BILLIARDS IN POLYHEDRA: A METHOD TO CONVERT 2-DIMENSIONAL UNIFORMITY TO 3-DIMENSIONAL UNIFORMITY

J. BECK, W.W.L. CHEN, AND Y. YANG

Abstract. The class of 2-dimensional non-integrable flat dynamical systems has a rather extensive literature with many deep results, but the methods developed for this type of problems, both the traditional approach via Teichmüller geometry and our recent shortline-ancestor method, appear to be exclusively plane-specific. Thus we know very little of any real significance concerning 3-dimensional systems.

Our purpose here is to describe some very limited extensions of uniformity in 2 dimensions to uniformity in 3 dimensions. We consider a 3-manifold which is the cartesian product of the regular octagonal surface with the unit torus. This is a restricted system, in the sense that one of the directions is integrable. However, this restriction also allows us to make use of a transference theorem for arithmetic progressions established earlier by Beck, Donders and Yang.

1. Introduction

For rational polygons where every angle is a rational multiple of π, we have the following fundamental result of Kerckhoff, Masur and Smillie [4] in 1986.

Theorem A. *Let P be a rational polygon. For almost every initial direction and for every non-pathological starting point for this direction, the half-infinite billiard orbit in P is uniformly distributed.*

Given any initial direction, a point $\mathbf{p}_0 \in P$ is called a pathological starting point for this direction if the half-infinite billiard orbit starting from $\mathbf{p}_0$ and with this direction hits a singularity of P. Otherwise the point $\mathbf{p}_0$ is called a non-pathological starting point for this direction. It is easy to see that for any given direction, almost every point in P is a non-pathological starting point.

The proof of Theorem A consists of essentially three steps. The first step is to establish the ergodicity of the corresponding interval exchange transformation. The second step is to use the well known Birkhoff ergodic theorem. The final step is to extend ergodicity to unique ergodicity.

Our aim is to convert Theorem A to a result concerning equidistribution of 3-dimensional billiard in some polyhedra. However, we need to restrict our discussion to rational polygonal right prisms. A rational polygonal right prism is a region in 3-dimensional cartesian space of the form

$$M = P \times I = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in P \text{ and } z \in I\}, \tag{1.1}$$

where P is a rational polygon and $I = [0, z_0]$ is an interval.

As the rational polygonal right prism $M = P \times I$ is integrable in the direction of the interval I, our extension is somewhat limited.

2010 *Mathematics Subject Classification.* 11K38, 37E35.
Key words and phrases. geodesics, billiards, uniformity.

Theorem 1. *Let M be a rational polygonal right prism of the form* (1.1)*, where P is a rational polygon and $I = [0, z_0]$ is an interval. For almost every pair of initial direction and starting point, the half-infinite billiard orbit in M is uniformly distributed.*

For illustration, we consider a special case where P is a right triangle. It is well known that the right triangle billiard with angle $\pi/4$ and the right triangle billiard with angle $\pi/6$ are the only right triangle billiards that are integrable, exhibiting stable and predictable behaviour. Perhaps the simplest non-integrable billiard is the right triangle billiard with angle $\pi/8$. It is also well known that unfolding in the spirit of König and Szücs [5] leads to a 16-fold covering of the triangle and shows that this billiard is equivalent to geodesic flow on the regular octagon surface $\mathcal{P}$ where parallel edges are identified in pairs. On the other hand, unfolding also leads to a 2-fold covering of the interval $I = [0, z_0]$. Thus billiard in the rational polygonal right prism $M = P \times I$, where P is the right triangle with angle $\pi/8$ and $I = [0, z_0]$, is equivalent to geodesic flow in the translation 3-manifold $\mathcal{M} = \mathcal{P} \times \mathcal{I}$, where $\mathcal{P}$ is the regular octagon translation surface and $\mathcal{I} = [0, 2z_0]$, treated as a torus. Figure 1 illustrates that $\mathcal{M} = \mathcal{P} \times \mathcal{I}$ gives a 32-fold covering of the rational polygonal right prism $M = P \times I$. It has 2 octagonal faces which are identified with each other, and 8 rectangular faces, with pairs of parallel ones identified with each other, analogous to the edge identification of the regular octagon translation surface $\mathcal{P}$.

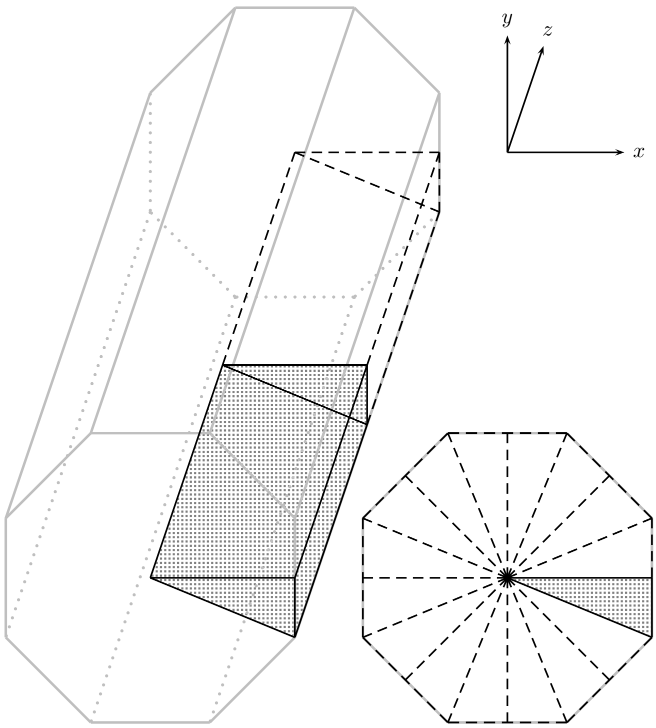

Figure 1: the translation 3-manifold $\mathcal{M} = \mathcal{P} \times \mathcal{I}$

