^2 + p_\eta^2}{2f^2(\cosh^2 \xi - \cos^2 \eta)} + V(\xi, \eta).$$

A Liouville integrable system.

In 1848 Joseph Liouville identified a broad class of dynamical systems that can be integrated [8, p. 67]. We confine attention here to systems with two degrees of freedom, with generalised coordinates (q_1, q_2) . If the kinetic and potential energies can be expressed in the form

$$T = \frac{1}{2}[\mathcal{U}_1(q_1) + \mathcal{U}_2(q_2)] \times [\mathcal{V}_1(q_1)\dot{q}_1^2 + \mathcal{V}_2(q_2)\dot{q}_2^2], \quad V = \frac{\mathcal{W}_1(q_1) + \mathcal{W}_2(q_2)}{\mathcal{U}_1(q_1) + \mathcal{U}_2(q_2)}. \quad (6)$$

then the solution can be solved in quadratures. This is proved in Appendix I, where explicit expressions for the solution integrals are given.

The kinetic energy term in the Lagrangian (5) is of the form (6) with

$$\mathcal{U}_1(\xi) = f^2 \cosh^2 \xi \quad \mathcal{U}_2(\eta) = -f^2 \cos^2 \eta \quad \mathcal{V}_1 \equiv 1 \quad \mathcal{V}_2 \equiv 1.$$

We can express the function $\mathcal{U}(\xi, \eta)$ in the form

$$\mathcal{U}(\xi, \eta) = \mathcal{U}_1(\xi) + \mathcal{U}_2(\eta) = r_1 r_2.$$

For small eccentricity, this depends only weakly on η .

If the potential energy function is of the form $V(\xi, \eta) = [\mathcal{W}_1(\xi) + \mathcal{W}_2(\eta)]/[\mathcal{U}_1(\xi) + \mathcal{U}_2(\eta)]$, the problem defined by the Lagrangian (5) is of Liouville type. The Hamiltonian may be written as

$$H = \left[\frac{\frac{1}{2}(p_\xi^2 + p_\eta^2) + (\mathcal{W}_1(\xi) + \mathcal{W}_2(\eta))}{f^2(\cosh^2 \xi - \cos^2 \eta)} \right].$$

It is a constant of the motion, equal in value to the total energy E .

We seek a potential surface that is close to constant within the elliptical region defined by $(x/a)^2 + (y/b)^2 = 1$ and that rises rapidly in a boundary zone. We define the potential surfaces by setting

$$\mathcal{W}_1(\xi) = V_N f^2 \cosh^N \xi \quad \mathcal{W}_2(\eta) = -V_N f^2 \cos^N \eta \quad (7)$$

where $V_N = V_0 \operatorname{sech}^N \xi_B$ with V_0 a constant, ξ_B the value of ξ defining the reference ellipse ($\cosh \xi_B = a/f = 1/e$) and N an even integer. The potential energy function is then

$$V(\xi, \eta) = \frac{\mathcal{W}_1(\xi) + \mathcal{W}_2(\eta)}{\mathcal{U}_1(\xi) + \mathcal{U}_2(\eta)} = V_N \left[\frac{\cosh^N \xi - \cos^N \eta}{\cosh^2 \xi - \cos^2 \eta} \right].$$

Two examples of potential energy surfaces are shown in figure 5.

For $N = 2$ we have

$$\mathcal{W} = V_2 f^2 (\cosh^2 \xi - \cos^2 \eta), \quad V \equiv 1$$

so that the potential energy is constant. The table is flat and the particle moves at constant velocity in a straight line; there is no finite boundary. For $N = 4$, we have

$$\mathcal{W} = V_4 f^2 (\cosh^4 \xi - \cos^4 \eta) \quad V = V_4 (\cosh^2 \xi + \cos^2 \eta)$$

