

DETERMINISTIC APPROXIMATIONS OF RANDOM REFLECTORS

OMER ANGEL, KRZYSZTOF BURDZY, AND SCOTT SHEFFIELD

ABSTRACT. Within classical optics, one may add microscopic “roughness” to a macroscopically flat mirror so that parallel rays of a given angle are reflected at different outgoing angles. Taking the limit (as the roughness becomes increasingly microscopic) one obtains a flat surface that reflects *randomly*, i.e., the transition from incoming to outgoing ray is described by a probability kernel (whose form depends on the nature of the microscopic roughness).

We consider two-dimensional optics (a.k.a. billiards) and show that *every* random reflector on a line that satisfies a necessary measure-preservation condition (well established in the theory of billiards) can be approximated by deterministic reflectors in this way.

1. INTRODUCTION

This article addresses the question of which random reflectors can be approximated by piecewise smooth surfaces that reflect light according to the classical rule of specular reflection which says that the angle of reflection is equal to the angle of incidence. One of the basic results on billiards says that a certain measure on the space of pairs consisting of location and angle of reflection is preserved by every reflecting surface (see [17, Thm. 3.1] or [3, Lemma 2.35]). Our main result, Theorem 2.3, shows that, except for this universal restriction that applies to all reflecting surfaces, one can approximate in a weak sense every random reflector by a sequence of specularly (deterministically) reflecting surfaces.

A special case of this theorem is that one can approximate every *deterministic* reflector that preserves the appropriate measure, including, for example, the reflector that reverses the direction of each incoming ray. To see how counterintuitive this is, imagine a higher dimensional analog of this reflector: a mirror that appears entirely the color of the observer’s eyeball, because the only light rays traveling from the mirror to the eyeball are those that bounced off of the eyeball before reaching the mirror. This striking effect is implemented in practice with limited accuracy (for theoretical and practical reasons) in “retroreflectors” (see [18]) and reflective paint (see [19]).

There are several sources of inspiration for our project. The article [1] studies reflected Brownian motion with inert drift. In a follow up project, Z.-Q. Chen and the second author plan to study the limiting situation when the diffusion coefficient

Received by the editors March 4, 2012 and, in revised form, April 7, 2012 and April 9, 2012.

2010 *Mathematics Subject Classification*. Primary 37D50.

The first author’s research was supported in part by NSERC and by the Sloan Foundation. The second author’s research was supported in part by NSF Grant DMS-0906743 and by grant N N201 397137, MNiSW, Poland. The third author’s research was supported in part by NSF Grant DMS 0645585.

of reflected Brownian motion goes to zero. It appears that in the limit, the particle will move along straight lines with a random angle of reflection.

The physics literature on random reflections is quite rich. We will not review it here; an excellent review can be found, for example, in [5].

Some early mathematical articles that considered random reflections were [14, 9], and more recent ones include [4, 5, 6, 7, 8, 11, 12, 15, 16]. Many of these articles are concerned with the so-called Knudsen reflection law. Knudsen proposed the cosine reflection law, in which the direction of the reflected molecule has a rotation-invariant distribution around the surface normal, with density proportional to the cosine of the angle with the normal, independent of the incidence direction. Knudsen's law is a model for gas dynamics. The same law is known as Lambert's cosine law in optics (see [2], pp. 147–148 or [13], Chap. 6).

On the technical side, our results seem to be related, at least at the intuitive level, to the “digital sundial” theorem proved by Falconer (see [10], Thm. 6.9). Roughly speaking, Falconer's theorem says that there exists a set with prescribed projections in almost all directions.

Although our article is close to the literature on billiards at the technical level, we will use the language of optics because our model is much closer to this circle of ideas at the intuitive level.

We will state our results in a rigorous way in Section 2. Section 3 contains short proofs of the most elementary results. The proof of our main result, Theorem 2.3, is a multistage construction presented in Sections 4–5.

2. THE MAIN RESULT, CONJECTURES AND OPEN PROBLEMS

For a very detailed and careful presentation of the billiards model in the plane, see Chap. 2 of [3]. We will be concerned with mirrors (walls of billiard tables) of very special shape. They are supposed to model macroscopically flat but rough reflecting surfaces. The paper of Feres [11] contains a rigorous mathematical presentation of this physical phenomenon and detailed analysis of its fundamental properties. Our setup is slightly different from that in [11].

Consider the following assumptions about a planar set M . These conditions contain, among other things, Assumptions A1–A4 from [3, Sections 2.1 and 2.4].

(M1) $M \subset \{(x_1, x_2) \in \mathbb{R}^2 : x_2 < 0\}$.

(M2) For every $k < \infty$, the set $\{(x_1, x_2) \in M : -k \leq x_1 \leq k\}$ is the union of a finite number of compact C^3 curves Γ_j .

(M3) The curves Γ_j intersect only at their endpoints. Each curve either is a line segment or has non-vanishing curvature of one sign (it has no inflection points). The curves do not form cusps at the intersection points; that is, the curves form an angle different from 0 at the point where they meet (if there is such a point).

We will say that $M \in \mathcal{M}_1$ if M satisfies (M1)–(M3).

Let $D = \mathbb{R} \times (-\pi, 0)$ and define a σ -finite measure Λ on D by $\Lambda(dx, d\alpha) = -dx \sin \alpha d\alpha$.

A light ray can be represented as $(x, \alpha) = (x(t), \alpha(t))$, where $x(t)$ is the location of the light ray at time t and $\alpha(t) \in (-\pi, \pi]$ is the angle between the direction of the light ray and the positive horizontal half-axis, measured in the counterclockwise direction from the half-axis. Time will play no role in our arguments, so it will be suppressed in the notation most of the time. We will always assume that light rays reflect from surfaces comprising $M \in \mathcal{M}_1$ according to the rule of specular

reflection; that is, the angle of incidence is equal to the angle of reflection, for every reflection.

Let $L_* := \{(x_1, x_2) \in \mathbb{R}^2 : x_2 = 0\}$. It will be convenient to identify L_* with \mathbb{R} ; for example, we will consider Λ to be a measure on $L_* \times (-\pi, 0)$. Consider the following natural condition.

(M4) Suppose that $M \in \mathcal{M}_1$ and for Λ -almost all $(x, \alpha) \in D$, a light ray starting from (x, α) and reflecting from surfaces comprising M will return to L_* after a finite number of reflections.

Condition (M4) is far from trivial; see, for example, Sec. 2.4 in [3] on accumulations of collisions. Moreover, some light rays reflecting from some mirror sets $M \in \mathcal{M}_1$ will not return to L_* . We will show that (M4) holds for a large class of sets M .

(M5) Suppose that $M \in \mathcal{M}_1$. Let $\{A'_k\}_{k \geq 1}$ be the family of all connected components of the open set $\{(x_1, x_2) \in \mathbb{R}^2 \setminus M : x_2 < 0\}$. There exists a subfamily $\{A_k\}_{k \geq 1}$ of $\{A'_k\}_{k \geq 1}$ (that is, every set A_k is equal to some set A'_j), such that every set A_k is bounded, $L_* \subset \bigcup_{k \geq 1} \partial A_k$, and the set $\{\partial A_j \cap \partial A_k \cap L_*, j, k \geq 1, j \neq k\}$ has no accumulation points in L_* .

We will say that $M \in \mathcal{M}_2$ if $M \in \mathcal{M}_1$ and it satisfies (M4).

Proposition 2.1. *If $M \in \mathcal{M}_1$ satisfies (M5), then it satisfies (M4) and, therefore, $M \in \mathcal{M}_2$.*

Consider some $M \in \mathcal{M}_2$. Suppose that a light ray starts from (x_0, α_0) with $x_0 \in L_*$ and $\alpha_0 \in (-\pi, 0)$ at time 0, reflects from surfaces of M and returns to L_* at a time t , i.e., $(x(t-), \alpha(t-)) = (x_1, \beta')$, $x_1 \in L_*$, and $t > 0$ is the smallest time with this property. Let $\beta = \beta' - \pi$. This defines a mapping $K : D \rightarrow D$, given by $K(x, \alpha) = (y, \beta)$. Clearly, K depends on M .

We will write $\mathbb{P}(x, \alpha; dy, d\beta)$ to denote a Markov transition kernel on D ; that is, for fixed $(x, \alpha) \in D$, $\mathbb{P}(x, \alpha; dy, d\beta)$ is a probability measure on D . We assume that \mathbb{P} satisfies the usual measurability conditions in all variables.

We will use $\delta_x(y)$ to denote Dirac's delta function. Recall the transformation K and let \mathbb{P}_K be defined by $\mathbb{P}_K(x, \alpha; dy, d\beta) = \delta_{K(x, \alpha)}(y, \beta) dy d\beta$. In other words, \mathbb{P}_K represents a deterministic Markov kernel, with the atom at $K(x, \alpha)$.

If $\mu_n, n \geq 1$, and μ_∞ are non-negative σ -finite measures on some measurable space Γ , then we will say that μ_n converge weakly to μ_∞ if there exists a sequence of sets $\Gamma_j, j \geq 1$, such that $\bigcup_{j \geq 1} \Gamma_j = \Gamma, \mu_n(\Gamma_j) < \infty, \mu_\infty(\Gamma_j) < \infty$ for all n and j , and for every fixed j , the sequence $\mu_n(\Gamma_j)$ converges weakly to $\mu_\infty(\Gamma_j)$.

