

Strong mixing of the Boca–Cobeli–Zaharescu map

Abstract

We prove that the Boca–Cobeli–Zaharescu map is strongly mixing for normalized Lebesgue measure on the Farey triangle. This answers the mixing question of Athreya and Cheung (Int. Math. Res. Not., 2014), which remained open after Cheung and Quas (arXiv, 2024) established weak mixing. Our main tool is a criterion for weakly mixing transformations whose induced maps on sets of measure tending to one are conjugate to the original transformation by maps converging to the identity. We decompose subsequential limits of graph measures according to the number of extra ordinary iterates needed to complete a fixed number of induced returns. Under a bounded mean-excess hypothesis at each scale, a finite invariant measure built from the positive-excess components forces each such component to be a multiple of the product measure; only the zero-excess component can obstruct mixing. For the BCZ map, we make this remaining mass arbitrarily small by estimating the probability that a long thin rectangle contains no primitive lattice point. The estimate combines a finite prime sieve, exact pairwise independence on an auxiliary torus, and a summable bound for the remaining prime divisors.

Note. This paper was generated entirely by AI, including MiMo, using an automated research pipeline developed by Chenghua Liu and Hanyu Li.

1 Introduction

The Boca–Cobeli–Zaharescu (BCZ) map was introduced to study correlations of Farey gaps [4]. Its iterates record consecutive normalized denominator pairs: if q_j, q_{j+1}, q_{j+2} are consecutive denominators in the Farey sequence of order Q , the map sends $(q_j/Q, q_{j+1}/Q)$ to $(q_{j+1}/Q, q_{j+2}/Q)$. Thus a piecewise linear map of a triangle encodes the progression through a Farey sequence. The joint spacing distributions of Augustin, Boca, Cobeli, and Zaharescu [3], and the correlation measures of Boca and Zaharescu [5], describe arithmetic statistics as the Farey order tends to infinity. The dynamical mixing question asks whether, in the resulting invariant system, observations separated by many iterates become asymptotically independent.

Athreya and Cheung [2] identified the BCZ map as a Poincaré return map for the horocycle flow on unimodular lattices. They proved that normalized Lebesgue measure on the Farey triangle is invariant and ergodic, and that the map has zero entropy. They also explicitly asked whether the map is mixing [2, §8.3]. Although the horocycle flow is mixing, this property does not in general pass to a return map: separation by a prescribed number of returns is different from separation by a prescribed flow time.

Cheung and Quas [6] proved weak mixing using the self-inducing geometry of the section. Restricting to lattices with a shorter horizontal vector gives an induced map conjugate to the original map, and the conjugacy approaches the identity as the restriction disappears. Their proof finds a positive proportion of points for which the induced and ordinary return counts differ by exactly one, and uses this event to exclude nonconstant eigenfunctions. Artiles [1] subsequently extended weak mixing to the corresponding return maps for lattice translation surfaces. Here we prove strong mixing for the BCZ map.

Theorem 1 (Strong mixing). *Let*

$$\Omega = \{(a, b) \in (0, 1]^2 : a + b > 1\}, \quad dm = 2 \, da \, db,$$

and let $B : \Omega \rightarrow \Omega$ be the Boca–Cobeli–Zaharescu map

$$B(a, b) = \left(b, \left\lfloor \frac{1+a}{b} \right\rfloor b - a \right).$$

Then, for all measurable sets $A, C \subset \Omega$,

$$\lim_{n \rightarrow \infty} m(B^{-n}A \cap C) = m(A)m(C).$$

Equivalently, for $f, g \in L^2(m)$,

$$\lim_{n \rightarrow \infty} \int_{\Omega} f(x)g(B^n x) \, dm(x) = \left(\int_{\Omega} f \, dm \right) \left(\int_{\Omega} g \, dm \right).$$

In particular, two prescribed measurable conditions on finite blocks of the stationary BCZ process become asymptotically independent as their separation grows. The convergence holds along all integer times, not merely after averaging over time. We do not obtain a rate of mixing.