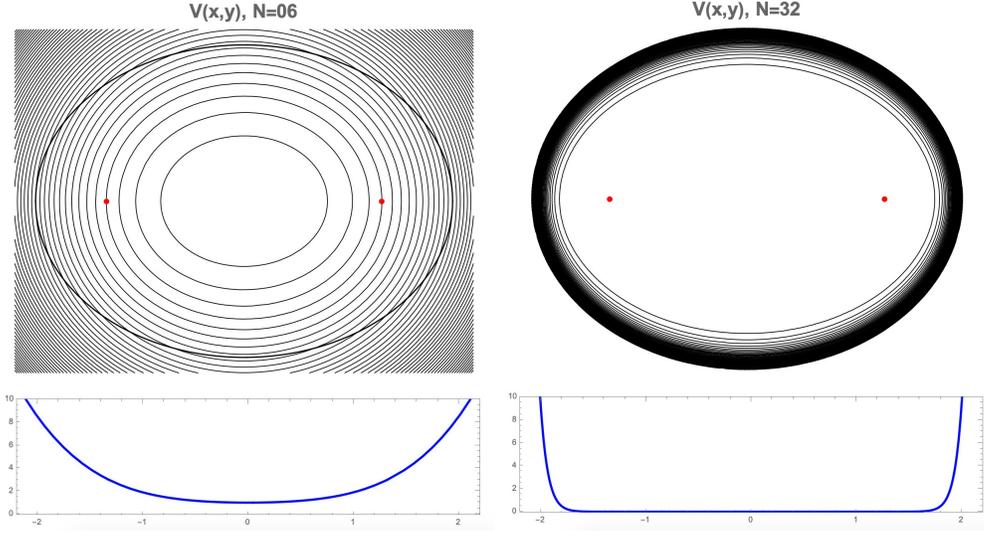


Figure 5. Potential energy surface for $N = 6$ (left) and $N = 32$ (right). Cross-sections along the major axis are also shown.

so the potential energy is proportional to $x^2 + y^2$. In this case, which corresponds to an isotropic simple harmonic oscillator in two dimensions, the orbits are closed ellipses.

The system is integrable for any value of N but, for definiteness, we let $N = 6$:

$$\mathcal{W} = V_6 f^2 (\cosh^6 \xi - \cos^6 \eta) = V_6 f^2 (\cosh^2 \xi - \cos^2 \eta) (\cosh^4 \xi + \cosh^2 \xi \cos^2 \eta + \cos^4 \eta).$$

The Hamiltonian function for this system is

$$H = \frac{\frac{1}{2}(p_\xi^2 + p_\eta^2)}{f^2(\cosh^2 \xi - \cos^2 \eta)} + V_6 (\cosh^4 \xi + \cosh^2 \xi \cos^2 \eta + \cos^4 \eta).$$

From Equations (A.1) and (A.2) in Appendix I we have

$$\frac{1}{2}\mathcal{U}^2 \dot{\xi}^2 = E\mathcal{U}_1(\xi) - \mathcal{W}_1(\xi) + \gamma_1 \quad (8)$$

$$\frac{1}{2}\mathcal{U}^2 \dot{\eta}^2 = E\mathcal{U}_2(\eta) - \mathcal{W}_2(\eta) + \gamma_2 \quad (9)$$

where γ_1 and γ_2 are constants of integration, and $\gamma_1 + \gamma_2 = 0$.

We partition the energy as $E = E_1 + E_2$, where

$$E_1 = \frac{1}{2}\mathcal{U}(\xi, \eta) \dot{\xi}^2 + \frac{\mathcal{W}_1(\xi)}{\mathcal{U}(\xi, \eta)} \quad \text{and} \quad E_2 = \frac{1}{2}\mathcal{U}(\xi, \eta) \dot{\eta}^2 + \frac{\mathcal{W}_2(\eta)}{\mathcal{U}(\xi, \eta)}.$$

Then the constants of motion can be written

$$\gamma_1 = \mathcal{U}E_1 - E\mathcal{U}_1 \quad \gamma_2 = \mathcal{U}E_2 - E\mathcal{U}_2.$$

Note that the components E_1 and E_2 of energy are not constants.