Theorem 2.2. (i) *Consider the transformation $K : D \rightarrow D$ corresponding to any $M \in \mathcal{M}_2$. The transformation K preserves measure Λ ; that is, for any $A \subset D$ with $\Lambda(A) < \infty$, we have $\Lambda(K^{-1}(A)) = \Lambda(A)$. Moreover, K is "time reversible" in the sense that if $K(A_1) = A_2$, then $K(A_2) = A_1$.*

(ii) *Suppose that for some sequence of sets $M_n \in \mathcal{M}_2$, corresponding transformations K_n , and some Markov transition kernels $\mathbb{P}(x, \alpha; dy, d\beta)$, we have*

$$(2.1) \quad \Lambda(dx, d\alpha) \mathbb{P}_{K_n}(x, \alpha; dy, d\beta) \rightarrow \Lambda(dx, d\alpha) \mathbb{P}(x, \alpha; dy, d\beta)$$

in the sense of weak convergence on D^2 as $n \rightarrow \infty$. Then \mathbb{P} is symmetric with respect to Λ in the sense that for any smooth functions f and g on D with compact

support, we have

$$(2.2) \quad \begin{aligned} & \int_{D^2} f(y, \beta) \mathbb{P}(x, \alpha; dy, d\beta) g(x, \alpha) \Lambda(dx, d\alpha) \\ &= \int_{D^2} g(y, \beta) \mathbb{P}(x, \alpha; dy, d\beta) f(x, \alpha) \Lambda(dx, d\alpha). \end{aligned}$$

In particular, Λ is invariant in the sense that

$$(2.3) \quad \int_{D^2} f(y, \beta) \mathbb{P}(x, \alpha; dy, d\beta) \Lambda(dx, d\alpha) = \int_D f(x, \alpha) \Lambda(dx, d\alpha).$$

See [11, Sect. 4] for a similar result stated in a slightly different setting. The first part of the theorem says that all specular reflections are time reversible and preserve a certain measure. For this reason, Λ is known as the invariant measure for the collision map in the theory of billiards (see [3], Sec. 2.12). This is related to Lambert’s cosine law in optics (see [2], pp. 147–148 or [13], Chap. 6), also known as Knudsen’s cosine reflection law in the context of gas dynamics (see [5]). The second part shows that this condition can be interpreted as symmetry for a Markov kernel (see (2.2)). This symmetry is preserved under weak limits of Markov kernels. The next theorem, which is our main result, says that the symmetry of the Markov kernel expressed in (2.2) is the only condition on a Markov kernel \mathbb{P} necessary for the existence of deterministic approximations of random reflections represented by \mathbb{P} .

Recall that $\delta_x(y)$ denotes Dirac’s delta function. Suppose that the probability kernel \mathbb{P} in Theorem 2.2(ii) satisfies $\mathbb{P}(x, \alpha; dy, d\beta) = \delta_x(y) dy \tilde{\mathbb{P}}(x, \alpha; d\beta)$ for some $\tilde{\mathbb{P}}$. Heuristically, this means that the light ray released at x is instantaneously reflected from a mirror located infinitesimally close to L_* . Then (2.2) and (2.3) imply that for all smooth bounded functions f and g on $(-\pi, 0)$, and almost all x ,

$$(2.4) \quad \int_{(-\pi, 0)^2} f(\beta) \tilde{\mathbb{P}}(x, \alpha; d\beta) g(\alpha) \sin \alpha \, d\alpha = \int_{(-\pi, 0)^2} g(\beta) \tilde{\mathbb{P}}(x, \alpha; d\beta) f(\alpha) \sin \alpha \, d\alpha$$

and

$$(2.5) \quad \int_{(-\pi, 0)^2} f(\beta) \tilde{\mathbb{P}}(x, \alpha; d\beta) \sin \alpha \, d\alpha = \int_{(-\pi, 0)} f(\alpha) \sin \alpha \, d\alpha.$$

Theorem 2.3. *Suppose that $\mathbb{P}(x, \alpha; dy, d\beta) = \delta_x(y) dy \tilde{\mathbb{P}}(x, \alpha; d\beta)$, where $\tilde{\mathbb{P}}$ satisfies (2.4). Then there exists a sequence of sets $M_n \in \mathcal{M}_2$ and corresponding transformations K_n such that*

$$\Lambda(dx, d\alpha) \mathbb{P}_{K_n}(x, \alpha; dy, d\beta) \rightarrow \Lambda(dx, d\alpha) \mathbb{P}(x, \alpha; dy, d\beta)$$

weakly on D^2 as $n \rightarrow \infty$. Moreover, M_n can be chosen in such a way that

- (a) $M_n \subset \{(x_1, x_2) : -1/n < x_2 < 0\}$ and
- (b) for every $\varepsilon > 0$ there exists $D_\varepsilon \subset D$ with $\Lambda(D \setminus D_\varepsilon) < \varepsilon$ such that all rays starting in D_ε reflect from M_n exactly twice before returning to L_* .

Remark 2.4. The sets M_n that we construct in the proof of Theorem 2.3 are not connected. We believe that the following conjecture can be proved using an iteration of the construction used in the proof of Theorem 2.3.

Conjecture 2.5. *One can construct sets M_n so that they satisfy Theorem 2.3 except for (b); they are connected and every light ray reflects from M_n four times, except for a set of rays of Λ measure $1/n$.*

The following problem is inspired by the “digital sundial” theorem of Falconer (see [10], Thm. 6.9).

Problem 2.6. Is it possible to construct sets M_n so that they satisfy Theorem 2.3 and every light ray reflects from M_n only once, except for a set of rays of Λ measure $1/n$?

We believe that an analogue of Theorem 2.3 holds in higher dimensions.

3. DETERMINISTIC REFLECTIONS

Proof of Proposition 2.1. Assumptions A1–A4 from [3, Sections 2.1 and 2.4] are built into (M1)–(M3). Hence, there are no accumulation points for reflections from M , by the results in [3, Sect. 2.4].

The remaining part of the proof is based on [17, Thm. 3.1 and Sect. 7.1]. We will only outline the main steps. Fix some k and consider A_k to be a bounded billiard table. We define the “billiard ball map” T as follows. Let $v(x)$ be a continuous unit tangent vector field on ∂A_k . For a light ray starting from a point $x \in \partial A_k$, let α be the angle between its direction and $v(x)$. The light ray will hit ∂A_k at a point $y \in \partial A_k$. Let β be the angle formed by the direction of the light ray just after reflection at y and $v(y)$. Then we let $T(x, \alpha) = (y, \beta)$. Let dx represent the arc length measure on ∂A_k . By [17, Thm. 3.1], T preserves the measure $dx \sin \alpha d\alpha$ on $\partial A_k \times (0, \pi)$. Poincaré’s Recurrence Theorem (see [17, Thm. 7.4]) shows that $(dx d\alpha)$ -almost all rays starting at $(\partial A_k \cap L_*) \times (0, \pi)$ return to $(\partial A_k \cap L_*) \times (0, \pi)$ after a finite number of reflections; see [17, page 116] for more details. \square

Proof of Theorem 2.2. (i) The claim that $\Lambda(K^{-1}(A)) = \Lambda(A)$ is a special case of a well-known theorem; see [17, Thm. 3.1] or [3, Lemma 2.35]. The fact that the light reflection process is time reversible implies that if $K(A_1) = A_2$, then $K(A_2) = A_1$.

(ii) Suppose that $M_n \in \mathcal{M}_2$ and K_n is the corresponding transformation. Consider some sets $A_1, A_2 \in D$ and let $f(x, \alpha) = \mathbf{1}_{A_1}(x, \alpha)$ and $g(x, \alpha) = \mathbf{1}_{A_2}(x, \alpha)$. Then, using part (i),

$$\begin{aligned} & \int_{D^2} f(y, \beta) \mathbb{P}_{K_n}(x, \alpha; dy, d\beta) g(x, \alpha) \Lambda(dx, d\alpha) \\ &= \int_{D^2} \mathbf{1}_{A_1}(y, \beta) \delta_{K_n(x, \alpha)}(dy, d\beta) \mathbf{1}_{A_2}(x, \alpha) \Lambda(dx, d\alpha) \\ &= \int_{D^2} \mathbf{1}_{K_n^{-1}(A_1)}(x, \alpha) \mathbf{1}_{A_2}(x, \alpha) \Lambda(dx, d\alpha) \\ &= \Lambda(K_n^{-1}(A_1) \cap A_2) = \Lambda(K_n^{-1}(A_2) \cap A_1). \end{aligned}$$

For the same reason,

$$\int_{D^2} g(y, \beta) \mathbb{P}_{K_n}(x, \alpha; dy, d\beta) f(x, \alpha) \Lambda(dx, d\alpha) = \Lambda(K_n^{-1}(A_2) \cap A_1),$$

so

$$\begin{aligned} & \int_{D^2} f(y, \beta) \mathbb{P}_{K_n}(x, \alpha; dy, d\beta) g(x, \alpha) \Lambda(dx, d\alpha) \\ &= \int_{D^2} g(y, \beta) \mathbb{P}_{K_n}(x, \alpha; dy, d\beta) f(x, \alpha) \Lambda(dx, d\alpha). \end{aligned}$$

A standard argument based on finite linear combinations of step functions and bounded convergence shows that (2.2) holds for smooth f and g with compact

support and \mathbb{P} of the form \mathbb{P}_{K_n} . This and weak convergence imply that (2.2) also holds for Markov kernels \mathbb{P} that satisfy (2.1). We obtain (2.3) from (2.2) by the monotone convergence theorem applied to a sequence $g_n \uparrow 1$ as $n \rightarrow \infty$. \square

4. TRANSPOSITION REFLECTOR

This section is devoted to the construction of a set of mirrors that transpose thin bundles of light of appropriate angular width. The first challenge is to guide bundles of light rays so that only an arbitrarily small amount of light is scattered in an unaccounted for way. The second challenge (easier to present in a rigorous way than the first one) is to place various sets of mirrors so that they do not interfere with each other.