Theorem 2. *Let $\mathcal{M} = \mathcal{P} \times \mathcal{I}$ be a rational octagonal right prism translation* 3*-manifold, where $\mathcal{P}$ is the regular octagon translation surface and $\mathcal{I} = [0, 2z_0]$, treated as a torus. For almost every pair of direction and starting point, the half-infinite geodesic in $\mathcal{M}$ is uniformly distributed.*

2. Proof of Theorem 2

Suppose that a half-infinite geodesic

$$\mathcal{L}_{S_0,\mathbf{v}}(t) = (s_1 + v_1 t, s_2 + v_2 t, s_3 + v_3 t), \quad t \geqslant 0,$$

in $\mathcal{M}$ has a non-pathological starting point $\mathcal{L}_{S_0,\mathbf{v}}(0) = S_0 = (s_1, s_2, s_3)$, and direction given by the unit vector

$$\mathbf{v} = (v_1, v_2, v_3) \in \mathbb{R}^3, \quad \text{where} \quad v_1^2 + v_2^2 + v_3^2 = 1,$$

with arc-length parametrization. The coordinates $(s_1 + v_1 t, s_2 + v_2 t)$ are modulo $\mathcal{P}$ and the coordinate $s_3 + v_3 t$ is modulo $\mathcal{I}$.

We may assume without loss of generality that $v_3 > 0$. Then the geodesic $\mathcal{L}_{S_0,\mathbf{v}}(t)$, $t \geqslant 0$, hits the octagon face of $\mathcal{M}$ for the very first time at time $t = t_0$, where $s_3 + v_3 t_0 = 2z_0$, so that $t_0 = (2z_0 - s_3)/v_3$. Indeed, the geodesic hits the octagon face of $\mathcal{M}$ for the $(k+1)$-th time at time $t = t_k$, where

$$t_k = \frac{2kz_0 + (2z_0 - s_3)}{v_3} = k\theta + \lambda, \quad k = 0, 1, 2, 3, \ldots, \tag{2.1}$$

with parameters

$$\theta = \frac{2z_0}{v_3} \quad \text{and} \quad \lambda = \frac{2z_0 - s_3}{v_3}.$$

This gives rise to an arithmetic progression

$$\lambda < \theta + \lambda < 2\theta + \lambda < 3\theta + \lambda < \ldots,$$

with common gap θ between consecutive terms.

We need the following result on arithmetic progressions; see [1, Theorem 2.2.2].

Lemma 2.1 (transference theorem for arithmetic progressions). *Let $\mathcal{S} \subset \mathbb{R}$ be a measurable set. For every $\ell \in \mathbb{Z}$, there exists a constant $c_1(\ell) > 0$, dependent only on ℓ, such that for almost every pair θ, λ satisfying $2^\ell \leqslant \theta < 2^{\ell+1}$ and $0 \leqslant \lambda < \theta$, the inequality*

$$\left| \sum_{\substack{k \geqslant 0 \\ k\theta+\lambda \in \mathcal{S}\cap[0,n]}} 1 - \frac{1}{\theta}\,\mathrm{meas}(\mathcal{S} \cap [0,n]) \right| \leqslant c_1(\ell) n^{3/4} (\log n)^{1/2}$$

holds for every sufficiently large positive integer n.

We have following immediate consequence.

Lemma 2.2. *Suppose that the set $\mathcal{S} \subset \mathbb{R}$ is measurable. Suppose further that $\mathcal{S}$ has asymptotic density $d = d(\mathcal{S}) \in [0,1]$, so that there exists a monotonic sequence $\varepsilon(n) = \varepsilon(\mathcal{S}; n) \to 0$ as $n \to \infty$ such that*

$$\left| \frac{1}{n}\,\mathrm{meas}(\mathcal{S} \cap [0,n]) - d(\mathcal{S}) \right| < \varepsilon(n), \quad n = 1, 2, 3, \ldots.$$

For every $\ell \in \mathbb{Z}$, there exists a constant $c_2(\ell) > 0$, dependent only on ℓ, such that for almost every pair θ, λ satisfying $2^\ell \leqslant \theta < 2^{\ell+1}$ and $0 \leqslant \lambda < \theta$, the inequality

$$\left| \frac{\theta}{n} \sum_{\substack{k \geqslant 0 \\ k\theta+\lambda \in \mathcal{S}\cap[0,n]}} 1 - d(\mathcal{S}) \right| \leqslant \varepsilon(n) + c_2(\ell) \frac{(\log n)^{1/2}}{n^{1/4}} \tag{2.2}$$

holds for every sufficiently large positive integer n.

Meanwhile, every point in $\mathcal{M}$ is of the form (x,y,z), where $(x,y) \in \mathcal{P}$ and $z \in \mathcal{I}$. We consider the projection

$$\phi : \mathcal{M} \to \mathcal{P} : (x,y,z) \mapsto (x,y). \tag{2.3}$$

Then the image of the geodesic $\mathcal{L}_{S_0,\mathbf{v}}(t)$, $t \geqslant 0$, under this projection is a half-infinite geodesic

$$\mathcal{H}_{S_0,\mathbf{v}}(t) = (s_1 + v_1 t, s_2 + v_2 t), \quad t \geqslant 0, \tag{2.4}$$

on the regular octagon translation surface $\mathcal{P}$. Clearly the key parameters s_3 and v_3, particularly concerning the hitting times given in (2.1), are lost under this projection (2.3). However, we know the arithmetic progression (2.1) of the time instances when the geodesic $\mathcal{L}_{S_0,\mathbf{v}}(t)$, $t \geqslant 0$, hits the octagon face of $\mathcal{M}$. This gives rise to an infinite sequence of points $\mathcal{H}_{S_0,\mathbf{v}}(t_k)$, $k = 0,1,2,3,\ldots$, on $\mathcal{P}$. For any S_0 and $\mathbf{v}$, if we can show that this sequence of points is uniformly distributed on $\mathcal{P}$, then the half-infinite geodesic $\mathcal{L}_{S_0,\mathbf{v}}(t)$, $t \geqslant 0$, is uniformly distributed in $\mathcal{M}$.