Equations (8) and (9) can be integrated:

$$\int^{\xi} \frac{\mathcal{U}_1(\xi) d\xi}{\sqrt{2[E\mathcal{U}_1(\xi) - \mathcal{W}_1(\xi) + \gamma_1]}} = \int^t dt \quad (10)$$

$$\int^{\eta} \frac{\mathcal{U}_2(\eta) d\eta}{\sqrt{2[E\mathcal{U}_2(\eta) - \mathcal{W}_2(\eta) + \gamma_2]}} = \int^t dt \quad (11)$$

Analytical evaluation of these integrals may or may not be possible. For the case $N = 4$, the solutions can be expressed in terms of elliptical integrals.

For the case $N = 6$, we get:

$$\int_{\xi_0}^{\xi} \frac{f^2 \cosh^2 \xi d\xi}{\sqrt{2[Ef^2 \cosh^2 \xi - V_6 f^2 \cosh^6 \xi + \gamma]}} = t - t_0 \quad (12)$$

$$\int_{\eta_0}^{\eta} \frac{-f^2 \cos^2 \eta d\eta}{\sqrt{2[-Ef^2 \cos^2(\eta) + V_6 f^2 \cos^6 \eta - \gamma]}} = t - t_0 \quad (13)$$

where we have written $\gamma = \gamma_1 = -\gamma_2$. These do not have solutions in closed form.

The Angular Momentum Integral.

For the billiard dynamics, the product of the angular momenta about the foci, $L_1 L_2$, is an integral of the motion. We seek a corresponding integral for the general case N of a ballyard. In elliptical coordinates, we can write

$$L_1 L_2 = f^2 \mathcal{U}(\xi, \eta) [\sinh^2 \xi \dot{\eta}^2 - \sin^2 \eta \dot{\xi}^2]$$

We use (8) and (9) to substitute for $\dot{\xi}^2$ and $\dot{\eta}^2$ and, after some manipulation, find that

$$L_1 L_2 + \frac{2f^2(\sinh^2 \xi \mathcal{W}_2 - \sin^2 \eta \mathcal{W}_1)}{\mathcal{U}} = -2(f^2 E + \gamma). \quad (14)$$

Since the right hand side is constant, the same is true of the left hand side. If we define the quantity

$$\Lambda(\xi, \eta) = \frac{2f^2[\sinh^2 \xi \mathcal{W}_2(\eta) - \sin^2 \eta \mathcal{W}_1(\xi)]}{\mathcal{U}(\xi, \eta)}$$

then the relationship (14) becomes

$$\mathbb{L} \equiv L_1 L_2 + \Lambda = -2(f^2 E + \gamma) = \text{constant}, \quad (15)$$

and \mathbb{L} is an integral of the motion.

It is easy to show that $\Lambda(\xi, \eta)$ is equal to $\Lambda_0 = -2f^2 V_N$, a constant, on the major axis ($y = 0$). This implies that $L_1 L_2$ is also constant there. But it is clear on physical grounds that $L_1 L_2 < 0$ on the inter-focal segment $-f < x < f$ and $L_1 L_2 > 0$ when $x < -f$ or $x > f$. Therefore, if the trajectory passes through the inter-focal segment, it can never cross the major axis outside it. If it crosses outside, it cannot pass between the foci. If a trajectory passes through a focus then $L_1 L_2$ must vanish there. It can

never cross the axis at points other than the foci. This special case of motions $\mathbb{L} = \Lambda_0$ separates the box orbits from the loop orbits.

We note that as $N \rightarrow \infty$,

$$\mathcal{W}_1 = O\left(\frac{\cosh \xi}{\cosh \xi_B}\right)^N \quad \text{and} \quad \mathcal{W}_2 = O\left(\frac{1}{\cosh \xi_B}\right)^N$$

so that, for $|\xi| < |\xi_B|$, we have $\lim_{N \rightarrow \infty} \mathbb{L} = L_1 L_2$. That is, the integral \mathbb{L} corresponds in this limit to the product of the angular moments about the foci, the quantity conserved for motion on a billiard.