Let \bar{A} denote the closure of a set A .

Definition 4.1. We will call $K : D \rightarrow D$ a *simple symmetric function* if there exists a countable family of rectangles $Q_k = (x_1^k, x_2^k) \times (\alpha_1^k, \alpha_2^k) \subset D$, $k \geq 1$, such that

- (i) $Q_k \cap Q_j = \emptyset$ for $k \neq j$,
- (ii) $\bigcup_{k \geq 1} \bar{Q}_k = D$,
- (iii) for any j, k , either $(x_1^k, x_2^k) = (x_1^j, x_2^j)$ or $(x_1^k, x_2^k) \cap (x_1^j, x_2^j) = \emptyset$,
- (iv) for every $a < \infty$, there is only a finite number of k such that $(x_1^k, x_2^k) \cap (-a, a) \neq \emptyset$,
- (v) for every k there exists α such that $K(Q_k) = (x_1^k, x_2^k) \times \{\alpha\}$,
- (vi) for every k there exists j such that

$$\int_{(\alpha_1^k, \alpha_2^k)} \sin \alpha d\alpha = \int_{(\alpha_1^j, \alpha_2^j)} \sin \alpha d\alpha,$$

$$(x_1^k, x_2^k) = (x_1^j, x_2^j), K(Q_k) \subset Q_j \text{ and } K(Q_j) \subset Q_k.$$

We will call $\sup_k(x_2^k - x_1^k) \vee \sup_k(\alpha_2^k - \alpha_1^k)$ the *mesh* of K .

Recall from Section 2 that for $K : D \rightarrow D$, the kernel \mathbb{P}_K is defined by $\mathbb{P}_K(x, \alpha; dy, d\beta) = \delta_{K(x, \alpha)}(y, \beta) dy d\beta$. Suppose that $\tilde{\mathbb{P}}$ satisfies (2.4). Then standard arguments (see the proof of Theorem 2.3 in Section 5) show that there exists a sequence $\{\hat{K}_n\}_{n \geq 1}$ of simple symmetric functions such that $\mathbb{P}_{\hat{K}_n}(x, \alpha; dy, d\beta) \rightarrow \delta_x(y) dy \tilde{\mathbb{P}}(x, \alpha; d\beta)$ weakly as $n \rightarrow \infty$. Note that we do not claim that \hat{K}_n arise as functions associated to reflecting sets in \mathcal{M}_2 . To prove Theorem 2.3, it will suffice to show that for any simple symmetric function $\hat{K}_n : D \rightarrow D$, there exists a sequence of sets $M_n \in \mathcal{M}_2$, $n \geq 1$, satisfying conditions (a) and (b) of Theorem 2.3 and such that if the functions $K_n : D \rightarrow D$ correspond to M_n 's, then $K_n - \hat{K}_n \rightarrow 0$ pointwise. The rest of the paper is devoted to the construction of sets M_n .

For $\rho > 0$, let Λ_ρ be the measure Λ restricted to $D_\rho := (-\rho, \rho) \times (-\pi, 0)$ and note that the total mass of Λ_ρ is 4ρ .

Lemma 4.2. *For any $\varepsilon_0 > 0$ there exist $n_0 < \infty$ and $\rho_0 > 0$ such that for any $n > n_0$, $\rho \in (0, \rho_0)$ and any simple symmetric function K with mesh less than or equal to $1/n$ such that $(-\rho, \rho)$ is one of the intervals (x_1^k, x_2^k) , the following holds:*

- (i) *There exist a bounded set $N \in \mathcal{M}_1$ and a set $D_\rho^* \subset D_\rho$ such that $\Lambda_\rho(D_\rho^*) < 4\rho\varepsilon_0$.*
- (ii) *The function $K_N(x, \alpha)$ corresponding to N is defined on $D_\rho \setminus D_\rho^*$.*
- (iii) *For all $(x, \alpha) \in D_\rho \setminus D_\rho^*$, we have $|K_N(x, \alpha) - K(x, \alpha)| \leq \varepsilon_0$.*

The function $K_N(x, \alpha)$ is not necessarily defined for all $(x, \alpha) \in D_\rho$ because some light rays starting from L_* and reflected in N may never come back to L_* .

Proof of Lemma 4.2. We will omit some details of our arguments and estimates that are totally elementary but are tedious to write down. However, we will now provide a solid justification for the approximate formulas that will form the core of the proof. Suppose that two (parts of) ellipses are fixed and serve as mirrors for light rays. Suppose that a light ray starts from $x \in L_*$ at an angle $\alpha \in (-\pi, 0)$, then makes two reflections from the elliptic mirrors, and then returns to L_* at $y \in L_*$ and angle β . An important (but elementary) observation is that the function $(x, \alpha) \rightarrow (y, \beta)$ is analytic. The reason is that the equations that determine the points and angles of reflection of the light ray are quadratic with “parameters” that are analytic functions of x and α . A similar remark applies to other quantities that are functions of x and α , such as the location and angle of incidence when the light ray hits one of the ellipses, or the distance from the light ray to a fixed point.

Step 1. In this step, we will construct a pair of elliptic mirrors that interchange two very thin bundles of light rays. This step is devoted only to the construction (or, in other words, definition) of the two mirrors. The next step will contain the proof that the mirrors reflect the light rays in the desired directions.

We will use letters A, B and C (with subscripts and superscripts) to denote points in the plane. The notation \overline{AB} will refer to a line segment with given endpoints; \overline{ABC} will denote an arc of an ellipse passing through the three points.

Consider some (small) $c_1 > 0, \alpha, \beta \in (-\pi + c_1, -c_1)$ and (small) $\Delta\alpha, \Delta\beta \in (0, 1)$, such that $\Delta\alpha \sin \alpha = \Delta\beta \sin \beta$. Suppose that $\rho > 0$ and (see Figure 1),

$$A_1 = (-\rho, 0), A_2 = (0, 0), A_3 = (\rho, 0).$$

First assume that $\alpha = \beta$. In this case, let $\mathcal{E}_\alpha = \mathcal{E}_\beta$ be the arc of a circle represented in complex notation as $\mathcal{E}_\alpha = \{re^{i\gamma} : \alpha - \Delta\alpha \leq \gamma \leq \alpha + \Delta\alpha\}$ for some $r > 0$. Note that the circle is centered at A_2 . We will consider \mathcal{E}_α to be a mirror. Light rays starting from (x, γ) in $D_\alpha^1 := \overline{A_1A_3} \times (\alpha - \Delta\alpha, \alpha + \Delta\alpha)$ will mostly return to the same set if ρ is small. More precisely,

$$\forall c_1 > 0 \forall \eta_1 > 0 \exists \rho_0 < \infty \forall \rho \in (0, \rho_0) \forall \alpha \in (-\pi + c_1, -c_1) : \\ \frac{\Lambda_\rho(\{(x, \gamma) \in D_\alpha^1 : K_{\mathcal{E}_\alpha}(x, \gamma) \in D_\alpha^1\})}{\Lambda_\rho(D_\alpha^1)} > 1 - \eta_1.$$

The case when $\alpha = \beta$ is rather easy, so we will focus on the case $\alpha \neq \beta$ in the rest of the proof.

Recall notation from Definition 4.1 and let $\hat{\alpha}^k = (\alpha_1^k + \alpha_2^k)/2$. Consider j and k such that $K(Q_k) \subset Q_j$ and $K(Q_j) \subset Q_k$. If $(\alpha_2^k - \alpha_1^j) \vee (\alpha_2^j - \alpha_1^k) \leq \varepsilon_0/2$, then let $K'(x, \alpha) = (x, \hat{\alpha}^k)$ for $(x, \alpha) \in Q_k$, and $K'(x, \alpha) = (x, \hat{\alpha}^j)$ for $(x, \alpha) \in Q_j$. Otherwise, we let $K'(x, \alpha) = K(x, \alpha)$. Note that K' is a simple symmetric function and $|K'(x, \alpha) - K(x, \alpha)| \leq \varepsilon_0/2$. Hence, it will suffice to find a set N and the corresponding function K_N such that $|K_N(x, \alpha) - K'(x, \alpha)| \leq \varepsilon_0/2$.

Consider the case when $|\alpha - \beta| \geq \varepsilon_0/2$ and $\alpha - \Delta\alpha, \alpha + \Delta\alpha, \beta - \Delta\beta, \beta + \Delta\beta \in (-\pi, 0)$. For any angle $\gamma \in (-\pi, 0)$, let L_γ be the line passing through A_2 , with the slope $\tan(-\gamma)$. Unless stated otherwise, we will consider only points and sets below the line L_* .