The step beyond weak mixing is to use all positive excess return counts, rather than only the event of one extra return. Consider a subsequential limit ν of the distributions of $(x, B^n x)$, with x distributed according to m . Near-identity self-induction gives two decompositions of this same limit. In one decomposition a component corresponds to an excess of j ordinary iterates; in the other, that component is shifted by j iterates in its second coordinate. Summing the intermediate shifts over $j > 0$ produces a nonnegative measure of finite mass, with finiteness supplied by Kac’s return-time formula [7]. The equality of the decompositions makes this measure invariant under $\text{id} \times B$. Ergodicity identifies it as a multiple of $m \times m$, and weak mixing then does the same for each positive-excess component. Any nonproduct part of ν must therefore come from zero excess. Theorem 4 formulates this argument for general induced maps.

For the BCZ map, an extra ordinary return is recorded by a primitive lattice vector in the strip removed from the smaller section. We show that the probability of finding no such vector can be made arbitrarily small. The required estimate concerns rectangles of width α/n and height $2n$, with $n \rightarrow \infty$ before $\alpha \rightarrow \infty$. A finite prime sieve makes the relevant counting variables pairwise independent on an auxiliary torus. Comparing this auxiliary measure with the section measure and controlling the remaining prime divisors gives the estimate in Lemma 5. Thus the proof combines a criterion for controlling every subsequential correlation limit with an arithmetic argument that eliminates its zero-excess component. The section geometry and weak mixing remain established inputs; the latter is also proved in Appendix A, following Cheung–Quas.

2 Self-induction of the horocycle section

For $x = (a, b) \in \Omega$, write

$$\Lambda_x = \mathbb{Z}(a, 0) + \mathbb{Z}(b, a^{-1}), \quad h_u = \begin{pmatrix} 1 & 0 \\ -u & 1 \end{pmatrix}.$$

A nonzero vector of Λ_x is *primitive* if it is not an integer multiple, by a factor of absolute value greater than one, of another lattice vector. In the displayed basis, this is equivalent to coprimality

of its two integer coefficients. The matrix h_u preserves the first coordinate of a vector and decreases its slope by u . The Farey triangle parametrizes the section of lattices having a primitive horizontal vector with positive length at most 1.

We recall the return description from [2]. A lattice vector with positive second coordinate has the form

$$v = (-pa + qb, q/a), \quad p \in \mathbb{Z}, \quad q \in \mathbb{N} = \{1, 2, \dots\}.$$

If its first coordinate belongs to $(0, 1]$, then $p \geq 0$: otherwise $-pa + qb \geq a + b > 1$. Its slope is consequently at least $1/(ab)$, with equality first attained by the primitive vector (b, a^{-1}) . At time $1/(ab)$ this vector becomes horizontal. If $k = \lfloor (1 + a)/b \rfloor$, changing the basis to

$$h_{1/(ab)}(b, a^{-1}) = (b, 0), \quad h_{1/(ab)}(-(a, 0) + k(b, a^{-1})) = (kb - a, b^{-1})$$

identifies the returned lattice with $\Lambda_{B(a,b)}$. Thus the return time is

$$\tau(a, b) = \frac{1}{ab}, \quad \int_{\Omega} \tau \, dm = 2 \int_0^1 \frac{-\log(1-a)}{a} \, da = \frac{\pi^2}{3}. \quad (1)$$

More generally, future returns correspond, in increasing order of slope, to primitive vectors in Λ_x with first coordinate in $(0, 1]$ and second coordinate positive. Distinct such primitive vectors cannot have the same slope.

The map B is invertible modulo null sets, with inverse

$$B^{-1}(a, b) = \left(\left\lfloor \frac{1+b}{a} \right\rfloor a - b, a \right).$$

Its linear branches have determinant one, so it preserves m . Ergodicity follows from the section construction and ergodicity of the horocycle flow, as proved in [2, Theorem 1.2]. All maps below are understood modulo null sets. Extending m by zero to $\bar{\Omega}$ and choosing a measurable extension of B to the added boundary gives a compact metric probability-space model.

For $0 < r < 1$, let

$$Y_r = \{(a, b) \in \Omega : a \leq r\}, \quad m(Y_r) = r^2, \quad m_r = r^{-2}m|_{Y_r}.$$

Write B_r for the first-return map of B to Y_r , and $\mathcal{R}_{r,n}(y)$ for the number of ordinary iterates until the n th return to Y_r . Thus

$$B_r^n y = B^{\mathcal{R}_{r,n}(y)} y, \quad \mathcal{R}_{r,n}(y) \geq n.$$

The following form of the self-induction in [2, 6] also records the normalization of measure and flow time.

