Further thoughts on a focusing property of the ellipse

J.B. Wilker

1 Introduction

If a billiard ball is fired from one focus of an elliptical table, its trajectory will pass alternately through the two foci of the table and will quickly become indistinguishable from a repeated tracing of the major axis of the table. In 3-dimensions, an ellipsoid of revolution has focusing properties similar to those of an ellipse. Frantz [1] has used this fact, together with geometrical optics, to follow a spherical wave emitted from one focus of a reflecting ellipsoidal cavity. His analysis predicts a growth in the density of energy tracing the major axis of the cavity which is exponential in time or, more precisely, in the number of reflections which have occurred to the wave. This dramatic prediction is mathematically sound but may not be physically correct since one of the caveats of geometrical optics is to avoid singular situations where spherical disturbances converge to a point (See e.g. [2] p.p. 6, 219). In spite of the fact that the energy interpretation of the model may be open to question, it is interesting to pursue the mathematics because it is attractive and because it does admit other interpretations, for example probabilistic ones.

In this paper we shall derive a probability density at all points and a related estimate concerning the concentration of probability along the major axis. Both of these results generalise the earlier one on density in the direction of the major axis. In order to give a full treatment of the 2-dimensional situation, which is interesting in its own right, we base our approach on the return map to a small circle centred at the first focus. This map turns out to be a Möbius transformation (of hyperbolic

Received by the editors May 1994
Communicated by J. Thas
AMS Mathematics Subject Classification : 78A05, 51N25.
Keywords : reflecting ellipsoidal cavity, return map, Möbius transformation.

type) of the circle in question and so our subsequent results have a helpful conceptual base. To reconcile our results with those of Frantz, one must bear in mind that the n^{th} return occurs after $2n$ reflections.

2 The return map

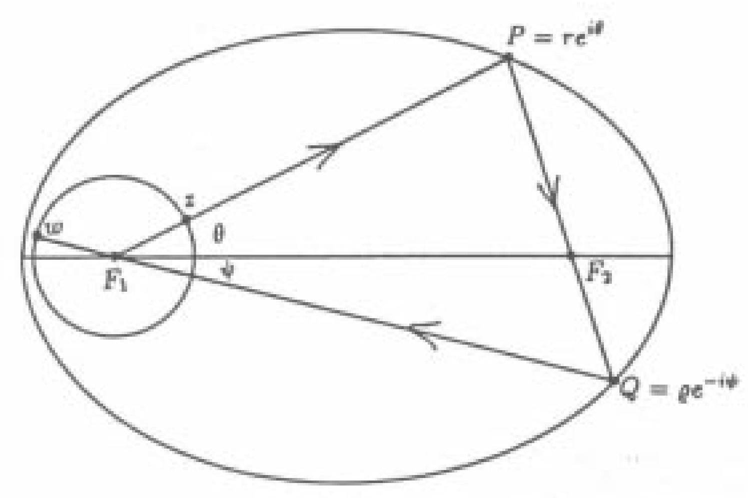
A ray emanating from one focus of an ellipse is reflected by the ellipse in such a way as to pass through its other focus. After $2n$ reflections of this kind, an initial ray from the first focus F_1 is related to a final ray from F_1 and the point z where the initial ray meets a circle centred at F_1 is related to the point w where the final ray meets this circle. The mapping $z \to w$ is called the n^{th} return map. We introduce complex coordinates with $|z| = |w| = 1$ to show that

Theorem 1 *The n^{th} return map is given by the Möbius transformation*

$$w = \frac{(\cosh n\delta)z - \sinh n\delta}{-(\sinh n\delta)z + \cosh n\delta}$$

where δ is related to the eccentricity ε of the ellipse by the formula $\cosh \delta = \frac{1+\varepsilon^2}{1-\varepsilon^2}$.

Proof: The usual polar equation of an ellipse with eccentricity ε and semi latus rectum ℓ is $\frac{\ell}{r} = 1 - \varepsilon \cos \theta$. The first focus F_1 of this ellipse is the origin; its major axis lies along the ray $\theta = 0$; and the distance between its foci is $F_1F_2 = \frac{2\varepsilon\ell}{1-\varepsilon^2}$. In this context, it is natural to think of the return map as acting on the circle $|z| = 1$. Our result for $n > 1$ will follow from familiar properties of Möbius transformations once we establish it for $n = 1$.



The first return map $z \to w$.