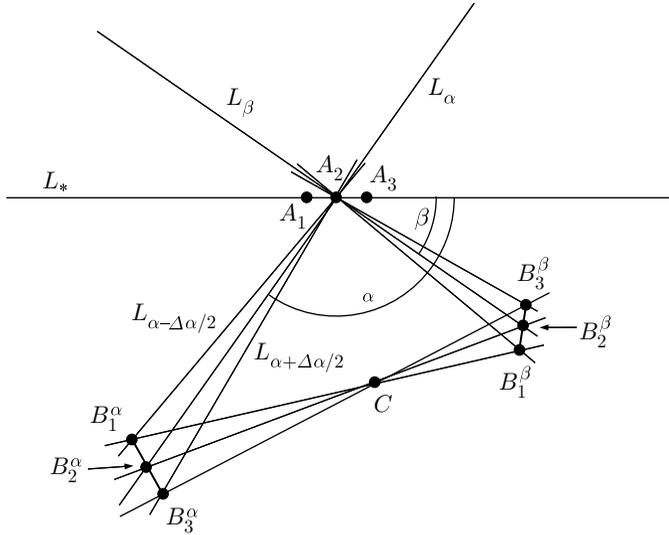


FIGURE 1. A single light ray interchange mirror system.

Let B_2^α be a point on L_α , and $r_\alpha := |A_2 - B_2^\alpha|$. Similarly, let B_2^β be a point on L_β , and $r_\beta := |A_2 - B_2^\beta|$ (values of r_α and r_β will be specified later). Let L' be the line passing through B_2^α and B_2^β , and let C be the point on L' such that $s_1 := |C - B_2^\alpha| = |C - B_2^\beta|$.

Recall that for any ellipse, if a light ray leaves one of the foci and meets a point on that ellipse, it will reflect off the ellipse and pass through the other focus.

Let $B_1^\alpha \in L_{\alpha-\Delta\alpha/2}$ and $B_3^\alpha \in L_{\alpha+\Delta\alpha/2}$ be points such that $\mathcal{E}_\alpha := \overline{B_1^\alpha B_2^\alpha B_3^\alpha}$ is an arc of an ellipse with the foci A_2 and C (see Figure 1). Similarly, let $B_1^\beta \in L_{\beta-\Delta\beta/2}$ and $B_3^\beta \in L_{\beta+\Delta\beta/2}$ be points such that $\mathcal{E}_\beta := \overline{B_1^\beta B_2^\beta B_3^\beta}$ is an arc of an ellipse with the foci A_2 and C .

The rest of this step is devoted to informal estimates aimed at finding values of r_α and r_β that will define mirrors with desirable properties. Consider the bundle \mathcal{B} of light rays starting at A_2 along lines between $L_{\alpha-\Delta\alpha/2}$ and $L_{\alpha+\Delta\alpha/2}$. We will find parameters of our construction such that for the resulting mirrors \mathcal{E}_α and \mathcal{E}_β , “most” light rays in \mathcal{B} reflect from \mathcal{E}_α , then reflect from \mathcal{E}_β and then hit A_2 . Let ρ_1 be the width of \mathcal{B} close to \mathcal{E}_α , just before hitting \mathcal{E}_α . Since $\Delta\alpha$ is assumed to be small, the rays in \mathcal{B} are almost parallel and \mathcal{E}_α is almost flat. Hence, the width of \mathcal{B} close to \mathcal{E}_α , just after hitting \mathcal{E}_α is $\rho_1(1 + o(\Delta\alpha))$. We have $\rho_1 = r_\alpha(\Delta\alpha + o(\Delta\alpha))$ for small $\Delta\alpha$.

Since \mathcal{E}_α is an arc of an ellipse with foci A_2 and C , all light rays in \mathcal{B} will pass through C and then they will hit \mathcal{E}_β , assuming that \mathcal{E}_β is large enough. Since \mathcal{E}_β is an arc of an ellipse with foci A_2 and C , the light rays in \mathcal{B} reflected from \mathcal{E}_β will hit L_* at A_2 .

Let ρ_2 be the width of \mathcal{B} close to \mathcal{E}_β , just before hitting \mathcal{E}_β . The width of \mathcal{B} close to \mathcal{E}_β , just after hitting \mathcal{E}_β , is $\rho_2(1 + o(\Delta\alpha))$. We want the bundle of light rays \mathcal{B} to form a cone with angle $\Delta\beta$ at the point A_2 on the way out, so we would like to have $\rho_2 = r_\beta(\Delta\beta + o(\Delta\beta))$ for small $\Delta\beta$. The point C is half way between B_2^α and

B_2^β , so we would like to have $\rho_1 = \rho_2$. Therefore, we choose r_α and r_β so that they satisfy

$$r_\alpha(\Delta\alpha + o(\Delta\alpha)) = \rho_1 = \rho_2 = r_\beta(\Delta\beta + o(\Delta\beta)).$$

More precisely, we choose r_α and r_β so that

$$(4.1) \quad \frac{r_\alpha}{r_\beta} = \frac{\Delta\beta}{\Delta\alpha} = \frac{\sin \alpha}{\sin \beta}.$$

The above formula incorporates our previously made assumption that $\Delta\alpha \sin \alpha = \Delta\beta \sin \beta$.

Step 2. We will now argue that a pair of mirrors defined in the previous step guides a thin bundle of light rays to the appropriate exit location and exit angle, with arbitrarily little “waste”, for appropriate values of parameters of the construction. Our argument is based on detailed analysis of small changes in the initial conditions (the starting location and angle of the light ray) on the “output”, that is, the location and angle for the light ray exiting the lower half-plane.

Let \mathbf{E}_α be the ellipse of which \mathcal{E}_α is a part and let \mathbf{E}_β have the analogous meaning. The informal estimates of Step 1 have the following rigorous version. For any $c_1, \varepsilon > 0$, there exists $\varepsilon_1 \in (0, \varepsilon)$ such that for all $\alpha, \beta \in (-\pi + c_1, -c_1)$ with $|\alpha - \beta| \geq \varepsilon_0/2$, $\Delta\alpha, \Delta\beta \in (0, \varepsilon_1)$, and r_α and r_β satisfying (4.1), we have the following. If a light ray starts from A_2 , follows $L_{\alpha+\alpha_1}$ with $\alpha_1 \in [-(1 - \varepsilon_1)\Delta\alpha/2, (1 - \varepsilon_1)\Delta\alpha/2]$ and reflects from \mathcal{E}_α , then it will intersect \mathbf{E}_β at a point x such that

$$(4.2) \quad |x - B_2^\beta| < (1 - \varepsilon)(|B_1^\beta - B_2^\beta| \vee |B_3^\beta - B_2^\beta|).$$

Similarly, if a light ray starts from A_2 , follows $L_{\beta+\beta_1}$ with $\beta_1 \in [-(1 - \varepsilon_1)\Delta\beta/2, (1 - \varepsilon_1)\Delta\beta/2]$ and reflects from \mathcal{E}_β , then it will intersect \mathbf{E}_α at a point y such that

$$(4.3) \quad |y - B_2^\alpha| < (1 - \varepsilon)(|B_1^\alpha - B_2^\alpha| \vee |B_3^\alpha - B_2^\alpha|).$$

It follows from (4.2) and (4.3) that all light rays starting from A_2 at angles in $[-(1 - \varepsilon_1)\Delta\alpha/2, (1 - \varepsilon_1)\Delta\alpha/2]$ or $[-(1 - \varepsilon_1)\Delta\beta/2, (1 - \varepsilon_1)\Delta\beta/2]$ reflect from both mirrors \mathcal{E}_α and \mathcal{E}_β and then reach L_* again at A_2 .

In the following part of the argument, we will consider α, β, r_α and r_β to be “fixed” parameters satisfying (4.1) and we will choose $\Delta\alpha, \Delta\beta$ and ρ sufficiently small, so that certain conditions are satisfied.

We will define a number of arcs, lines and points which will be used in the next part of the proof; see Figure 2. Consider any point $A_4 \in \overline{A_1A_3}$ and light ray R_1 emanating from A_4 at an angle η . Let $B_4 \in \mathbf{E}_\alpha$ and $B_6 \in \mathbf{E}_\beta$ denote the points where this light ray reflects in \mathbf{E}_α and \mathbf{E}_β . Let R_4 be the line between B_4 and B_6 . Let R_6 be the line that contains this light ray after reflection in \mathbf{E}_β and let B_8 be the point where R_6 intersects L_* . Let R_2 be the line passing through A_2 and parallel to R_1 and let B_5 be the intersection point of R_2 with \mathbf{E}_α . Let R_3 be the line passing through A_2 and B_4 . Let R_5 be the line that represents light ray R_3 after reflecting from \mathbf{E}_α . Let B_7 be the point where R_5 intersects \mathbf{E}_β . Let R_7 be the line passing through A_2 and B_6 .

Let α_1 be the angle between $\overline{A_2B_2^\alpha}$ and the tangent to \mathcal{E}_α at B_2^α . The angle α_1 is a function of α, β, r_α and r_β only. Assume that $\eta \in [\alpha - (1 - \varepsilon_1)\Delta\alpha/2,$

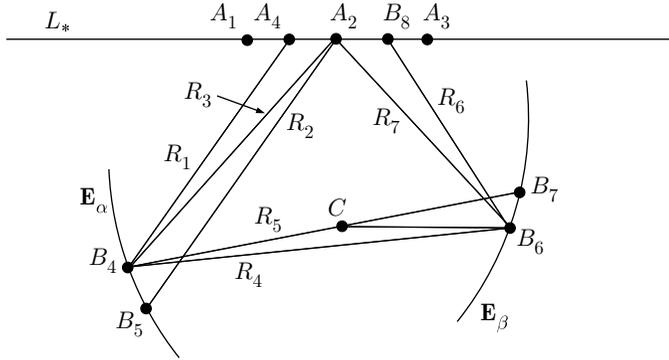


FIGURE 2. Trajectories of various light rays.