Lemma 2 (Self-induction and time scaling). *Define*

$$\phi_r(a, b) = (ra, r(b + ja)), \quad j = \left\lfloor \frac{r^{-1} - b}{a} \right\rfloor. \quad (2)$$

Then $\phi_r : \Omega \rightarrow Y_r$ is an almost-everywhere bijection satisfying

$$(\phi_r)_* m = m_r, \quad B_r \phi_r = \phi_r B.$$

Moreover, $\phi_r \rightarrow \text{id}$ in m -measure as $r \uparrow 1$.

Let $H_n(x) = \sum_{i=0}^{n-1} \tau(B^i x)$, and let $H_{r,n}(y)$ be the horocycle flow time until the n th return to the smaller section. Then

$$H_{r,n}(\phi_r x) = r^{-2} H_n(x). \quad (3)$$

Proof. Let $g_r = \text{diag}(r, r^{-1})$. It maps the section with horizontal cutoff 1 onto the section with horizontal cutoff r . In the first row above the horizontal axis of $g_r \Lambda_x$, choose the rightmost vector whose first coordinate is at most 1. Its first coordinate is $r(b + ja)$ with j as in (2); it belongs to $(1 - ra, 1]$. This gives the stated coordinates.

For $(s, t) \in Y_r$, put

$$j = \left\lceil \frac{t - r}{s} \right\rceil, \quad (a, b) = \left(\frac{s}{r}, \frac{t - js}{r} \right).$$

The inequalities $r - s < t - js \leq r$ show that $(a, b) \in \Omega$. Also $1 - s < t \leq 1$, so $\lfloor (r^{-1} - b)/a \rfloor = j$. This proves the inverse formula. On each piece of (2), the Jacobian is r^2 ; hence $(\phi_r)_* m = r^{-2} m|_{Y_r} = m_r$.

On $\{a + b > r^{-1}\}$, the integer j is zero and $\phi_r(a, b) = r(a, b)$. The measure of this set tends to one, proving convergence in measure. Finally,

$$h_u g_r = g_r h_{r^2 u}.$$

Thus g_r preserves the order of section returns and multiplies their flow times by r^{-2} . This proves both the conjugacy and (3). $\square$

An important consequence is that Birkhoff's theorem can be applied without a uniform ergodic theorem for the changing sections. For any sequence $r_n \in (0, 1)$ with $r_n \rightarrow 1$, equations (1) and (3) give

$$\frac{H_{r_n, n}}{n} \longrightarrow \frac{\pi^2}{3} \quad \text{in } m_{r_n}\text{-probability.} \quad (4)$$

Indeed, its distribution is that of $r_n^{-2} H_n / n$ under the fixed measure m , and $H_n / n \rightarrow \pi^2 / 3$ almost everywhere by ergodicity.

Proposition 3 (Cheung–Quas [6]). *The BCZ map is weakly mixing with respect to m . Equivalently, $B \times B$ is ergodic with respect to $m \times m$.*

For completeness, Appendix A gives their counting and eigenfunction argument with the return-time conventions used here.

3 A mixing criterion from excess returns

For an invertible ergodic probability-preserving transformation T and a measurable set Y of positive measure, write

$$r_Y(y) = \min\{k \geq 1 : T^k y \in Y\}, \quad T_Y(y) = T^{r_Y(y)} y,$$

and

$$R_Y^{(n)}(y) = \sum_{i=0}^{n-1} r_Y(T_Y^i y).$$

Thus $R_Y^{(n)}$ is the discrete time of the n th induced return, and $R_Y^{(n)} - n \geq 0$ counts the extra ordinary iterates. These maps and return times are defined almost everywhere on Y .