Consider a typical half-infinite geodesic $\mathcal{H}_{\mathbf{w}}(\tau)$, $\tau \geqslant 0$, on $\mathcal{P}$, with direction given by the unit vector $\mathbf{w} = (w_1, w_2) \in \mathbb{R}^2$. Suppose that this geodesic is the image on $\mathcal{P}$ of $\mathcal{L}_{S_0,\mathbf{v}}(t)$, $t \geqslant 0$, under the projection (2.3). Then

$$\mathcal{H}_{\mathbf{w}}(\tau) = (s_1 + w_1 \tau, s_2 + w_2 \tau), \quad \tau \geqslant 0. \tag{2.5}$$

In view of the different parametrizations of (2.4) and (2.5), we have

$$\mathcal{H}_{S_0,\mathbf{v}}(t) = \mathcal{H}_{\mathbf{w}}(\tau) \quad \text{if and only if} \quad \tau = (v_1^2 + v_2^2)^{1/2} t.$$

Corresponding to the arithmetic progression t_k, $k = 0,1,2,3,\ldots$, of hitting times given by (2.1) is the arithmetic progression

$$\tau_k = (v_1^2 + v_2^2)^{1/2} t_k = \frac{(v_1^2 + v_2^2)^{1/2}(2kz_0 - s_3)}{v_3} = k\theta + \lambda, \quad k = 0,1,2,3,\ldots,$$

with parameters

$$\theta = \frac{2z_0(v_1^2 + v_2^2)^{1/2}}{v_3} \quad \text{and} \quad \lambda = \frac{(2z_0 - s_3)(v_1^2 + v_2^2)^{1/2}}{v_3}. \tag{2.6}$$

By the geodesic analogue of Theorem A, for almost every direction $\mathbf{w}$ and for every non-pathological starting point for this direction, the geodesic $\mathcal{H}_{\mathbf{w}}(\tau)$, $\tau \geqslant 0$, is uniformly distributed on $\mathcal{P}$. Let RP denote an arbitrary polygon on $\mathcal{P}$ where all the vertices have rational coordinates, and let the measurable set

$$\mathcal{S} = \mathcal{S}(\mathcal{H}_{\mathbf{w}}; \mathrm{RP}) = \{\tau \geqslant 0 : \mathcal{H}_{\mathbf{w}}(\tau) \in \mathrm{RP}\} \tag{2.7}$$

denote the set of time instances when this geodesic visits RP. The uniformity of the geodesic then implies that $\mathcal{S}$ has asymptotic density

$$d = d(\mathcal{S}) = \frac{\text{area}(\mathrm{RP})}{\text{area}(\mathcal{P})} \in [0,1].$$

The uniformly distributed geodesic $\mathcal{H}_{\mathbf{w}}(\tau)$, $\tau \geqslant 0$, is clearly the image on $\mathcal{P}$ under the projection (2.3) of infinitely many different geodesics $\mathcal{L}_{S_0,\mathbf{v}}(t)$, $t \geqslant 0$, in the 3-manifold $\mathcal{M}$, as there are only two requirements, namely $\mathcal{H}_{\mathbf{w}}(0) = (s_1, s_2)$ concerning the starting point, where $S_0 = (s_1, s_2, s_3)$, and $w_1 v_2 = w_2 v_1$ concerning equality of the relevant directions. Applying Lemma 2.2, we see that for every $\ell \in \mathbb{Z}$ and for almost every pair θ, λ of the form (2.6) satisfying $2^\ell \leqslant \theta < 2^{\ell+1}$ and $0 \leqslant \lambda < \theta$, the inequality (2.2) with $\mathcal{S}$ given by (2.7) holds for every sufficiently large positive integer n.

This means that for almost every pair of starting point S_0 and unit direction vector $\mathbf{v}$, the infinite sequence

$$\mathcal{L}_{S_0;\mathbf{v}}(t_k), \quad k = 0, 1, 2, 3, \ldots, \tag{2.8}$$

of points, where the sequence t_k, $k = 0, 1, 2, 3, \ldots$, of time instances is given by (2.1), is uniformly distributed on $\mathcal{P}$ relative to the single test set RP.

The set of all polygons RP on $\mathcal{P}$ where all the vertices have rational coordinates is countable. On the other hand, a countable union of sets of measure zero has measure zero. It follows that for almost every pair of starting point S_0 and unit direction vector $\mathbf{v}$, the infinite sequence (2.8) of points is uniformly distributed on $\mathcal{P}$ relative to every polygon RP on $\mathcal{P}$ where all the vertices have rational coordinates. This guarantees uniformity in general, in the classical Weyl sense, and completes the proof of Theorem 2.

Remark. Theorem 1 is a result on time-qualitative uniformity, and does not say anything about the speed of convergence to uniform distribution, as a key ingredient of the proof is the geodesic analogue of Theorem A which is also time-qualitative in nature. There are instances, however, when we can establish time-quantitative results. This happens, for example, when it is possible to establish time-quantitative uniform distribution results for some geodesics on the underlying rational polygonal translation surface $\mathcal{P}$. We can establish such extensions of the geodesic analogue of Theorem A in [1, 2] for the L-surface and in [3] for finite polysquare translation surfaces and for regular polygonal translation surfaces. These in turn lead to extensions of various analogues of Theorem 2, and hence also Theorem 1, to time-quantitative results.

3. Proof of Lemma 2.1

Throughout the proof, the set $\mathcal{S} \subset \mathbb{R}$ is measurable.

Let a non-negative integer ℓ be chosen and fixed.