$\alpha + (1 - \varepsilon_1)\Delta\alpha/2]$. Then

$$\begin{aligned} |B_5 - B_2^\alpha| &\leq (1 - \varepsilon_1 + o(\varepsilon_1))(|B_1^\alpha - B_2^\alpha| \vee |B_3^\alpha - B_2^\alpha|) \\ &\leq (1 - \varepsilon_1 + o(\varepsilon_1)) \frac{r_\alpha(\Delta\alpha/2)(1 + o(\Delta\alpha))}{\sin \alpha_1}. \end{aligned}$$

We also have

$$|B_4 - B_5| \leq (1 + o(\rho)) \frac{\rho \sin \alpha}{\sin \alpha_1}$$

and, therefore,

$$|B_4 - B_2^\alpha| \leq (1 - \varepsilon_1 + o(\varepsilon_1)) \frac{r_\alpha(\Delta\alpha/2)(1 + o(\Delta\alpha))}{\sin \alpha_1} + (1 + o(\rho)) \frac{\rho \sin \alpha}{\sin \alpha_1}.$$

It follows that for any $c_1, \varepsilon > 0$, there exists $\varepsilon_1 \in (0, \varepsilon)$, $\varepsilon_2 > 0$, such that for all $\alpha, \beta \in (-\pi + c_1, -c_1)$ with $|\alpha - \beta| \geq \varepsilon_0/2$ and $\Delta\alpha, \Delta\beta \in (0, \varepsilon_1)$, r_α and r_β satisfying (4.1), all

$$\rho \in (0, \varepsilon_2(\Delta\alpha \wedge \Delta\beta)) \text{ and } \eta \in [\alpha - (1 - \varepsilon_1)\Delta\alpha/2, \alpha + (1 - \varepsilon_1)\Delta\alpha/2],$$

we have $B_4 \in \mathcal{E}_\alpha$. By symmetry, if $\eta \in [\beta - (1 - \varepsilon_1)\Delta\beta/2, \beta + (1 - \varepsilon_1)\Delta\beta/2]$, then R_1 intersects \mathbf{E}_β at a point in \mathcal{E}_β .

Suppose that $\eta \in [\alpha - (1 - \varepsilon_1)\Delta\alpha/2, \alpha + (1 - \varepsilon_1)\Delta\alpha/2]$. Let β_1 be the angle between $\overline{A_2B_2^\beta}$ and the tangent to \mathcal{E}_β at B_2^β . Note that β_1 is also the angle between $\overline{CB_2^\beta}$ and the tangent to \mathcal{E}_β at B_2^β . Let γ_1 be the angle between R_1 and R_3 . The angle between R_4 and R_5 is also γ_1 . We have

$$\gamma_1 = \sin \alpha |A_4 - A_2|(1 + o(\rho) + o(\Delta\alpha))/r_\alpha \leq \rho \sin \alpha(1 + o(\rho) + o(\Delta\alpha))/r_\alpha.$$

Thus

$$\begin{aligned} (4.4) \quad |B_6 - B_7| &= \frac{2s_1 \sin \gamma_1}{\sin \beta_1} (1 + o(\rho) + o(\Delta\beta)) \\ &\leq \frac{2s_1 \rho \sin \alpha(1 + o(\rho) + o(\Delta\alpha))}{r_\alpha \sin \beta_1} (1 + o(\rho) + o(\Delta\beta)) \\ &= \frac{2s_1 \rho \sin \alpha}{r_\alpha \sin \beta_1} (1 + o(\rho) + o(\Delta\beta)). \end{aligned}$$

It follows from (4.2), (4.3) and (4.4) that for any $c_1, \varepsilon > 0$ there exists $\varepsilon_1 \in (0, \varepsilon)$, $\varepsilon_2 > 0$ such that for all $\alpha, \beta \in (-\pi + c_1, -c_1)$ with $|\alpha - \beta| \geq \varepsilon_0/2$ and $\Delta\alpha, \Delta\beta \in (0, \varepsilon_1)$, r_α and r_β satisfying (4.1), and all $\rho \in (0, \varepsilon_2(\Delta\alpha \wedge \Delta\beta))$ we have the following. If a light ray starts from a point in $\overline{A_1A_3}$ at an angle $\eta \in [\alpha - (1 - \varepsilon_1)\Delta\alpha/2, \alpha + (1 - \varepsilon_1)\Delta\alpha/2]$, then it reflects from \mathcal{E}_α and then it intersects \mathcal{E}_β . Similarly, if a light ray starts from a point in $\overline{A_1A_3}$ at an angle $\eta \in [\beta - (1 - \varepsilon_1)\Delta\beta/2, \beta + (1 - \varepsilon_1)\Delta\beta/2]$, then it reflects from \mathcal{E}_β and then it intersects \mathcal{E}_α .

Let a denote the distance between C and R_4 . We have

$$a = s_1 \sin \gamma_1(1 + o(\rho)) = s_1 \sin \alpha |A_4 - A_2|(1 + o(\rho) + o(\Delta\alpha))/r_\alpha.$$

Note that if a light ray moved along $\overline{CB_6}$ and reflected from \mathbf{E}_β , then it stays on R_7 . The angle between R_6 and R_7 is equal to the angle between $\overline{CB_6}$ and R_4 and, therefore, it is equal to $\gamma_2 := (a/s_1)(1 + o(\rho) + o(\Delta\alpha))$. The following estimate for the distance between A_2 and B_8 uses (4.1),

$$\begin{aligned} |B_8 - A_2| &= \frac{\gamma_2 r_\beta}{\sin \beta}(1 + o(\Delta\alpha)) = \frac{ar_\beta}{s_1 \sin \beta}(1 + o(\rho) + o(\Delta\alpha)) \\ &= \frac{s_1 \sin \alpha |A_4 - A_2| r_\beta}{r_\alpha s_1 \sin \beta}(1 + o(\rho) + o(\Delta\alpha)) \\ &= |A_4 - A_2|(1 + o(\rho) + o(\Delta\alpha)). \end{aligned}$$

We combine this estimate with the conclusions obtained so far to see that for any $c_1, \varepsilon > 0$ there exists $\varepsilon_1 \in (0, \varepsilon)$, $\varepsilon_2 > 0$ such that for all $\alpha, \beta \in (-\pi + c_1, -c_1)$ with $|\alpha - \beta| \geq \varepsilon_0/2$ and $\Delta\alpha, \Delta\beta \in (0, \varepsilon_1)$, r_α and r_β satisfying (4.1), all $\rho \in (0, \varepsilon_2(\Delta\alpha \wedge \Delta\beta))$, and all $A_4 \in \overline{A_1A_3}$ satisfying $|A_4 - A_2| \leq (1 - \varepsilon_2)\rho$ we have the following. If a light ray starts from A_4 at an angle $\eta \in [\alpha - (1 - \varepsilon_1)\Delta\alpha/2, \alpha + (1 - \varepsilon_1)\Delta\alpha/2]$, then it reflects from \mathcal{E}_α , next it reflects from \mathcal{E}_β and then it intersects L_* between A_1 and A_3 . Similarly, if a light ray starts from A_4 at an angle $\eta \in [\beta - (1 - \varepsilon_1)\Delta\beta/2, \beta + (1 - \varepsilon_1)\Delta\beta/2]$, then it reflects from \mathcal{E}_β , next it reflects from \mathcal{E}_α and then it intersects L_* between A_1 and A_3 .

Step 3. We will now assemble a finite family of pairs of mirrors so that they properly guide light rays entering the system at “most” angles and the mirrors do not interfere with one another. The lack of interference will be achieved by inductive scaling of the mirrors, that is, making them large, so that a light ray traveling between a pair of mirrors takes a path far beyond all the mirrors constructed earlier in the inductive procedure.

Recall the notation from the statement of the lemma. Let $c_1 > 0$ be so small that

$$\Lambda_\rho(D_\rho \setminus ((-\rho, \rho) \times (-\pi + c_1, -c_1))) < 4\rho\varepsilon_0/16.$$

Note that c_1 satisfying this condition can be chosen independently of ρ .

Recall the simple symmetric function K' defined in Step 1 and let $\{Q'_k\}_{k \geq 1} = \{(x_1^k, x_2^k) \times (\alpha_1^k, \alpha_2^k)\}_{k \geq 1}$ be the family of rectangles as in Definition 4.1. Let \mathbf{J} be the family of pairs (j, k) such that $j \leq k$,

$$K'((-\rho, \rho) \times (\alpha_1^j, \alpha_2^j)) \subset (-\rho, \rho) \times (\alpha_1^k, \alpha_2^k)$$

and, therefore,

$$K'((-\rho, \rho) \times (\alpha_1^k, \alpha_2^k)) \subset (-\rho, \rho) \times (\alpha_1^j, \alpha_2^j).$$

Note that according to Definition 4.1(iv), the set \mathbf{J} is finite. Let \mathcal{J} be the set of all j such that $[\alpha_1^j, \alpha_2^j] \cap (-\pi + c_1, -c_1) \neq \emptyset$. For $(j, k) \in \mathbf{J}$, let

$$\begin{aligned} \widehat{\alpha}^j &= (\alpha_1^j + \alpha_2^j)/2, \\ \Delta \widehat{\alpha}^j &= \alpha_2^j - \alpha_1^j, \\ \Delta \widehat{\alpha}^k &= \Delta \widehat{\alpha}^j \sin \widehat{\alpha}^j / \sin \widehat{\alpha}^k. \end{aligned}$$