Note that (15) may be written $\mathbb{L} + 2f^2 E + 2\gamma = 0$, so the three constants E , γ and \mathbb{L} are not independent. Given that the system has two degrees of freedom, there must be another independent integral of the motion.

4. Numerical Results

Numerical integrations confirm the dichotomy between box and loop orbits for the ballyard potentials. A large number of numerical experiments were performed with $N = 6$, and several for larger values of N . Figure 6 shows a typical box orbit (left panel) and loop orbit (right panel).

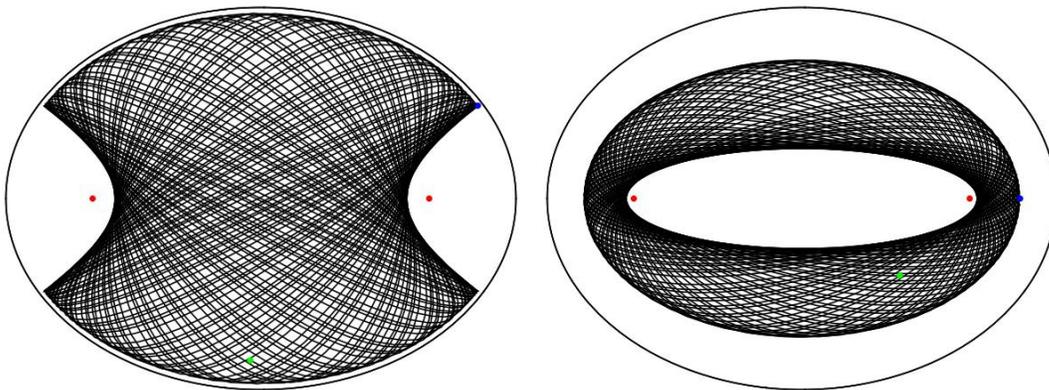


Figure 6. Numerical solutions for $N = 6$. Left: Box orbit, crossing the major axis between the foci. Right: Loop orbit, crossing the major axis outside the foci.

For the special case of orbits passing through the foci, the equations in (ξ, η) coordinates are singular. A re-coding using cartesian coordinates enabled numerical integrations along homoclinic orbits. A typical orbit is shown in figure 7. The orbit rapidly approaches an oscillation along the major axis.

In general, the orbits are dense in a region bounded by an ellipse and hyperbola (box orbits) or between two ellipses (loop orbits). However, for delicately chosen initial conditions, the solutions may be periodic. A periodic box orbit is shown in figure 8 (left panel) and a pure elliptic loop orbit is shown in figure 8 (right panel).

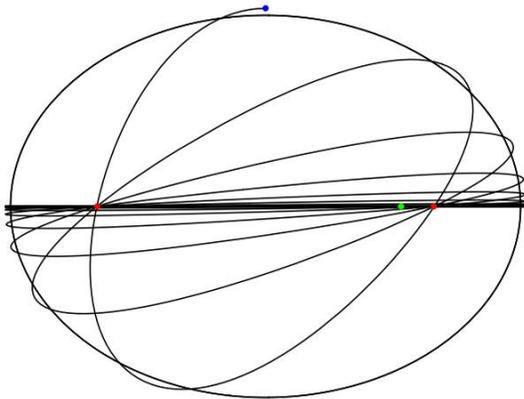


Figure 7. Numerical solution ($N = 6$) for the homoclinic orbit, crossing the major axis at the foci.