Consider an infinite sequence $N_1, N_2, N_3, \ldots$ of positive integers satisfying

$$1 < N_1 < N_2 < N_3 < \ldots < N_h < \ldots,$$

and another infinite sequence $M_1, M_2, M_3, \ldots$ of positive integers satisfying

$$1 < M_h < N_h, \quad h = 1, 2, 3, \ldots,$$

both to be specified later in terms of the parameter h and the chosen integer ℓ. For every positive integer h, let $S(h) \subset [0, 1]$ denote the *contraction* of $\mathcal{S} \cap [0, N_h]$ to the unit interval, so that

$$x \in S(h) \quad \text{if and only if} \quad N_h x \in \mathcal{S} \cap [0, N_h]. \tag{3.1}$$

Since the characteristic function

$$\chi_{S(h)}(x) = \begin{cases} 1, & \text{if } x \in S(h), \\ 0, & \text{if } x \notin S(h), \end{cases}$$

defined over $[0, 1]$ and extended periodically over the whole real line with period 1, is measurable, we can consider its Fourier series

$$\chi_{S(h)}(x) = \sum_{j \in \mathbb{Z}} a_j \mathrm{e}^{2\pi \mathrm{i} j x}, \tag{3.2}$$

with Fourier coefficients a_j, $j \in \mathbb{Z}$. In particular,

$$a_0 = \lambda_1(S(h)).$$

Remark. For a measurable set $S(h)$, the infinite Fourier series (3.2) may diverge at some points. However, Lemma 2.1 is a measure theoretic statement which ignores sets of measure zero. So it suffices to have pointwise convergence almost everywhere. Fourier analysis provides at least two options to settle this issue. We can use the very deep Carleson theorem. Alternatively, we can use the much simpler Lebesgue theorem with Cesàro summability.

The Parseval formula gives

$$\sum_{j\in\mathbb{Z}} |a_j|^2 = \lambda_1(S(h)) \leqslant 1,$$

so that

$$\sum_{j\in\mathbb{Z}\setminus\{0\}} |a_j|^2 = \lambda_1(S(h)) - \lambda_1^2(S(h)) = \lambda_1(S(h))(1-\lambda_1(S(h))) < 1. \tag{3.3}$$

Lemma 2.1 for the chosen value ℓ concerns the arithmetic progression $k\theta+\eta$, $k \geqslant 0$, where $2^\ell \leqslant \theta < 2^{\ell+1}$ and $0 \leqslant \eta < \theta$. The contraction (3.1) leads to a new arithmetic progression $k\alpha+\beta$, $k\geqslant 0$, where $2^\ell/N_h \leqslant \alpha < 2^{\ell+1}/N_h$ and $0 \leqslant \beta < \alpha$.

For any $\alpha \in [2^\ell/N_h, 2^{\ell+1}/N_h)$, let $K(\alpha)$ be the unique integer satisfying

$$(K(\alpha)-1)\alpha < 1 \leqslant K(\alpha)\alpha. \tag{3.4}$$

Using the Fourier series (3.2), we have

$$\sum_{k=0}^{K(\alpha)-1} \chi_{S(h)}(k\alpha+\beta) - K(\alpha)\lambda_1(S(h)) = \sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi \mathrm{i} j(k\alpha+\beta)} \tag{3.5}$$

for every α and β satisfying $2^\ell/N_h \leqslant \alpha < 2^{\ell+1}/N_h$ and $0 \leqslant \beta < \alpha$.

To study (3.5), we consider the integral

$$J(\mathbf{a};N_h;M_h) = \int_{2^\ell/N_h}^{2^{\ell+1}/N_h} \int_{-2^{\ell-1}/M_h}^{2^{\ell-1}/M_h} \left| \sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi \mathrm{i} j(k\alpha+\gamma)} \right|^2 \mathrm{d}\gamma\,\mathrm{d}\alpha. \tag{3.6}$$

To obtain a bound on this integral, we observe that for $\gamma \in [-2^{\ell-1}/M_h, 2^{\ell-1}/M_h]$, the inequality $2(1-|\gamma|M_h/2^\ell) \geqslant 1$ holds. It then follows that

$$J(\mathbf{a};N_h;M_h) \leqslant 2J^*(\mathbf{a};N_h;M_h), \tag{3.7}$$

where for integers N and M satisfying $1 < M < N$,

$$J^*(\mathbf{a};N;M) = \int_{2^\ell/N}^{2^{\ell+1}/N} \int_{-2^\ell/M}^{2^\ell/M} \left| \sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi \mathrm{i} j(k\alpha+\gamma)} \right|^2 \left(1-\frac{|\gamma|M}{2^\ell}\right) \mathrm{d}\gamma\,\mathrm{d}\alpha. \tag{3.8}$$

At the end of this section, we establish the following bound for this integral.

Lemma 3.1. *For any sequence* $\mathbf{a}$ *satisfying* (3.3)*, the inequality*

$$J^*(\mathbf{a};N;M) \leqslant 2^{\ell+11}. \tag{3.9}$$

holds uniformly for integers M *and* N *satisfying* $1 < M < N$*, where* $K(\alpha)$ *is the integer defined by* (3.4) *for every* $\alpha \in [2^\ell/N, 2^{\ell+1}/N)$.

Combining (3.6), (3.7) and (3.9), we deduce that

$$J(\mathbf{a};N_h;M_h)\leqslant 2^{\ell+12}. \tag{3.10}$$

The inequality (3.10) is a quadratic average result, from which we can derive information concerning the majority of pairs

$$(\alpha,\gamma)\in\left[\frac{2^\ell}{N_h},\frac{2^{\ell+1}}{N_h}\right)\times\left[-\frac{2^{\ell-1}}{M_h},\frac{2^{\ell-1}}{M_h}\right]. \tag{3.11}$$

Let

$$B(N_h;M_h)=\left\{(\alpha,\gamma)\in\left[\frac{2^\ell}{N_h},\frac{2^{\ell+1}}{N_h}\right)\times\left[-\frac{2^{\ell-1}}{M_h},\frac{2^{\ell-1}}{M_h}\right]:\text{(3.12) holds}\right\}$$

denote the collection of pairs (α,γ) satisfying (3.11) such that

$$\left|\sum_{j\in\mathbb{Z}\setminus\{0\}}a_j\sum_{k=0}^{K(\alpha)-1}\mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\gamma)}\right|\geqslant(hN_hM_h)^{1/2}\log(1+h). \tag{3.12}$$

Then it follows from (3.10) that

$$\lambda_2(B(N_h;M_h))\leqslant\frac{2^{\ell+12}}{hN_hM_h\log^2(1+h)}, \tag{3.13}$$

where λ_2 denotes 2-dimensional Lebesgue measure.