Theorem 4 (Mixing from excess returns). *Let T be an invertible, probability-preserving, weakly mixing transformation of a compact metric space X with Borel probability measure μ . For each $\alpha > 0$ and all sufficiently large integers n , let $Y_{n, \alpha}$ be a measurable set of measure $\rho_{n, \alpha} > 0$. Denote its induced map by $T_{n, \alpha}$, its normalized measure by $\mu_{n, \alpha} = \rho_{n, \alpha}^{-1} \mu|_{Y_{n, \alpha}}$, and its n th return time by $R_{n, \alpha}$. Suppose that:*

(i) There are almost-everywhere bijections $\phi_{n,\alpha} : X \rightarrow Y_{n,\alpha}$ such that

$$(\phi_{n,\alpha})_*\mu = \mu_{n,\alpha}, \quad T_{n,\alpha}\phi_{n,\alpha} = \phi_{n,\alpha}T,$$

and, for each fixed α , $\phi_{n,\alpha} \rightarrow \text{id}$ in μ -measure.

(ii) For each fixed α , $\rho_{n,\alpha} \rightarrow 1$ and

$$\limsup_{n \rightarrow \infty} n(\rho_{n,\alpha}^{-1} - 1) < \infty.$$

(iii) The probability of zero excess satisfies

$$\lim_{\alpha \rightarrow \infty} \limsup_{n \rightarrow \infty} \mu_{n,\alpha}(R_{n,\alpha} = n) = 0. \quad (5)$$

Then T is strongly mixing.

Proof. Fix α and suppress it in the subscripts. Set

$$D_n = R_n - n, \quad E_{n,j} = \{y \in Y_n : D_n(y) = j\}, \quad j = 0, 1, \dots$$

Kac's formula [7] gives $\int_{Y_n} r_{Y_n} d\mu = 1$. Since the induced map preserves μ_n ,

$$\int_{Y_n} D_n d\mu_n = n(\rho_n^{-1} - 1). \quad (6)$$

By (ii), after discarding finitely many n these first moments are bounded by a constant $C_\alpha < \infty$. In particular,

$$\mu_n(D_n > J) \leq \frac{C_\alpha}{J+1} \quad (J \geq 0). \quad (7)$$

We first check that the sets $E_{n,j}$ are almost invariant under T . For any induced map T_Y , if

$$x \in Y, \quad Tx \in Y, \quad TT_Y^n x \in Y,$$

then the first and the $(n+1)$ st induced return each take one ordinary iterate. Consequently,

$$R_Y^{(n)}(Tx) = R_Y^{(n+1)}(x) - 1 = R_Y^{(n)}(x).$$

The complement of these three conditions has μ -measure at most $3\mu(X \setminus Y)$. For the third condition, restricted to Y , this follows from preservation of $\mu|_Y$ by T_Y^n and $\mu(Y \cap T^{-1}(X \setminus Y)) \leq \mu(X \setminus Y)$. It follows that, for every j ,

$$\mu(E_{n,j} \triangle T^{-1}E_{n,j}) \leq 3(1 - \rho_n) \longrightarrow 0. \quad (8)$$

Let

$$\nu_n = (\text{id}, T^n)_*\mu, \quad \Pi = \mu \times \mu,$$

and fix an arbitrary weakly convergent subsequence $\nu_{n_k} \Rightarrow \nu$. Compactness permits passage to a further subsequence on which, simultaneously for all $j \geq 0$,

$$\nu_{n,j} := (\text{id}, T^n)_*(\mu_n|_{E_{n,j}}) \Rightarrow \nu_j.$$

We continue to write n for this subsequence, and put $p_j = \nu_j(X \times X)$.

Both marginals of $\nu_{n,j}$ are bounded above by $\rho_n^{-1}\mu$, so both marginals of ν_j are bounded above by μ . This domination also allows us to test weak convergence against fixed products of $L^2(\mu)$ functions. Indeed, if the marginals of a measure η are bounded above by $C\mu$, Cauchy–Schwarz gives

$$\left| \int (f(x)g(y) - f_0(x)g_0(y)) d\eta \right| \leq C(\|f - f_0\|_2 \|g\|_2 + \|f_0\|_2 \|g - g_0\|_2).$$

Approximating f, g by continuous f_0, g_0 proves the assertion. In particular, pushforwards by the measurable maps $T \times T$ and $\text{id} \times T^j$, for fixed j , pass through these limits.