The figure shows how $z = e^{i\theta}$ is carried by the first return map to $w = -e^{-i\psi}$. Two expressions for the area of triangle F_1PQ yield the equation

$$\tfrac{1}{2}r\varrho\sin(\theta+\psi) = \tfrac{1}{2}F_1F_2(r\sin\theta + \varrho\sin\psi).$$

With the help of the equation of the ellipse we transform this to

$$\sin(\theta+\psi) = \frac{2\varepsilon}{1-\varepsilon^2}\left(\frac{\ell}{\varrho}\sin\theta + \frac{\ell}{r}\sin\psi\right) = \frac{2\varepsilon}{1-\varepsilon^2}(\sin\theta + \sin\psi - \varepsilon\sin(\theta+\psi))$$

and hence to

$$(1+\varepsilon^2)\sin(\theta+\psi) = 2\varepsilon(\sin\theta + \sin\psi).$$

Direct substitution for $e^{i\theta}$ and $e^{i\psi}$ followed by multiplication by $2i$ yields

$$(1+\varepsilon^2)\left(-\frac{z}{w} + \frac{w}{z}\right) = 2\varepsilon\left(z - \frac{1}{z} - \frac{1}{w} + w\right)$$

and hence

$$(1+\varepsilon^2)(w^2 - z^2) = 2\varepsilon zw(z+w) - 2\varepsilon(z+w).$$

The figure shows that the return map fixes the points 1 and -1. Otherwise z and w both lie on the open semi circle with positive imaginary part or on the one with negative imaginary part. It follows that $z + w \neq 0$ and hence our equation can be simplified to

$$(1+\varepsilon^2)(w-z) = 2\varepsilon zw - 2\varepsilon$$

or

$$w = \frac{(1+\varepsilon^2)z - 2\varepsilon}{-2\varepsilon z + (1+\varepsilon^2)}\ .$$

By normalizing the determinant of the coefficient matrix to 1 we put this in the form

$$w = \frac{(\cosh\delta)z - \sinh\delta}{-(\sinh\delta)z + \cosh\delta}\ ,$$

where $\cosh\delta = \frac{1+\varepsilon^2}{1-\varepsilon^2}$.

For later developments in 3-dimensions it is useful to have an expression for the return map in terms of the arguments of z and w. If we write $z = e^{i\varphi}$ and $w = e^{i\varphi'}$ then we have

Corollary 1 *The* n^{th} *return map can be written*

$$\cos\varphi' = \frac{(\cosh 2n\delta)\cos\varphi - \sinh 2n\delta}{-(\sinh 2n\delta)\cos\varphi + \cosh 2n\delta}\ .$$

Proof: This expression can be verified by a straight forward calculation based on the formula $\cos\varphi' = \frac{1}{2}(w + w^{-1})$ and on identities for the hyperbolic trigonometric functions.

It is somewhat surprising that the Möbius transformation in Corollary 1 is the square of the Möbius transformation in Theorem 1. A proof of Corollary 1 that "explains" this can be based on the fact that stereographic projection of the unit circle from -1 to the imaginary axis is given by $e^{i\varphi} \to i \tan \frac{\varphi}{2}$. The original Möbius transformation on the unit circle is conjugated by this transformation to the dilatation $\tan \frac{\varphi'}{2} = e^{2n\delta} \tan \frac{\varphi}{2}$ on the imaginary axis. Corollary 1 now follows from the half angle formula $\cos \varphi = \dfrac{1 - \tan^2 \frac{\varphi}{2}}{1 + \tan^2 \frac{\varphi}{2}}$ which involves squaring in an essential way.

3 Probability densities

The n^{th} return map $w = h(z)$ of Theorem 1 transforms an initial probability density on the unit circle into a final probability density. If the initial density is a uniform one we can take it to be identically equal to 1, in non-normalized form. Then the transformed density, also in non-normalized form, is $\varrho(n, \delta, \varphi)$, where

$$\varrho(n, \delta, \varphi') d\varphi' = 1 d\varphi.$$

This non-normalized form saves a factor $\frac{1}{2\pi}$ in later formulae and invites density interpretations beyond our probabilistic one.