Next, note that as we move from (3.5) to (3.6), we replace the parameter β over a short interval $[0,\alpha)$ by a parameter $\gamma\in[-2^{\ell-1}/M_h,2^{\ell-1}/M_h]$ over a longer interval. For any pair

$$(\alpha,\beta)\in\left[\frac{2^\ell}{N_h},\frac{2^{\ell+1}}{N_h}\right)\times[0,\alpha), \tag{3.14}$$

there are at least $N_h/2M_h$ values of $\gamma\in[-2^{\ell-1}/M_h,2^{\ell-1}/M_h]$ where $\{\gamma/\alpha\}\alpha=\beta$. For each of these values of γ, consider the two arithmetic progressions

$$k\alpha+\gamma,\quad k=0,1,2,3,\ldots,K(\alpha)-1, \tag{3.15}$$

and

$$k\alpha+\beta,\quad k=0,1,2,3,\ldots,K(\alpha)-1. \tag{3.16}$$

Since

$$\gamma=\left[\frac{\gamma}{\alpha}\right]\alpha+\beta,$$

the arithmetic progression (3.15) is obtained by simply advancing the arithmetic progression (3.16) by $[\gamma/\alpha]$ terms. More precisely, the arithmetic progression (3.15) is given by

$$k\alpha+\beta,\quad k=\left[\frac{\gamma}{\alpha}\right],\left[\frac{\gamma}{\alpha}\right]+1,\left[\frac{\gamma}{\alpha}\right]+2,\left[\frac{\gamma}{\alpha}\right]+3,\ldots,\left[\frac{\gamma}{\alpha}\right]+K(\alpha)-1. \tag{3.17}$$

Lemma 3.2. *If a pair* (α,β) *such that* (3.14) *holds satisfies the inequality*

$$\left|\sum_{j\in\mathbb{Z}\setminus\{0\}}a_j\sum_{k=0}^{K(\alpha)-1}\mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\beta)}\right|\geqslant\frac{2N_h}{M_h}+(hN_hM_h)^{1/2}\log(1+h), \tag{3.18}$$

then each pair (α,γ) *such that* (3.11) *and* $\{\gamma/\alpha\}\alpha=\beta$ *hold satisfies the inequality* (3.12)*.*

Proof. It clearly suffices to prove that

$$\left|\sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\beta)} - \sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\gamma)}\right| \leqslant \frac{2N_h}{M_h}. \tag{3.19}$$

Since

$$\left|\left[\frac{\gamma}{\alpha}\right]\right| \leqslant \frac{N_h}{2M_h},$$

it follows from (3.16) and (3.17) that those terms that belong to one of the arithmetic progressions (3.15) or (3.16) but not both then form two arithmetic progressions of the form

$$k\alpha+\rho, \quad k=0,1,2,3,\ldots,K-1,$$

where $K \leqslant N_h/2M_h$. Hence

$$\sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\beta)} - \sum_{k=0}^{K(\alpha)-1} \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\gamma)}$$

is the sum of two sums of the form

$$\sum_{k=0}^{K-1} \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\rho)},$$

where $K \leqslant N_h/2M_h$. Now for each of the two sums, we have

$$\begin{aligned}\left|\sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \sum_{k=0}^{K-1} \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\rho)}\right| &= \left|\sum_{k=0}^{K-1} \sum_{j\in\mathbb{Z}\setminus\{0\}} a_j \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\rho)}\right| \\ &= \left|\sum_{k=0}^{K-1}\left(\sum_{j\in\mathbb{Z}} a_j \mathrm{e}^{2\pi\mathrm{i}j(k\alpha+\rho)} - a_0\right)\right| \\ &\leqslant \sum_{k=0}^{K-1}\left|\chi_{S(h)}(k\alpha+\rho)-\lambda_1(S_h)\right| \\ &\leqslant 2K.\end{aligned}$$

This clearly leads to (3.19) and completes the proof. □

Let

$$A(N_h;M_h) = \left\{(\alpha,\beta)\in\left[\frac{2^\ell}{N_h},\frac{2^{\ell+1}}{N_h}\right]\times[0,\alpha) : (3.18) \text{ holds}\right\}.$$

Then the above argument leads to the inequality

$$\lambda_2(B(N_h;M_h)) \geqslant \frac{N_h}{2M_h}\lambda_2(A(N_h;M_h)). \tag{3.20}$$

Combining (3.13) and (3.20), we obtain the upper bound

$$\lambda_2(A(N_h;M_h)) \leqslant \frac{2^{\ell+13}}{hN_h^2\log^2(1+h)}.$$

Combining this with (3.5), it is not difficult to see that apart from a set of measure $\lambda_2(A(N_h;M_h))$, every pair (α,β) such that (3.14) holds satisfies the inequality

$$\left|\sum_{k=0}^{K(\alpha)-1}\chi_{S(h)}(k\alpha+\beta)-K(\alpha)\lambda_1(S(h))\right|<\frac{2N_h}{M_h}+(hN_hM_h)^{1/2}\log(1+h).$$