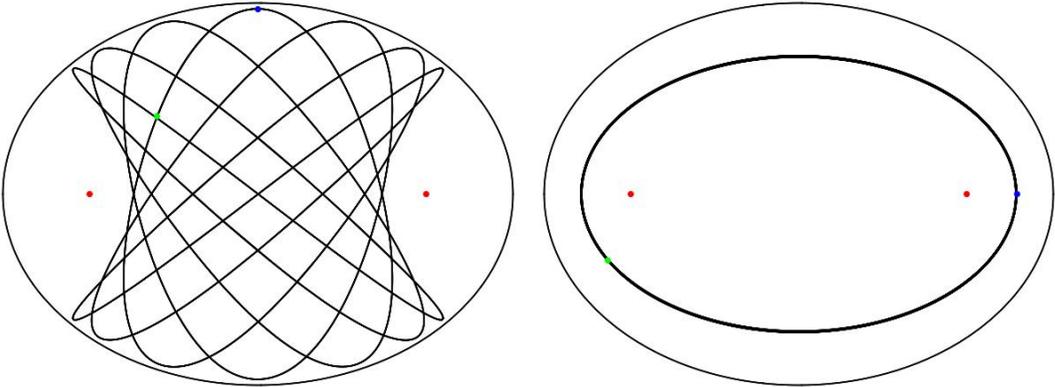


Figure 8. Numerical solutions for $N = 6$. Left: Periodic box orbit Right: Pure elliptic loop orbit.

5. Conclusion

We have reviewed the well-known problem of motion on an elliptical billiard, and shown how the dynamics may be regularized by approximating the flat-bedded, hard-edged billiard by a smooth surface, a ballyard. The ballyard Lagrangians are of Liouville type and so are completely integrable. A new constant of the motion (\mathbb{L}) was found, allowing us to show that the trajectories split into boxes and loops, separated by homoclinic orbits. The discriminant that determines the character of the solution is the sign of L_1L_2 on the major axis.

Appendix I: Integrable Systems of Liouville Type

Following [8, §43], we show how a system of Liouville type, where the kinetic and potential energies can be expressed in the form

$$T = \frac{1}{2}[\mathcal{U}_1(q_1) + \mathcal{U}_2(q_2)][\dot{q}_1^2 + \dot{q}_2^2], \quad V = \frac{[\mathcal{W}_1(q_1) + \mathcal{W}_2(q_2)]}{[\mathcal{U}_1(q_1) + \mathcal{U}_2(q_2)]},$$

can be solved in quadratures. Letting $\mathcal{U}(q_1, q_2) = \mathcal{U}_1(q_1) + \mathcal{U}_2(q_2)$ we have

$$T = \frac{1}{2}\mathcal{U}[\dot{q}_1^2 + \dot{q}_2^2], \quad V = \frac{1}{\mathcal{U}}[\mathcal{W}_1(q_1) + \mathcal{W}_2(q_2)].$$

The Lagrangian equation for the coordinate q_1 is

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_1} \right) - \frac{\partial T}{\partial q_1} = -\frac{\partial V}{\partial q_1}$$

or, more explicitly,

$$\frac{d}{dt} (\mathcal{U}\dot{q}_1) - \frac{1}{2} \frac{\partial \mathcal{U}}{\partial q_1} [\dot{q}_1^2 + \dot{q}_2^2] = -\frac{\partial V}{\partial q_1}$$

Multiplying both sides by $2\mathcal{U}\dot{q}_1$, this becomes

$$\frac{d}{dt} (\mathcal{U}^2\dot{q}_1^2) - \mathcal{U}\dot{q}_1 \frac{\partial \mathcal{U}}{\partial q_1} [\dot{q}_1^2 + \dot{q}_2^2] = -2\mathcal{U}\dot{q}_1 \frac{\partial V}{\partial q_1}$$

But the expression for total energy E implies $\mathcal{U}(\dot{q}_1^2 + \dot{q}_2^2) = 2(E - V)$. Thus the q_1 -equation may be written

$$\begin{aligned} \frac{d}{dt} (\mathcal{U}^2\dot{q}_1^2) &= 2(E - V)\dot{q}_1 \frac{\partial \mathcal{U}}{\partial q_1} - 2\mathcal{U}\dot{q}_1 \frac{\partial V}{\partial q_1} \\ &= 2\dot{q}_1 \left[(E - V) \frac{\partial \mathcal{U}}{\partial q_1} + \mathcal{U} \frac{\partial}{\partial q_1} (E - V) \right] \\ &= 2\dot{q}_1 \frac{\partial}{\partial q_1} [(E - V)\mathcal{U}] = 2\dot{q}_1 \frac{\partial}{\partial q_1} [E\mathcal{U}_1(q_1) - \mathcal{W}_1(q_1)] \\ &= 2 \frac{d}{dt} [E\mathcal{U}_1(q_1) - \mathcal{W}_1(q_1)]. \end{aligned}$$