Theorem 2 *For the n^{th} return map on the circle,*

$$\varrho(n, \delta, \varphi) = \frac{1}{\cosh 2n\delta + (\sinh 2n\delta) \cos \varphi} .$$

Proof: The defining equation shows that

$$\varrho(n, \delta, \varphi') = \frac{d\varphi}{d\varphi'} = \frac{d\varphi}{dz} \cdot \frac{dz}{dw} \cdot \frac{dw}{d\varphi'} .$$

For the factors on the right we have

$$\begin{aligned} \frac{d\varphi}{dz} &= \left(\frac{dz}{d\varphi}\right)^{-1} = (iz)^{-1}, \\ \frac{dz}{dw} &= \frac{d}{dw}\left(h^{-1}(w)\right) = ((\sinh n\delta) w + \cosh n\delta)^{-2}, \\ \frac{dw}{d\varphi'} &= iw. \end{aligned}$$

It follows that

$$\begin{aligned} \varrho(n, \delta, \varphi') &= w \cdot \frac{1}{z} \cdot ((\sinh n\delta) w + \cosh n\delta)^{-2} \\ &= w \cdot \frac{(\sinh n\delta) w + \cosh n\delta}{(\cosh n\delta) w + \sinh n\delta} \cdot ((\sinh n\delta) w + \cosh n\delta)^{-2} \end{aligned}$$

$$= \frac{w}{(\sinh n\delta)(\cosh n\delta)(w^2+1)+(\cosh^2 n\delta+\sinh^2 n\delta)w}$$
$$= \frac{1}{(\sinh 2n\delta)\cos\varphi' + \cosh 2n\delta} \ .$$

In the proof of Theorem 2 we have used the fact that the inverse of our Möbius transformation can be written down immediately by changing n to $-n$ in the original formula. Also, we have used without comment the fact that the derivative of $f(z) = \frac{az+b}{cz+d}$, $ad-bc=1$, is $f'(z) = (cz+d)^{-2}$.

In dealing with the 3-dimensional situation it is convenient to use spherical polar coordinates. Then with the usual colatitude interpretation of φ we can take the n^{th} return map on the sphere to be the mapping $\cos\varphi' = g(\cos\varphi)$ of Corollary 1. This transforms an initial area density 1 on the sphere to a final density $\sigma(n,\delta,\varphi)$ where

$$\sigma(n,\delta,\varphi')\sin\varphi' d\varphi' d\theta' = 1 \sin\varphi d\varphi d\theta.$$

Theorem 3 *For the n^{th} return map on the sphere,*

$$\sigma(n,\delta,\varphi) = \frac{1}{(\cosh 2n\delta + (\sinh 2n\delta)\cos\varphi)^2} \ .$$

Proof: Since the return map involves only φ, the defining equation shows that

$$\sigma(n,\delta,\varphi') = \frac{\sin\varphi}{\sin\varphi'}\cdot\frac{d\varphi}{d\varphi'} \ .$$

By a happy accident, the right hand side can be computed by simply differentiating the equation $\cos\varphi = g^{-1}(\cos\varphi')$ with respect to φ'.

4 Exponential concentration of probability

Theorem 2 shows that the probability density outward bound along the major axis of the ellipse grows exponentially:

$$\varrho(n,\delta,\pi) = (\cosh 2n\delta + (\sinh 2n\delta)\cos\pi)^{-1} = (\cosh 2n\delta - \sinh 2n\delta)^{-1} = e^{2n\delta} \ .$$

Theorem 3 gives a companion formula for the ellipsoid of revolution in 3-dimensions:

$$\sigma(n,\delta,\pi) = (\cosh 2n\delta + (\sinh 2n\delta)\cos\pi)^{-2} = e^{4n\delta} \ .$$

This can be reconciled with Frantz's Proposition 3 by observing that his factor μ is the same as our $e^{2\delta}$.

Since we have an explicit formula for the probability density in any direction we can examine the asymptotic behaviour of the probability of any event. We shall see that in 2-dimensions and also in 3-dimensions the probability of any event that avoids an arbitrarily small neighbourhood of $\varphi = \pi$ must eventually tend to zero. We begin with 2-dimensions.

Theorem 4 *Let* $\varrho(\varphi) = \frac{1}{a+b\cos\varphi}$ *with* $a > b > 0$ *and let* $\lambda > 1$*. Then, if* $E = \{\varphi : |\varphi - \pi| \geq \frac{\pi}{\sqrt{\lambda}}\}$,

$$\int_E \varrho(\varphi) d\varphi < \frac{2\pi\lambda}{a+b} \ .$$

Proof: We shall apply the inequality $\frac{2}{\pi}\theta \leq \sin\theta$ which holds for $0 \leq \theta \leq \frac{\pi}{2}$.