Next, note that the two expressions

$$\sum_{\substack{k\geqslant 0\\ k\alpha+\beta\in S(h)}}1-\frac{1}{\alpha}\lambda_1(S(h))\quad\text{and}\quad\sum_{k=0}^{K(\alpha)-1}\chi_{S(h)}(k\alpha+\beta)-K(\alpha)\lambda_1(S(h))$$

differ by at most 2, due to the possibility that $(K(\alpha)-1)\alpha+\beta>1$ and the difference $|K(\alpha)-1/\alpha|<1$, in view of (3.4). It follows that on reversing the contraction, we see that apart from a set of measure at most

$$\frac{2^{\ell+13}}{h\log^2(1+h)},$$

every pair (θ,η) such that $2^\ell\leqslant\theta\leqslant 2^{\ell+1}$ and $\eta\in[0,\theta)$ satisfies the inequality

$$\left|\sum_{\substack{k\geqslant 0\\ k\theta+\eta\in\mathcal{S}\cap[0,N_h]}}1-\frac{1}{\theta}\lambda_1(\mathcal{S}\cap[0,N_h])\right|<\frac{2N_h}{M_h}+(hN_hM_h)^{1/2}\log(1+h)+2. \qquad (3.21)$$

Lemma 3.3 (Borel–Cantelli lemma)**.** *Let (X,Σ,μ) be a measure space, and suppose that E_h, $h=1,2,3,\ldots$, is a sequence of Σ-measurable sets. If*

$$\sum_{h=1}^{\infty}\mu(E_h)<\infty,$$

then

$$\mu\left(\bigcap_{h=1}^{\infty}\bigcup_{i=h}^{\infty}E_i\right)=0.$$

Since

$$\sum_{h=1}^{\infty}\frac{1}{h\log^2(1+h)}<\infty,$$

we conclude that for almost every pair (θ,η) such that $2^\ell\leqslant\theta\leqslant 2^{\ell+1}$ and $\eta\in[0,\theta)$, the inequality (3.21) holds for all sufficiently large positive integers h.

Finally, we specify the integers N_h and M_h in terms of the parameter $h\geqslant 1$ and the chosen integer ℓ. Choosing them to satisfy

$$N_h\leqslant 2^\ell(h^4\log^2(1+h))<N_h+1\quad\text{and}\quad M_h=2^\ell h \qquad (3.22)$$

ensures that the two dominant terms on the right hand side of (3.21) have the same order of magnitude in terms of h. For an arbitrary sufficiently integer n, we choose

h to satisfy $N_h \leqslant n < N_{h+1}$. Then it follows from (3.21) and (3.22) that

$$\begin{aligned}
&\left| \sum_{\substack{k\geqslant 0\\ k\theta+\eta\in\mathcal{S}\cap[0,n]}} 1 - \frac{1}{\theta}\lambda_1(\mathcal{S}\cap[0,n]) \right| \\
&\quad < \frac{N_{h+1}-N_h}{\theta} + \frac{2N_h}{M_h} + (hN_hM_h)^{1/2}\log(1+h) + 2 \\
&\quad \leqslant \frac{2^\ell((h+1)^4\log^2(2+h) - h^4\log^2(1+h))}{\theta} + (2+2^\ell)h^3\log^2(1+h) + 3 \\
&\quad \leqslant 2^{\ell+3}h^3\log^2(1+h) + O_\ell(h^3\log(1+h)) \leqslant c_1(\ell)n^{3/4}(\log n)^{1/2},
\end{aligned}$$

provided that n, and hence also h, is sufficiently large.

This completes the proof of Lemma 2.1.

Proof of Lemma 3.1. For any fixed $\delta \in (0,1/2)$, we define the roof function $R_\delta : \mathbb{R} \to \mathbb{R}$ by writing

$$R_\delta(x) = \begin{cases} 0, & \text{if } |x| > \delta, \\ 1 - (|x|/\delta), & \text{if } 0 \leqslant |x| \leqslant \delta. \end{cases}$$

For every integer $j \in \mathbb{Z}$, we consider the integral

$$I(\delta; j) = \int_{-1/2}^{1/2} R_\delta(x)\mathrm{e}^{2\pi\mathrm{i}jx}\,\mathrm{d}x. \tag{3.23}$$

Then

$$I(\delta;0) = \delta \quad \text{and} \quad I(\delta;j) = \delta\left(\frac{\sin\pi j\delta}{\pi j\delta}\right)^2, \quad j \in \mathbb{Z}\setminus\{0\}. \tag{3.24}$$

For any integers $j_1, j_2 \in \mathbb{Z}\setminus\{0\}$ and positive integer N, let

$$\begin{aligned}
\mathfrak{B}(j_1;j_2;N) &= \int_{2^\ell/N}^{2^{\ell+1}/N} \left(a_{j_1} \sum_{k_1=0}^{K(\alpha)-1} \mathrm{e}^{2\pi\mathrm{i}j_1k_1\alpha} \right) \left(\overline{a_{j_2}} \sum_{k_2=0}^{K(\alpha)-1} \mathrm{e}^{-2\pi\mathrm{i}j_2k_2\alpha} \right) \mathrm{d}\alpha \\
&= \int_{2^\ell/N}^{2^{\ell+1}/N} \left(a_{j_1} \frac{\mathrm{e}^{2\pi\mathrm{i}j_1K(\alpha)\alpha}-1}{\mathrm{e}^{2\pi\mathrm{i}j_1\alpha}-1} \right) \left(\overline{a_{j_2}} \frac{\mathrm{e}^{-2\pi\mathrm{i}j_2K(\alpha)\alpha}-1}{\mathrm{e}^{-2\pi\mathrm{i}j_2\alpha}-1} \right) \mathrm{d}\alpha,
\end{aligned} \tag{3.25}$$

so that

$$\begin{aligned}
&|\mathfrak{B}(j_1;j_2;N)| \\
&\quad \leqslant \frac{1}{2}\int_{2^\ell/N}^{2^{\ell+1}/N} \left(\left| a_{j_1} \frac{\mathrm{e}^{2\pi\mathrm{i}j_1K(\alpha)\alpha}-1}{\mathrm{e}^{2\pi\mathrm{i}j_1\alpha}-1} \right|^2 + \left| a_{j_2} \frac{\mathrm{e}^{2\pi\mathrm{i}j_2K(\alpha)\alpha}-1}{\mathrm{e}^{2\pi\mathrm{i}j_2\alpha}-1} \right|^2 \right) \mathrm{d}\alpha.
\end{aligned} \tag{3.26}$$