Note that, typically, it is not true that $\Delta \widehat{\alpha}^k = \alpha_2^k - \alpha_1^k$ because if $(j, k) \in \mathbf{J}$ and $j \neq k$, then $(k, j) \notin \mathbf{J}$. The quantity $\Delta \widehat{\alpha}^k$ is well defined because if $(j, k) \in \mathbf{J}$ and $j \neq k$, then there is no i such that $(k, i) \in \mathbf{J}$. It is easy to see that for fixed $c_1, \varepsilon_1 > 0$, we can choose n_0 so large that for all $n \geq n_0$, if the mesh of K is smaller than $1/n$, then for all $j \in \mathcal{J}$ we have

$$[\widehat{\alpha}^j - (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2, \widehat{\alpha}^j + (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2] \subset (\alpha_1^j, \alpha_2^j).$$

Let

$$\Theta(\varepsilon_1) = (-\pi + c_1, -c_1) \cap \bigcup_{j \in \mathcal{J}} [\widehat{\alpha}^j - (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2, \widehat{\alpha}^j + (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2].$$

For a fixed $c_1 > 0$, we can make $\varepsilon_1 > 0$ smaller and n_0 larger, if necessary, so that

$$\Lambda_\rho((-\rho, \rho) \times \Theta(\varepsilon_1)) > 4\rho(1 - \varepsilon_0/8).$$

We make n_0 larger, if necessary, so that for $n \geq n_0$ and all $j \in \mathcal{J}$ we have $\Delta \widehat{\alpha}^j < \varepsilon_1$.

Choose $\varepsilon_3 > 0$ so small that

$$\Lambda_\rho((- (1 - \varepsilon_3)\rho, (1 - \varepsilon_3)\rho) \times \Theta(\varepsilon_1)) > 4\rho(1 - \varepsilon_0/4).$$

We will write $r(\alpha)$ instead of r_α and $\mathcal{E}(\alpha)$ instead of \mathcal{E}_α , for typographical reasons.

Let $\Delta^* \alpha = \min_{j \in \mathcal{J}} \Delta \widehat{\alpha}^j$. It follows from what we have shown in Step 2 that for $c_1, \varepsilon > 0$ there exists $n_0, \varepsilon_1 \in (0, \varepsilon), \varepsilon_2 > 0$ such that for all $\rho \in (0, \varepsilon_2 \Delta^* \alpha), n \geq n_0$ and all $j \in \mathcal{J}$ we have the following. If $(j, k) \in \mathbf{J}$ and a light ray starts from a point in $(-\rho, \rho) \times L_*$ at an angle

$$\eta \in [\widehat{\alpha}^j - (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2, \widehat{\alpha}^j + (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2],$$

then it reflects from $\mathcal{E}(\widehat{\alpha}^j)$, next it reflects from $\mathcal{E}(\widehat{\alpha}^k)$ and then it intersects L_* at a point in $(-\rho, \rho) \times L_*$. Moreover, if a light ray starts from a point in $(-\rho, \rho) \times L_*$ at an angle

$$\eta \in [\widehat{\alpha}^k - (1 - \varepsilon_1)\Delta \widehat{\alpha}^k/2, \widehat{\alpha}^k + (1 - \varepsilon_1)\Delta \widehat{\alpha}^k/2],$$

then it reflects from $\mathcal{E}(\widehat{\alpha}^k)$, next it reflects from $\mathcal{E}(\widehat{\alpha}^j)$ and then it intersects L_* at a point in $(-\rho, \rho) \times L_*$.

Let $(j_1, k_1), (j_2, k_2), \dots$ be an arbitrary ordering of pairs in \mathbf{J} . Choose $r(\widehat{\alpha}^{j_1})$ and $r(\widehat{\alpha}^{k_1})$ so that they satisfy (4.1). Suppose that $r(\widehat{\alpha}^{j_m})$ and $r(\widehat{\alpha}^{k_m})$ have been chosen for $m = 1, 2, \dots, i$. We choose $r(\widehat{\alpha}^{j_{i+1}})$ and $r(\widehat{\alpha}^{k_{i+1}})$ so that they satisfy (4.1) and they are so large that lines connecting any points in $\mathcal{E}(\widehat{\alpha}^{j_{i+1}})$ and $\mathcal{E}(\widehat{\alpha}^{k_{i+1}})$ do not intersect any $\mathcal{E}(\widehat{\alpha}^{j_m})$ and $\mathcal{E}(\widehat{\alpha}^{k_m})$ for $m = 1, 2, \dots, i$.

Recall that \mathbf{J} is finite and let n_1 be its cardinality. For $(j, k) \in \mathbf{J}$, let $\widehat{D}_{\rho, (j, k)}$ be the set of light rays in D_ρ such that the ray starts from a point in $(-\rho, \rho) \times L_*$ at an angle

$$\eta \in [\widehat{\alpha}^j - (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2, \widehat{\alpha}^j + (1 - \varepsilon_1)\Delta \widehat{\alpha}^j/2]$$

or at an angle

$$\eta \in [\widehat{\alpha}^k - (1 - \varepsilon_1)\Delta\widehat{\alpha}^k/2, \widehat{\alpha}^k + (1 - \varepsilon_1)\Delta\widehat{\alpha}^k/2].$$

Let $\widetilde{D}_{\rho,(j,k)}$ be the set of light rays in $\widehat{D}_{\rho,(j,k)}$ such that at some time the ray reflects from a set $\mathcal{E}(\widehat{\alpha}^m)$, for some $m \neq j, k$. If $\rho = 0$, that is, if $[-(1 - \varepsilon_3)\rho, (1 - \varepsilon_3)\rho]$ is the single point A_2 , then $\widetilde{D}_{\rho,(j,k)} = \emptyset$, by the claim made in the previous paragraph. The proportion (in terms of the measure Λ_ρ) of light rays in $\widehat{D}_{\rho,(j,k)}$ which belong to $\widetilde{D}_{\rho,(j,k)}$ is a continuous function of $\rho > 0$, with zero limit when $\rho \downarrow 0$; that is, $\lim_{\rho \downarrow 0} \Lambda_\rho(\widetilde{D}_{\rho,(j,k)})/\Lambda_\rho(\widehat{D}_{\rho,(j,k)}) = 0$. It follows that for sufficiently small $\rho > 0$, we have $\Lambda_\rho(\widetilde{D}_{\rho,(j,k)}) < 4\rho\varepsilon_0/(16n_1)$. Let $\widetilde{D}_\rho = \bigcup_{(j,k) \in \mathbf{J}} \widetilde{D}_{\rho,(j,k)}$. We see that for fixed $r(\widehat{\alpha}^j)$'s, we can make $\rho > 0$ so small that $\Lambda_\rho(\widetilde{D}_\rho) < 4\rho\varepsilon_0/16$.

Let $N = \bigcup_{j \in \mathcal{J}} \mathcal{E}(\widehat{\alpha}^j)$. Our arguments have shown that the function K_N satisfies the assertions listed in the lemma, with function K' in place of K , and $\varepsilon_0/2$ in place of ε_0 . We have observed in Step 1 that this is sufficient for the proof of the lemma. \square

5. LIGHT REFLECTOR CELLS IN CANTOR SET HOLES

The system of mirrors constructed in the previous section is very “inefficient” in that it has a large diameter compared to the length of the segment $(-\rho, \rho)$ of L_* , where the light enters and exits the system. In this section, we will build a “compact” version of the reflector using scaling and a Cantor-like construction. The following result is very similar to Lemma 4.2 except that we use smaller mirrors.

Lemma 5.1. *Fix any $\rho, \rho_1 > 0$, $a \in \mathbb{R}$ and let Λ_ρ be the measure Λ restricted to $D_\rho := (a - \rho, a + \rho) \times (-\pi, 0)$. The total mass of Λ_ρ is 4ρ . For any $\varepsilon_0 > 0$, there exists $n_0 < \infty$ such that for any $n > n_0$ and any simple symmetric function K with parameter n such that $[a - \rho, a + \rho]$ is one of the intervals $[x_1^k, x_2^k]$, the following holds:*

- (i) *There exist a compact set $N \in \mathcal{M}_1$ and a set $D_\rho^* \subset D_\rho$ such that $\Lambda_\rho(D_\rho^*) < 4\rho\varepsilon_0$.*
- (ii) *The function $K_N(x, \alpha)$ corresponding to N is defined on $D_\rho \setminus D_\rho^*$.*
- (iii) *For all $(x, \alpha) \in D_\rho \setminus D_\rho^*$, we have $|K_N(x, \alpha) - K(x, \alpha)| \leq \varepsilon_0$.*
- (iv) *$N \subset (a - \rho, a + \rho) \times (-\rho_1, 0)$.*

Proof. By translation invariance, it is enough to discuss the case $a = 0$.

Recall that we consider $(-\rho, \rho)$ to be a subset of the horizontal axis L_* . Let N_0 be a new name for the set satisfying Lemma 4.2 and let M be the closure of the convex hull of $N_0 \cup (-\rho, \rho)$. Since c_1 in the proof of Lemma 4.2 is strictly positive, $M \cap L_* = [-\rho, \rho]$. Let r_0 be the diameter of M and let $M_r = \{x \in \mathbb{R}^2 : \exists y \in M \text{ such that } x = (r/r_0)y\}$, i.e., M_r is a dilation of M .

We will define sets M_r^j . Each of these sets will be a horizontal shift of M_r ; that is, $M_r^j = M_r + (b, 0)$ for some $b \in \mathbb{R}$ depending on j and r . For any such set, we let $\widetilde{M}_r^j = M_r^j \cap L_*$. Note that the length of \widetilde{M}_r^j is $2(r/r_0)\rho$.