Put $V = T \times T$ and $W = \text{id} \times T$. Equation (8) implies that $V_*\nu_{n,j} - \nu_{n,j}$ tends to zero in total variation, and hence $V_*\nu_j = \nu_j$. Each marginal of ν_j is therefore a T -invariant measure dominated by μ . Ergodicity of T makes its Radon–Nikodym density constant, so both marginals equal $p_j\mu$.

There are two ways to recover the limiting graph measure from these components. First,

$$\sum_{j \geq 0} \nu_{n,j} = (\text{id}, T^n)_* \mu_n,$$

which differs from ν_n by a measure whose total variation tends to zero. Second, the return identity and the conjugacy give

$$\sum_{j \geq 0} W_*^j \nu_{n,j} = (\text{id}, T_n^n)_* \mu_n = (\phi_n \times \phi_n)_* \nu_n.$$

The last expression also converges weakly to ν . To see this, let F be continuous on $X \times X$. Uniform continuity and $\phi_n \rightarrow \text{id}$ in measure imply

$$\int (F(\phi_n x, \phi_n y) - F(x, y)) d\nu_n(x, y) \longrightarrow 0,$$

since both marginals of ν_n equal μ . The tail bound (7) now permits passage to the limit in both sums. We obtain

$$\nu = \sum_{j \geq 0} \nu_j = \sum_{j \geq 0} W_*^j \nu_j, \quad \sum_{j \geq 0} p_j = 1, \quad \sum_{j \geq 1} j p_j \leq C_\alpha. \quad (9)$$

The last inequality follows by taking limits in finite partial sums of the first moment and then increasing the number of terms.

The nonnegativity of the excess counts allows us to turn these two identities into a finite invariant measure. Define

$$\mathcal{F} = \sum_{j \geq 1} \sum_{k=0}^{j-1} W_*^k \nu_j, \quad \gamma = \mathcal{F}(X \times X) = \sum_{j \geq 1} j p_j < \infty. \quad (10)$$

Both marginals of $\mathcal{F}$ equal $\gamma\mu$. Since V and W commute, $\mathcal{F}$ is V -invariant. Moreover, telescoping gives

$$W_* \mathcal{F} - \mathcal{F} = \sum_{j \geq 1} (W_*^j \nu_j - \nu_j) = 0$$

by (9). All these series converge in total variation because $\mathcal{F}$ has finite mass.

We claim that $\mathcal{F} = \gamma\Pi$. For bounded f, g , its W -invariance gives, for every $N \geq 1$,

$$\int f(x)g(y) d\mathcal{F} = \int f(x) \left(\frac{1}{N} \sum_{k=0}^{N-1} g(T^k y) \right) d\mathcal{F}.$$

By the mean ergodic theorem the average in parentheses converges in $L^1(\mu)$ to $\int g d\mu$. Its integration error against $f(x) d\mathcal{F}$ is at most $\gamma\|f\|_\infty$ times this L^1 error. The first marginal of $\mathcal{F}$ is $\gamma\mu$, so the claimed product identity follows.

For $j \geq 1$, the term with $k = 0$ in (10) shows that $\nu_j \leq \mathcal{F} = \gamma\Pi$. The measure ν_j is V -invariant, and weak mixing says that V is ergodic for Π . Its density with respect to Π must therefore be constant. Thus

$$\nu_j = p_j\Pi \quad (j \geq 1), \quad \nu = \nu_0 + (1 - p_0)\Pi. \quad (11)$$

This also covers $\gamma = 0$, when all the positive-excess components vanish.

Finally, let

$$v(\alpha) = \limsup_{n \rightarrow \infty} \mu_{n,\alpha}(R_{n,\alpha} = n).$$

For every fixed α , the preceding construction yields $p_0 \leq v(\alpha)$ and hence

$$\nu \geq (1 - v(\alpha))\Pi.$$

The further subsequence and its components may depend on α , but the initially fixed limit ν does not. By (iii), $\nu \geq \Pi$; equality follows because both measures have mass one. Every subsequential limit of ν_n is therefore Π , proving $\nu_n \Rightarrow \Pi$. The same continuous approximation argument, using the fixed marginals μ , gives convergence for all L^2 product tests and in particular for indicators of measurable sets. This is strong mixing. $\square$

4 Primitive lattice points and the mixing theorem

A primitive vector whose horizontal coordinate belongs to $(r, 1]$ records a return to the original section that is omitted from the smaller one. We estimate the probability that no such vector occurs within a height proportional to the number of returns.