Then it follows from (3.8), (3.23) with $\delta = 2^\ell/M$ and $j = j_1 - j_2$ and from (3.25) that

$$J^*(\mathbf{a};N;M) = \sum_{j_1\in\mathbb{Z}\setminus\{0\}} \sum_{j_2\in\mathbb{Z}\setminus\{0\}} I\left(\frac{2^\ell}{M}; j_1 - j_2\right) \mathfrak{B}(j_1;j_2;N).$$

We write

$$J^*(\mathbf{a};N;M) = J_1^*(\mathbf{a};N;M) + J_2^*(\mathbf{a};N;M), \tag{3.27}$$

where $J_1^*(\mathbf{a};N;M)$ contains all the diagonal terms in $J^*(\mathbf{a};N;M)$ with $j_1=j_2$, while $J_2^*(\mathbf{a};N;M)$ contains all the off-diagonal terms in $J^*(\mathbf{a};N;M)$ with $j_1\neq j_2$. Noting that $I(2^\ell/M;0)=2^\ell/M$, we see that

$$J_1^*(\mathbf{a};N;M)=\frac{2^\ell}{M}\sum_{j\in\mathbb{Z}\setminus\{0\}}|a_j|^2E(j;N), \tag{3.28}$$

where

$$E(j;N)=\int_{2^\ell/N}^{2^{\ell+1}/N}\left|\frac{\mathrm{e}^{2\pi\mathrm{i}jK(\alpha)\alpha}-1}{\mathrm{e}^{2\pi\mathrm{i}j\alpha}-1}\right|^2\mathrm{d}\alpha. \tag{3.29}$$

Meanwhile, noting (3.24), we see that

$$J_2^*(\mathbf{a};N;M)=\frac{2^\ell}{M}\sum_{j_1\in\mathbb{Z}\setminus\{0\}}\sum_{\substack{j_2\in\mathbb{Z}\setminus\{0\}\\ j_1\neq j_2}}\left(\frac{\sin\pi(j_1-j_2)2^\ell M^{-1}}{\pi(j_1-j_2)2^\ell M^{-1}}\right)^2\mathfrak{B}(j_1;j_2;N). \tag{3.30}$$

Combining (3.26) and (3.30), we deduce that

$$|J_2^*(\mathbf{a};N;M)|\leqslant\frac{2^\ell}{M}\sum_{j\in\mathbb{Z}\setminus\{0\}}\sum_{\zeta\in\mathbb{Z}\setminus\{0\}}\left(\frac{\sin\pi\zeta 2^\ell M^{-1}}{\pi\zeta 2^\ell M^{-1}}\right)^2|a_j|^2E(j;N). \tag{3.31}$$

It then follows from (3.27), (3.28) and (3.31) that

$$|J^*(\mathbf{a};N;M)|\leqslant\frac{2^\ell}{M}\left(1+\sum_{\zeta\in\mathbb{Z}\setminus\{0\}}\left(\frac{\sin\pi\zeta 2^\ell M^{-1}}{\pi\zeta 2^\ell M^{-1}}\right)^2\right)\sum_{j\in\mathbb{Z}\setminus\{0\}}|a_j|^2E(j;N). \tag{3.32}$$

Next, note that

$$\begin{aligned}\frac{2^\ell}{M}\left(1+\sum_{\zeta\in\mathbb{Z}\setminus\{0\}}\left(\frac{\sin\pi\zeta 2^\ell M^{-1}}{\pi\zeta 2^\ell M^{-1}}\right)^2\right)&\leqslant\frac{2^\ell}{M}\left(\sum_{|\zeta|\leqslant M/2^\ell}1+\sum_{|\zeta|>M/2^\ell}\left(\frac{M}{\pi\zeta 2^\ell}\right)^2\right)\\&\leqslant\frac{2^\ell}{M}\left(\frac{2M}{2^\ell}+1+\frac{2M^2}{\pi^2 4^\ell}\int_{M/2^\ell}^\infty\frac{\mathrm{d}x}{x^2}\right)\leqslant 6.\end{aligned} \tag{3.33}$$

To complete the proof of Lemma 3.1, in view of (3.3), (3.32) and (3.33), it suffices to show that for every $j\in\mathbb{Z}\setminus\{0\}$ and integer $N>1$, we have

$$E(j;N)\leqslant 2^{\ell+8}. \tag{3.34}$$

Consider first small values of j, where $1\leqslant|j|\leqslant N/2^{\ell+2}$. Suppose first that $j>0$. Since $\alpha\in[2^\ell/N,2^{\ell+1}/N)$, it follows that $0\leqslant j\alpha\leqslant 1/2$, so that $\mathrm{e}^{2\pi\mathrm{i}j\alpha}$ is a point on the upper half circle of unit radius. On the other hand, (3.4) implies the inequality $0\leqslant jK(\alpha)\alpha-j<j\alpha$. Hence $\mathrm{e}^{2\pi\mathrm{i}jK(\alpha)\alpha}=\mathrm{e}^{2\pi\mathrm{i}(jK(\alpha)\alpha-j)}$ is a point on the circular arc of the upper half circle of unit radius joining the points 1 and $\mathrm{e}^{2\pi\mathrm{i}j\alpha}$. As shown in Figure 4.1, we clearly have $|\mathrm{e}^{2\pi\mathrm{i}jK(\alpha)\alpha}-1|<|\mathrm{e}^{2\pi\mathrm{i}j\alpha}-1|$.