Suppose $\frac{\pi}{\sqrt{\lambda}} \leq |\varphi - \pi|$. Then

$$\frac{1}{\lambda} \leq \left(\frac{2}{\pi}|\frac{\pi - \varphi}{2}|\right)^2 \leq \sin^2\frac{\pi-\varphi}{2} = \cos^2\frac{\varphi}{2}$$

and hence

$$1 - \cos\varphi = 2\left(1 - \cos^2\frac{\varphi}{2}\right) \leq 2\left(1 - \frac{1}{\lambda}\right) \leq \frac{a+b}{b} \cdot \frac{\lambda - 1}{\lambda} \ .$$

From this it follows in turn that

$$\begin{aligned} \lambda b - \lambda b\cos\varphi &\leq \lambda a + \lambda b - (a+b), \\ a + b &\leq \lambda(a + b\cos\varphi), \qquad \text{and} \\ \varrho(\varphi) &\leq \frac{\lambda}{a+b} \ . \end{aligned}$$

We can now estimate the integral as

$$\int_E \varrho(\varphi) d\varphi \leq \int_E \frac{\lambda}{a+b} d\varphi < \int_0^{2\pi} \frac{\lambda}{a+b} d\varphi = \frac{2\pi\lambda}{a+b} \ .$$

Corollary 2 *In 2-dimensions the probability becomes concentrated at* $\varphi = \pi$*. We have the quantitative estimate*

$$\int\limits_{|\varphi-\pi|<\pi e^{-n\delta/2}} \varrho(n,\delta,\varphi) d\varphi > 2\pi(1 - e^{-n\delta}).$$

Proof: Since there is conservation from return to return,

$$\int_0^{2\pi} \varrho(n,\delta,\varphi) d\varphi = \int_0^{2\pi} \varrho(0,\delta,\varphi) d\varphi = \int_0^{2\pi} d\varphi = 2\pi.$$

For any event E with complement E' we have

$$\int_{E'} \varrho(n,\delta,\varphi) d\varphi = 2\pi - \int_E \varrho(n,\delta,\varphi) d\varphi.$$

To obtain the desired result, we estimate $\int_E \varrho(n,\delta,\varphi) d\varphi$ by applying Theorem 4 to $\varrho(\varphi) = \frac{1}{\cosh 2n\delta + (\sinh 2n\delta)\cos\varphi}$ with $\lambda = e^{n\delta}$ so that $\frac{\lambda}{a+b} = \frac{e^{n\delta}}{e^{2n\delta}} = e^{-n\delta}$.

The argument in Theorem 4 makes it clear that if $\sigma(\varphi) = \frac{1}{(a+b\cos\varphi)^2}$ with $a > b > 0$ then, for $\lambda > 1$, the inequality $\frac{\pi}{\sqrt{\lambda}} \leq |\varphi - \pi|$ implies $\frac{1}{a+b\cos\varphi} \leq \frac{\lambda}{a+b}$ and hence $\sigma(\varphi) \leq \left(\frac{\lambda}{a+b}\right)^2$. It follows that for subsets E of the sphere defined by $|\varphi - \pi| \geq \frac{\pi}{\sqrt{\lambda}}$ we have

$$\iint_E \sigma(\varphi) dS < 4\pi \left(\frac{\lambda}{a+b}\right)^2.$$

As a companion to Corollary 2 we offer

Corollary 3 *In 3-dimensions the probability also becomes concentrated at $\varphi = \pi$. The quantitative estimate is*

$$\iint_{|\varphi-\pi|<\pi e^{-n\delta/2}} \sigma(n,\delta,\varphi) dS > 4\pi(1 - e^{-2n\delta}).$$

Proof: The method is the same as in the proof of Corollary 2. We again choose $\lambda = e^{n\delta}$ and find $\frac{\lambda}{a+b} = e^{-n\delta}$ but this time the final estimate is in terms of $\left(\frac{\lambda}{a+b}\right)^2 = e^{-2n\delta}$.

References

[1] Marc Frantz, A focusing property of the ellipse, *Amer. Math. Monthly* **101** (1994) 250–258.

[2] R.S. Longhurst, *Geometrical and Physical Optics*, 3rd edition, Longman, London (1973).

Physical Sciences Division
Scarborough College
University of Toronto
Scarborough, Ontario M1C 1A4
Canada