Lemma 5 (Primitive-point void probability). *For $\alpha > 0$ and integers $n > \alpha$, set $\varepsilon = \alpha/n$ and*

$$Z_{n,\alpha}(x) = \# \left\{ v \in \Lambda_x : \begin{array}{l} v \text{ is primitive,} \\ v \in (1 - \varepsilon, 1] \times (0, 2n] \end{array} \right\}.$$

Then

$$\lim_{\alpha \rightarrow \infty} \limsup_{n \rightarrow \infty} m(Z_{n,\alpha} = 0) = 0. \quad (12)$$

Proof. Use the coordinates

$$c = a^{-1} \geq 1, \quad \theta = b/a \pmod{1} \in [0, 1).$$

For fixed a , the ratio b/a ranges uniformly over $(c - 1, c]$. Consequently,

$$dm = 2c^{-3} dc d\theta, \quad m(c > K) = K^{-2} \quad (K > 1). \quad (13)$$

The lattice points with positive second coordinate are

$$\left(\frac{q\theta - p}{c}, qc \right), \quad p \in \mathbb{Z}, \quad q \in \mathbb{N}. \quad (14)$$

Replacing b/a by its representative modulo one changes p by an integer multiple of q , so primitivity is still equivalent to $\gcd(p, q) = 1$.

Fix $K > 1$, let $Q_n = \lfloor 2n/K \rfloor$, and count only pairs satisfying

$$1 \leq q \leq Q_n, \quad c - \varepsilon < q\theta - p \leq c. \quad (15)$$

If $1 \leq c \leq K$, their lattice vectors lie in the target rectangle:

$$1 - \varepsilon \leq 1 - \varepsilon/c < \frac{q\theta - p}{c} \leq 1, \quad qc \leq 2n.$$

The narrower band in (15) has width independent of c , which will allow an exact second-moment calculation.

Choose $P \geq 2$ such that

$$M = \prod_{\substack{\ell \leq P \\ \ell \text{ prime}}} \ell \geq K.$$

Let $G_M(c, \theta)$ count the pairs in (15) for which $\gcd(p, q, M) = 1$, and let $G_{\text{prim}}(c, \theta)$ count those with $\gcd(p, q) = 1$. Thus G_M imposes only that no prime at most P divide both coordinates. These counting functions make sense for every real c , not only for $c \geq 1$. Write

$$\varrho_M(q) = \prod_{\substack{\ell | \gcd(q, M) \\ \ell \text{ prime}}} (1 - \ell^{-1}), \quad \beta_n = \varepsilon \sum_{q \leq Q_n} \varrho_M(q).$$

Temporarily equip $[0, 1) \times [0, M)$, with coordinates (θ, c) , with normalized Lebesgue measure λ_M . Since $\varepsilon < 1$, the contribution for each fixed q is either zero or one and equals

$$I_q(c, \theta) = f_q(q\theta - c), \quad f_q(u) = \sum_{\substack{p \in \mathbb{Z} \\ \gcd(p, q, M) = 1}} \mathbf{1}_{(p-\varepsilon, p]}(u).$$

The function f_q is M -periodic, and also satisfies $f_q(u + q) = f_q(u)$ because addition of q preserves the coprimality condition. Thus I_q is 1-periodic in θ , as is every product $I_q I_{q'}$. Since M is an integer, their means on $[0, 1) \times [0, M)$ equal their means on the torus $(\mathbb{R}/M\mathbb{Z})^2$.