So we can write

$$\frac{d}{dt} \left[\frac{1}{2} \mathcal{U}^2 \dot{q}_1^2 \right] = \frac{d}{dt} [E\mathcal{U}_1(q_1) - \mathcal{W}_1(q_1)].$$

Integrating this we get

$$\frac{1}{2} \mathcal{U}^2 \dot{q}_1^2 = E\mathcal{U}_1(q_1) - \mathcal{W}_1(q_1) + \gamma_1 \tag{A.1}$$

where γ_1 is a constant of integration. We obtain a similar equation

$$\frac{1}{2} \mathcal{U}^2 \dot{q}_2^2 = E\mathcal{U}_2(q_2) - \mathcal{W}_2(q_2) + \gamma_2 \tag{A.2}$$

for the q_2 -component. Adding these together we find that $\gamma_1 + \gamma_2 = 0$.

Solution by Quadratures.

We partition the energy as $E = E_1 + E_2$, where

$$E_1 = \frac{1}{2}\mathcal{U}(q_1, q_2)\dot{q}_1^2 + \frac{\mathcal{W}_1(q_1)}{\mathcal{U}(q_1, q_2)} \quad \text{and} \quad E_2 = \frac{1}{2}\mathcal{U}(q_1, q_2)\dot{q}_2^2 + \frac{\mathcal{W}_2(q_2)}{\mathcal{U}(q_1, q_2)}.$$

Then the constants of motion can be written

$$\begin{aligned} \gamma_1 &= \mathcal{U}E_1 - E\mathcal{U}_1 \\ \gamma_2 &= \mathcal{U}E_2 - E\mathcal{U}_2 \end{aligned}$$

Note that the components E_1 and E_2 of energy are not constants.

Equations (A.1) and (A.2) can be integrated:

$$\int^{q_1} \frac{\mathcal{U}_1(q_1)dq_1}{\sqrt{2[E\mathcal{U}_1(q_1) - \mathcal{W}_1(q_1) + \gamma_1]}} = \int^t dt \quad (\text{A.3})$$

$$\int^{q_2} \frac{\mathcal{U}_2(q_2)dq_2}{\sqrt{2[E\mathcal{U}_2(q_2) - \mathcal{W}_2(q_2) + \gamma_2]}} = \int^t dt \quad (\text{A.4})$$

Analytical evaluation of these integrals may or may not be possible.

References

- [1] Berger, Marcel, 2005: *Geometry Revealed: a Jacob's Ladder to Higher Geometry*. Springer, 831pp. ISBN: 978-0-5217-4449-2.
- [2] Berry, Michael, 1981: Regularity and chaos in classical mechanics, illustrated by three deformations of a circular 'billiard'. *Eur. J. Phys.*, **2**, 91–102.
- [3] Binney, James and Scott Tremaine, 2008: *Galactic Dynamics* Princeton Univ. Press, Princeton and Oxford. 885pp.
- [4] Birkhoff, George D., 1927. *Dynamical Systems*. Amer. Math. Soc.. Revised Edn. 1966, Colloquium Publications Vol. 9, Amer. Math. Soc.
- [5] Coriolis, G.-G., 1835: Gaspard-Gustave de Coriolis, *Théorie Mathématique des Effets du Jeu de Billiard*, Carilian-Goeury, Paris.
- [6] Sommerfeld, Arnold, 1964: *Lectures on Theoretical Physics*. Vol. 1: Mechanics. Academic Press, 304 pp. ISBN: 978-0-1265-4670-5
- [7] Tabachnikov, Serge, 2005: *Geometry and Billiards*. Amer. Math. Soc., Student Mathematical Library, vol. 30. ISBN: 978-0-8218-3919-5 <http://www.ams.org/publications/authors/books/postpub/stml-30>
- [8] Whittaker, E.T., 1937 *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*. 4th Edn., Cambridge Univ. Press, 456pp.