The sets M_r^j will be grouped into a countable number of families, with all sets in one family having the same size. We will pack smaller sets among the bigger sets in a tight way, as much as possible. The intersections of the sets M_r^j with L_* will form a pattern qualitatively similar to holes in the Cantor set; see Figure 3.

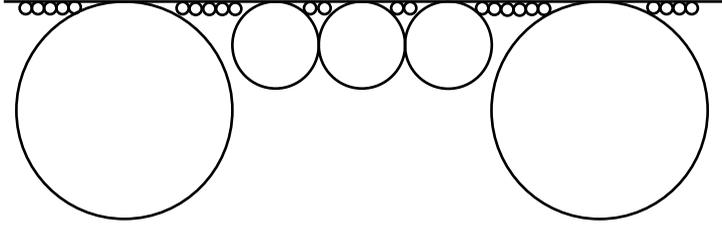


FIGURE 3. Sets $M_{r_k}^j$ of three “generations”. The sets M_r^j are convex by definition. They are pictured as circles but in fact they are not discs.

For $r \in (0, 1)$, let n_r be the maximum number k of disjoint sets $M_r^1, M_r^2, \dots, M_r^k$ such that $\bigcup_{1 \leq j \leq k} M_r^j \subset \{(x_1, x_2) : |x_1| < 1\}$. It is easy to see that for some $c_2, r_* > 0$ and all $r \in (0, r_*)$ we have $rn_r > c_2$. It follows that for some $c_3 > 0$, for every $r \in (0, r_*)$ one can find a family of disjoint sets $M_r^1, M_r^2, \dots, M_r^k$ such that $\bigcup_{1 \leq j \leq k} M_r^j \subset \{(x_1, x_2) : |x_1| < 1\}$ and the total length of $\bigcup_{1 \leq j \leq k} \widetilde{M}_r^j$ is greater than c_3 .

Choose $r_1 \in (0, \rho_1)$ and a family of disjoint sets $\{M_{r_1}^j\}_{1 \leq j \leq j_1}$ such that $\bigcup_{1 \leq j \leq j_1} M_{r_1}^j \subset \{(x_1, x_2) : |x_1| < \rho\}$ and the total length of $\bigcup_{1 \leq j \leq j_1} \widetilde{M}_{r_1}^j$ is greater than $c_3\rho$.

We proceed by induction. Let λ_1 be the Lebesgue measure of the set $[-\rho, \rho] \setminus \bigcup_{1 \leq j \leq j_1} \widetilde{M}_{r_1}^j$. We find $r_2 \in (0, r_1)$ and a family of sets $\{M_{r_2}^j\}_{1 \leq j \leq j_2}$ such that

- (i) $\bigcup_{1 \leq j \leq j_2} M_{r_2}^j \subset \{(x_1, x_2) : |x_1| < \rho\}$,
- (ii) all sets in the family $\{M_{r_k}^j\}_{k=1,2; 1 \leq j \leq j_k}$ are disjoint, and
- (iii) the total length of $\bigcup_{1 \leq j \leq j_2} \widetilde{M}_{r_2}^j$ is greater than $c_3\lambda_1/2$.

The general inductive step is the following. Suppose that we have defined $r_k > 0$ and families $\{M_{r_k}^j\}_{1 \leq j \leq j_k}$ for $k = 1, 2, \dots, i$. Let λ_i be the Lebesgue measure of the set $[-\rho, \rho] \setminus \bigcup_{1 \leq k \leq i} \bigcup_{1 \leq j \leq j_k} \widetilde{M}_{r_k}^j$. We find $r_{i+1} \in (0, r_i)$ and a family of sets $\{M_{r_{i+1}}^j\}_{1 \leq j \leq j_{i+1}}$ such that

- (i) $\bigcup_{1 \leq j \leq j_{i+1}} M_{r_{i+1}}^j \subset \{(x_1, x_2) : |x_1| < \rho\}$,
- (ii) all sets in the family $\{M_{r_k}^j\}_{1 \leq k \leq i+1; 1 \leq j \leq j_k}$ are disjoint, and
- (iii) the total length of $\bigcup_{1 \leq j \leq j_{i+1}} \widetilde{M}_{r_{i+1}}^j$ is greater than $c_3\lambda_i/2$.

Let T_r^j be the linear transformation that maps M onto M_r^j and let $N_r^j = T_r^j(N_0)$. Let $N_* = \bigcup_{k \geq 1} \bigcup_{1 \leq j \leq j_k} N_{r_k}^j$. It is easy to see that the Lebesgue measure of the set $[-\rho, \rho] \setminus \bigcup_{k \geq 1} \bigcup_{1 \leq j \leq j_k} \widetilde{M}_{r_k}^j$ is zero. It is also clear that $N_* \subset \{(x_1, x_2) : |x_1| \leq \rho\}$.

Let $\widehat{D} \subset D_\rho$ be such that $\Lambda_\rho(\widehat{D}) \geq 4\rho(1 - \varepsilon_0)$ and for all $(x, \alpha) \in \widehat{D}$, we have $|K_{N_0}(x, \alpha) - K(x, \alpha)| \leq \varepsilon_0$. Let $\widehat{D}_r = \{(x, \alpha) : \exists(y, \alpha) \in \widehat{D} \text{ such that } x = (r/r_0)y\}$. If $M_r^j = M_r + (b_r^j, 0)$, then we let $\widehat{D}_r^j = \{(x, \alpha) : \exists(y, \alpha) \in \widehat{D} \text{ such that } x = y + (b_r^j, 0)\}$. Let Λ_r^j be the measure Λ restricted to $(-(r/r_0)\rho + b_r^j, (r/r_0)\rho + b_r^j) \times (-\pi, 0)$. By the scaling and shift invariance of reflections, we see that for every $r > 0$ and j we have $\Lambda_r^j(\widehat{D}_r^j) \geq (r/r_0)4\rho(1 - \varepsilon_0)$ and for all $(x, \alpha) \in \widehat{D}_r^j$, we have $|K_{N_*}(x, \alpha) - K(x, \alpha)| \leq \varepsilon_0$. Let $D_\rho^* = D_\rho \setminus \bigcup_{k \geq 1} \bigcup_{1 \leq j \leq j_k} \widehat{D}_{r_k}^j$. Then $\Lambda_\rho(D_\rho^*) < 4\rho\varepsilon_0$, and for all $(x, \alpha) \in D_\rho \setminus D_\rho^*$, we have $|K_{N_*}(x, \alpha) - K(x, \alpha)| \leq \varepsilon_0$.

The set N_* satisfies all assertions in the lemma except that it does not belong to \mathcal{M}_1 because it consists of an infinite number of curves (arcs of ellipses) and hence it does not satisfy condition (M2) of Section 2. To address this problem, we let $N = \bigcup_{1 \leq k \leq k_0} \bigcup_{1 \leq j \leq j_k} N_{r_k}^j$ for some $k_0 < \infty$. It is easy to see that N and the accompanying function K_N satisfy all the conditions stated in the lemma if k_0 is sufficiently large. \square

Proof of Theorem 2.3. Recall the definition of a simple symmetric function from the beginning of Section 4. Suppose that $\tilde{\mathbb{P}}$ satisfies (2.4). Recall that $\int_{-\pi}^0 \sin \alpha d\alpha = -2$, let $m \geq 1$ be an integer and let γ_k^m be defined by $-\int_{-\pi}^{\gamma_k^m} \sin \alpha d\alpha = k2^{-m}$ for $k = 0, 1, \dots, 2^{m+1}$. If we take $f(\alpha) = \mathbf{1}_{(\gamma_n^m, \gamma_{n+1}^m]}(\alpha)$ and $g(\alpha) = \mathbf{1}_{(\gamma_j^m, \gamma_{j+1}^m]}(\alpha)$, then (2.4) yields

$$(5.1) \quad \int_{(\gamma_j^m, \gamma_{j+1}^m]} \int_{(\gamma_n^m, \gamma_{n+1}^m]} \tilde{\mathbb{P}}(x, \alpha; d\beta) \sin \alpha d\alpha = \int_{(\gamma_n^m, \gamma_{n+1}^m]} \int_{(\gamma_j^m, \gamma_{j+1}^m]} \tilde{\mathbb{P}}(x, \alpha; d\beta) \sin \alpha d\alpha.$$

For a fixed integer $-\infty < k < \infty$, let

$$a(n, j) = - \int_{[k/m, (k+1)/m]} \int_{(\gamma_n^m, \gamma_{n+1}^m]} \int_{(\gamma_j^m, \gamma_{j+1}^m]} \tilde{\mathbb{P}}(x, \alpha; d\beta) \sin \alpha d\alpha dx.$$

It follows from (5.1) that $a(n, j) = a(j, n)$. Let $\beta_{n,0}^m = \gamma_n^m$ and let $\beta_{n,i}^m$ be defined by

$$- \int_{(k/m, (k+1)/m]} \int_{\gamma_n^m}^{\beta_{n,i}^m} \sin \alpha d\alpha dx = \sum_{0 \leq j \leq i-1} a(n, j)$$

for $i = 1, \dots, 2^{m+1}$. Note that $\int_{(-\pi, 0]} \tilde{\mathbb{P}}(x, \alpha; d\beta) = 1$ so

$$\sum_{0 \leq j \leq 2^{m+1}-1} a(n, j) = - \int_{[k/m, (k+1)/m]} \int_{(\gamma_n^m, \gamma_{n+1}^m]} \sin \alpha d\alpha dx,$$

and, therefore, $\beta_{n,2^{m+1}}^m = \gamma_{n+1}^m$.