For $q \neq q'$, the map

$$(\theta, c) \mapsto (q\theta - c, q'\theta - c) \pmod{M}$$

is a surjective endomorphism of this torus: its integer matrix has nonzero determinant $q' - q$. It therefore pushes Haar probability measure to Haar probability measure. Its two output coordinates are independent and uniform, proving pairwise independence of I_q and $I_{q'}$ under λ_M . There are $M\varrho_M(q)$ admissible residue classes for p modulo M , so $\mathbb{E}_{\lambda_M} I_q = \varepsilon \varrho_M(q)$. Hence

$$\mathbb{E}_{\lambda_M} G_M = \beta_n, \quad \text{Var}_{\lambda_M}(G_M) = \sum_{q \leq Q_n} \text{Var}_{\lambda_M}(I_q) \leq \beta_n. \quad (16)$$

The function ϱ_M is M -periodic in q . By the Chinese remainder theorem its average over one period is

$$\delta_M = \prod_{\substack{\ell \leq P \\ \ell \text{ prime}}} (1 - \ell^{-2}) \geq \zeta(2)^{-1}.$$

Indeed, for a prime $\ell \mid M$, the factor is one for the $\ell - 1$ nonzero residues of q and $1 - \ell^{-1}$ for the zero residue, giving the average $1 - \ell^{-2}$. Therefore, with K, P, α fixed,

$$\beta_n \longrightarrow \frac{2\alpha}{K} \delta_M. \quad (17)$$

On $1 \leq c \leq K$, the density in (13) is at most 2, whereas the density of λ_M is $1/M$. Chebyshev's inequality and (16) give

$$m(c \leq K, G_M < \beta_n/2) \leq 2M\lambda_M(G_M < \beta_n/2) \leq \frac{8M}{\beta_n} \quad (18)$$

for all sufficiently large n .

It remains to control points surviving the finite sieve that are not primitive. Each such point has a prime divisor $\ell > P$ common to p and q . For fixed c and $q = \ell q'$, setting $p = \ell p'$ in (15) gives

$$c/\ell - \varepsilon/\ell < q'\theta - p' \leq c/\ell.$$

Since $q'\theta \pmod{1}$ is uniform for $\theta \in [0, 1)$, the mean number of such p' is exactly ε/ℓ . With

$$S_P = \sum_{\substack{\ell > P \\ \ell \text{ prime}}} \ell^{-2},$$

a union bound and integration against (13) yield

$$\mathbb{E}_m(G_M - G_{\text{prim}}) \leq \sum_{\substack{\ell > P \\ \ell \text{ prime}}} \left\lfloor \frac{Q_n}{\ell} \right\rfloor \frac{\varepsilon}{\ell} \leq \varepsilon Q_n S_P. \quad (19)$$

The estimate is uniform in c , so this integration introduces no additional factor.

If $c \leq K$ and $Z_{n,\alpha} = 0$, then $G_{\text{prim}} = 0$. Either $G_M < \beta_n/2$ or $G_M - G_{\text{prim}} \geq \beta_n/2$. Equations (13), (18), and (19), together with Markov's inequality, therefore give

$$m(Z_{n,\alpha} = 0) \leq K^{-2} + \frac{8M}{\beta_n} + \frac{2\varepsilon Q_n}{\beta_n} S_P. \quad (20)$$

Taking $n \rightarrow \infty$ and using (17),

$$\limsup_{n \rightarrow \infty} m(Z_{n,\alpha} = 0) \leq K^{-2} + \frac{4MK}{\alpha \delta_M} + \frac{2S_P}{\delta_M}. \quad (21)$$

Given a prescribed error, first choose K large and then P large so that $M \geq K$ and both K^{-2} and $2\zeta(2)S_P$ are small. This is possible since $S_P \leq \sum_{k > P} k^{-2} \rightarrow 0$. Keep K, P fixed and let $\alpha \rightarrow \infty$ to eliminate the middle term. This proves (12) in the stated order of limits. $\square$

We now compare the height in Lemma 5 with the flow time of the n th return to the smaller section.

Proof of Theorem 1. For fixed $\alpha > 0$ and $n > \alpha$, take

$$r_n = 1 - \alpha/n, \quad Y_{n,\alpha} = Y_{r_n}, \quad D_{n,\alpha} = \mathcal{R}_{r_n,n} - n.$$

Lemma 2 gives measure-preserving conjugacies to the induced systems that converge to the identity in measure. Also,

$$n(m(Y_{r_n})^{-1} - 1) = n((1 - \alpha/n)^{-2} - 1) \rightarrow 2\alpha.$$

Thus conditions (i) and (ii) of Theorem 4 hold, and Proposition 3 supplies weak mixing.