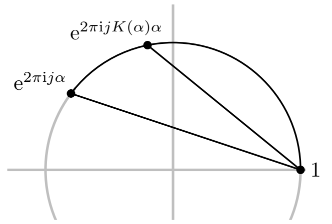

Figure 2: justifying the inequality (3.35)

Meanwhile, replacing j by $-j$ preserves this inequality. It follows that for every integer j satisfying $1 \leqslant |j| \leqslant N/2^{\ell+2}$, we have

$$\left| \frac{\mathrm{e}^{2\pi \mathrm{i} j K(\alpha)\alpha} - 1}{\mathrm{e}^{2\pi \mathrm{i} j\alpha} - 1} \right| < 1 \quad \text{and so} \quad E(j;N) \leqslant \frac{2^\ell}{N}. \tag{3.35}$$

Next, for any fixed integer j satisfying $|j| > N/2^{\ell+2}$, we use the inequalities

$$\left| \frac{\mathrm{e}^{2\pi \mathrm{i} j K(\alpha)\alpha} - 1}{\mathrm{e}^{2\pi \mathrm{i} j\alpha} - 1} \right| \leqslant \min\left\{K(\alpha), \frac{1}{\|j\alpha\|}\right\} \leqslant \min\left\{N, \frac{1}{\|j\alpha\|}\right\}, \tag{3.36}$$

where $\|x\|$ denotes the distance of a real number x from the nearest integer, and the inequality $K(\alpha) \leqslant N$ follows from $\alpha \in [2^\ell/N, 2^{\ell+1}/N)$ and (3.4).

Let r be the integer closest to $j\alpha$. Then

$$\|j\alpha\| < \frac{1}{N} \quad \text{if and only if} \quad \left|\alpha - \frac{r}{j}\right| < \frac{1}{|j|N},$$

so that

$$\|j\alpha\| < \frac{1}{N} \quad \text{if and only if} \quad \alpha \in \left(\frac{r}{j} - \frac{1}{|j|N}, \frac{r}{j} + \frac{1}{|j|N}\right). \tag{3.37}$$

For this fixed integer j satisfying $|j| > N/2^{\ell+2}$, there are at most

$$\max\left\{5, \frac{2^{\ell+2}|j|}{N}\right\} \leqslant \frac{2^{\ell+5}|j|}{N} \tag{3.38}$$

integers r such that

$$\left(\frac{r}{j} - \frac{1}{|j|N}, \frac{r}{j} + \frac{1}{|j|N}\right) \cap \left[\frac{2^\ell}{N}, \frac{2^{\ell+1}}{N}\right) \neq \emptyset. \tag{3.39}$$

On the other hand, suppose that n is a positive fixed integer satisfying $2^n \leqslant N$. Then analogous to (3.37), we have

$$\frac{2^{n-1}}{N} \leqslant \|j\alpha\| \leqslant \frac{2^n}{N} \quad \text{if and only if} \quad \alpha \in \mathfrak{I}(j;N;n), \tag{3.40}$$

where

$$\mathfrak{I}(j;N;n) = \left[\frac{r}{j} - \frac{2^n}{|j|N}, \frac{r}{j} - \frac{2^{n-1}}{|j|N}\right] \cup \left[\frac{r}{j} + \frac{2^{n-1}}{|j|N}, \frac{r}{j} + \frac{2^n}{|j|N}\right], \tag{3.41}$$

except when $2^n/N > 1/2$, in which case we have the modification

$$\mathfrak{I}(j;N;n) = \left[\frac{r}{j} - \frac{1}{2|j|}, \frac{r}{j} - \frac{2^{n-1}}{|j|N}\right] \cup \left[\frac{r}{j} + \frac{2^{n-1}}{|j|N}, \frac{r}{j} + \frac{1}{2|j|}\right]. \tag{3.42}$$

Similarly, for this fixed integer j satisfying $|j| > N/2^{\ell+2}$ and fixed positive integer n satisfying $2^n \leqslant N$, there are at most (3.38) integers r such that

$$\mathfrak{I}(j;N;n) \cap \left[\frac{2^\ell}{N}, \frac{2^{\ell+1}}{N}\right) \neq \emptyset. \tag{3.43}$$

Combining (3.29) and (3.36)–(3.43), we see that for any fixed integer j satisfying $|j| > N/2^{\ell+2}$, we have

$$\begin{aligned} E(j;N) &\leqslant \int_{2^\ell/N}^{2^{\ell+1}/N} \left(\min\left\{N, \frac{1}{\|j\alpha\|}\right\}\right)^2 \mathrm{d}\alpha \\ &\leqslant N^2\,\frac{2^{\ell+5}|j|}{N}\,\frac{2}{|j|N} + \sum_{\substack{n=1\\2^n\leqslant N}}^{\infty} \left(\frac{N}{2^{n-1}}\right)^2 \frac{2^{\ell+5}|j|}{N}\,\frac{2^n}{|j|N} \leqslant 2^{\ell+5}\left(2+4\sum_{n=1}^{\infty}\frac{1}{2^n}\right), \end{aligned}$$

confirming the assertion (3.34) and completing the proof. □

References

[1] J. Beck, M. Donders, Y. Yang. Quantitative behavior of non-integrable systems I. *Acta Math. Hungar.* **161** (2020), 66–184.

[2] J. Beck, M. Donders, Y. Yang. Quantitative behavior of non-integrable systems II. *Acta Math. Hungar.* **162** (2020), 220–324.

[3] J. Beck, W.W.L. Chen, Y. Yang. Quantitative behavior of non-integrable systems IV. *Acta Math. Hungar.* **167** (2022), 1–160.

[4] S. Kerckhoff, H. Masur, J. Smillie. Ergodicity of billiard flows and quadratic differentials. *Ann. of Math.* **124** (1986), 293–311.

[5] D. König, A. Szücs. Mouvement d'un point abondonne a l'interieur d'un cube. *Rend. Circ. Mat. Palermo* **36** (1913), 79–90.

Department of Mathematics, Hill Center for the Mathematical Sciences, Rutgers University, Piscataway NJ 08854, USA

Email address: jbeck@math.rutgers.edu

School of Mathematical and Physical Sciences, Faculty of Science and Engineering, Macquarie University, Sydney NSW 2109, Australia

Email address: william.chen@mq.edu.au

School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China

Email address: yangyx@bupt.edu.cn