Let $\hat{\beta}_{n,i}^m = (\beta_{n,i}^m + \beta_{n,i+1}^m)/2$ and $\hat{K}_m(x, \alpha) = (x, \hat{\beta}_{n,j}^m)$ for $(x, \alpha) \in (k/m, (k+1)/m] \times (\beta_{j,n}^m, \beta_{j,n+1}^m]$. Since $a(n, j) = a(j, n)$, we have

$$\int_{(\beta_{j,n}^m, \beta_{j,n+1}^m)} \sin \alpha d\alpha = \int_{(\beta_{n,j}^m, \beta_{n,j+1}^m)} \sin \alpha d\alpha.$$

This shows that $\hat{K}_m(x, \alpha)$ is a simple symmetric function according to Definition 4.1.

Recall that for $K : D \rightarrow D$, the kernel \mathbb{P}_K is defined by

$$\mathbb{P}_K(x, \alpha; dy, d\beta) = \delta_{K(x, \alpha)}(y, \beta) dy d\beta.$$

It is elementary to check that $\mathbb{P}_{\hat{K}_n}(x, \alpha; dy, d\beta) \rightarrow \delta_x(y) dy \tilde{\mathbb{P}}(x, \alpha; d\beta)$ weakly on D as $n \rightarrow \infty$. Note that we do not claim that \hat{K}_n arise as functions associated to reflecting sets $M \in \mathcal{M}_2$.

To prove Theorem 2.3, it will suffice to show that there exists a sequence of sets $M_n \in \mathcal{M}_2$, $n \geq 1$, satisfying conditions (a) and (b) of Theorem 2.3 and such that if the functions $K_n : D \rightarrow D$ correspond to M_n 's, then $K_n - \hat{K}_n \rightarrow 0$ pointwise.

Fix some \widehat{K}_n and let (x_1^k, x_2^k) be one of the corresponding intervals as in Definition 4.1. Let $D_k = (x_1^k, x_2^k) \times (-\pi, 0)$. According to Lemma 5.1 and its proof, there exist a compact set $N_k \in \mathcal{M}_1$ and a set $D_k^* \subset D_k$ such that $\Lambda(D_k^*) < \varepsilon_0 2^{-k}$, and the following assertions hold:

- (i) The function $K_{N_k}(x, \alpha)$ corresponding to N_k is defined on $D_k \setminus D_k^*$.
- (ii) For all $(x, \alpha) \in D_k \setminus D_k^*$, we have $|K_{N_k}(x, \alpha) - \widehat{K}_n(x, \alpha)| \leq \varepsilon_0$, and the light ray with the starting position and direction represented by (x, α) reflects exactly twice from N_k before returning to L_* .
- (iii) $N_k \subset (x_1^k, x_2^k) \times (-1/n, 0)$.

Note that for $j \neq k$, either the sets N_j and N_k do not intersect or they are identical. Moreover, the construction presented in Lemma 5.1 allows us to assume that light rays that start in $D_k \setminus D_k^*$ do not leave $(x_1^k, x_2^k) \times (-\infty, 0)$. In this sense, the families of mirrors N_k do not interfere with one another. Let $M_n^* = \bigcup_{k \geq 1} N_k$ and let $K_n^*(x, \alpha)$ be the function on D corresponding to M_n^* . Let $D^* = \bigcup_{k \geq 1} D_k^*$ and note that $\Lambda(D^*) \leq \varepsilon_0$. We have the following:

- (i) The function $K_n^*(x, \alpha)$ is defined on $D \setminus D^*$.
- (ii) For all $(x, \alpha) \in D \setminus D^*$, we have $|K_n^*(x, \alpha) - \widehat{K}_n(x, \alpha)| \leq \varepsilon_0$, and the light ray with the starting position and direction represented by (x, α) reflects exactly twice from M_n^* before returning to L_* .
- (iii) $M_n^* \subset \mathbb{R} \times (-1/n, 0)$.

It is easy to see that $M_n^* \in \mathcal{M}_1$ for every n , but it is not necessarily true (actually unlikely) that $M_n^* \in \mathcal{M}_2$. We can instead consider sets

$$M_n = M_n^* \cup \{(x_1, x_2) : x_2 = -1/n\} \cup \bigcup_k \{(x_1, x_2) : x_1 = x_1^k \text{ or } x_2^k, -1/n \leq x_2 \leq 0\}.$$

It is easy to see that functions K_n corresponding to M_n satisfy all properties (i)–(iii) listed above. Moreover, M_n satisfy condition (M5) of Section 2 so $M_n \in \mathcal{M}_2$, according to Proposition 2.1. This completes the proof of the theorem. \square

ACKNOWLEDGMENT

The authors are grateful to Persi Diaconis and John Sylvester for their most helpful advice. The authors thank the anonymous referee for many suggestions for improvement.

REFERENCES

- [1] K. Burdzy, R. Bass, Z. Chen and M. Hairer, Stationary distributions for diffusions with inert drift. *Probab. Theory Rel. Fields* **146** (2010) 1–47. MR2550357 (2011b:60226)
- [2] S. Chandrasekhar, *Radiative transfer*. Dover Publications, Inc., New York, 1960. MR0111583 (22:2446)
- [3] N. Chernov and R. Markarian, *Chaotic billiards*. Mathematical Surveys and Monographs, 127. American Mathematical Society, Providence, RI, 2006. MR2229799 (2007f:37050)
- [4] F. Comets and S. Popov, Ballistic regime for random walks in random environment with unbounded jumps and Knudsen billiards. *Ann. Inst. H. Poincaré Probab. Statist.* (2012) (to appear).
- [5] F. Comets, S. Popov, G. Schütz and M. Vachkovskaia, Billiards in a general domain with random reflections. *Arch. Rat. Mech. Anal.* **191**, (2009) 497–537. MR2481068 (2010h:37077)
- [6] F. Comets, S. Popov, G. Schütz and M. Vachkovskaia, Erratum: Billiards in a general domain with random reflections. *Arch. Ration. Mech. Anal.* **193**, (2009), 737–738. MR2525117 (2010h:37078)

- [7] F. Comets, S. Popov, G. Schütz and M. Vachkovskaia, Quenched invariance principle for the Knudsen stochastic billiard in a random tube *Ann. Probab.* **38** (2010) 1019–1061. MR2674993 (2012c:60244)
- [8] F. Comets, S. Popov, G. Schütz and M. Vachkovskaia, Knudsen gas in a finite random tube: transport diffusion and first passage properties. *J. Stat. Phys.* **140** (2010) 948–984. MR2673342 (2012b:60317)
- [9] S. Evans, Stochastic billiards on general tables. *Ann. Appl. Probab.* **11**, 419–437 (2001). MR1843052 (2002f:60134)
- [10] K. Falconer, *Fractal geometry. Mathematical foundations and applications. Second edition.* John Wiley & Sons, Inc., Hoboken, NJ, 1990. MR1102677 (92j:28008)
- [11] R. Feres, Random walks derived from billiards. *Dynamics, ergodic theory, and geometry*, 179–222, *Math. Sci. Res. Inst. Publ.*, **54**, Cambridge Univ. Press, Cambridge, 2007. MR2369447 (2009c:37033)
- [12] R. Feres and H.-K. Zhang, The spectrum of the billiard Laplacian of a family of random billiards. *J. Stat. Phys.* **141**, (2010) 1039–1054. MR2740402 (2011h:37052)
- [13] R. Klette, K. Schlüns and A. Koschan, *Computer vision. Three-dimensional data from images.* Springer-Verlag Singapore, Singapore, 1998. MR1729491 (2000j:68190)
- [14] S. Lalley and H. Robbins, Stochastic search in a convex region. *Probab. Theory Relat. Fields* **77**, 99–116 (1988). MR921821 (89f:90078)
- [15] M. Lapidus and R. Niemeyer, Towards the Koch snowflake fractal billiard: computer experiments and mathematical conjectures. *Gems in experimental mathematics*, 231–263, *Contemp. Math.*, **517**, Amer. Math. Soc., Providence, RI, 2010. MR2731085 (2012b:37101)
- [16] M. Lapidus and R. Niemeyer, Families of Periodic Orbits of the Koch Snowflake Fractal Billiard. [arXiv:1105.0737](https://arxiv.org/abs/1105.0737)
- [17] S. Tabachnikov, *Geometry and billiards.* Student Mathematical Library, 30. American Mathematical Society, Providence, RI; Mathematics Advanced Study Semesters, University Park, PA, 2005. MR2168892 (2006h:51001)
- [18] Wikipedia, *Retroreflector*, <http://en.wikipedia.org/wiki/Retroreflector> Online; accessed 4-March-2012.
- [19] Wikipedia, *Scotchlite*, <http://en.wikipedia.org/wiki/Scotchlite> Online; accessed 4-March-2012.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, 121 - 1984 MATHEMATICS ROAD, VANCOUVER, BRITISH COLUMBIA, CANADA V6T 1Z2

E-mail address: angel@math.ubc.ca

DEPARTMENT OF MATHEMATICS, BOX 354350, UNIVERSITY OF WASHINGTON, SEATTLE, WASHINGTON 98195

E-mail address: burdzy@math.washington.edu

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, 2-180, 77 MASSACHUSETTS AVENUE, CAMBRIDGE, MASSACHUSETTS 02139

E-mail address: sheffield@math.mit.edu