By (4) and $\pi^2/3 > 3$,

$$m_{r_n}(H_{r_n,n} > 3n) \rightarrow 1.$$

For large n , one has $r_n > 2/3$, and every vector in $(r_n, 1] \times (0, 2n]$ has slope at most $2n/r_n < 3n$. On the event $H_{r_n, n} > 3n$, a primitive vector in this rectangle therefore gives a return to the original section strictly before the n th return to the smaller section, but is not itself a return to the smaller section. Consequently, on Y_{r_n} ,

$$\{D_{n,\alpha} = 0\} \subset \{Z_{n,\alpha} = 0\} \cup \{H_{r_n, n} \leq 3n\}.$$

Since $m_{r_n}(E) \leq r_n^{-2}m(E)$ for every measurable E , Lemma 5 implies

$$\lim_{\alpha \rightarrow \infty} \limsup_{n \rightarrow \infty} m_{r_n}(D_{n,\alpha} = 0) = 0.$$

Condition (iii) holds as well, and Theorem 4 proves strong mixing of B . □

References

- [1] A. Artiles, *The return map of the cross section of horizontally short lattice surfaces is weakly mixing*, arXiv:2510.15450 (2025). <https://arxiv.org/abs/2510.15450>.
- [2] J. S. Athreya and Y. Cheung, *A Poincaré section for the horocycle flow on the space of lattices*, Int. Math. Res. Not. IMRN **2014** (2014), no. 10, 2643–2690. doi:10.1093/imrn/rnt003.
- [3] V. Augustin, F. P. Boca, C. Cobeli, and A. Zaharescu, *The h -spacing distribution between Farey points*, Math. Proc. Cambridge Philos. Soc. **131** (2001), no. 1, 23–38. doi:10.1017/S0305004101005187.
- [4] F. P. Boca, C. Cobeli, and A. Zaharescu, *A conjecture of R. R. Hall on Farey points*, J. Reine Angew. Math. **535** (2001), 207–236. doi:10.1515/crll.2001.049.
- [5] F. P. Boca and A. Zaharescu, *The correlations of Farey fractions*, J. London Math. Soc. (2) **72** (2005), no. 1, 25–39. doi:10.1112/S0024610705006629.
- [6] Y. Cheung and A. Quas, *BCZ map is weakly mixing*, arXiv:2403.14976 (2024). <https://arxiv.org/abs/2403.14976>.
- [7] M. Kac, *On the notion of recurrence in discrete stochastic processes*, Bull. Amer. Math. Soc. **53** (1947), no. 10, 1002–1010. doi:10.1090/S0002-9904-1947-08927-8.

A The weak-mixing input

We give the Cheung–Quas argument [6] for Proposition 3, with explicit counting bounds and the time scaling from Lemma 2. The required geometric fact is that there are constants $\alpha_0, \eta > 0$ such that, for $r_n = 1 - \alpha_0/n$ and all sufficiently large n ,

$$m\{y \in Y_{r_n} : \mathcal{R}_{r_n, n}(y) = n + 1\} \geq \eta. \quad (22)$$

Fix for now $0 < \alpha < 1/32$, put $\varepsilon = \alpha/n$ and $r = 1 - \varepsilon$, and let

$$\mathcal{A}_r = \{(s, t) \in \Omega : 1/2 \leq s \leq r\}.$$

Let F_1, F_2 count primitive points of $\Lambda_{s, t}$ in

$$(r, 1] \times (0, 2n], \quad (r, 1] \times (0, 4n],$$

respectively.

To bound the first moment of F_1 from below, restrict to pairs

$$n/6 < p \leq n/3, \quad n/2 \leq q \leq n, \quad \gcd(p, q) = 1. \quad (23)$$

Their number is $n^2/(2\pi^2) + O(n \log n)$. Indeed, if μ_{ar} denotes the Möbius function, the identity $\mathbf{1}_{\gcd(p, q)=1} = \sum_{d|p, d|q} \mu_{\text{ar}}(d)$ gives this count as

$$\begin{aligned} & \sum_{d \leq n} \mu_{\text{ar}}(d) \left(\left\lfloor \frac{n}{3d} \right\rfloor - \left\lfloor \frac{n}{6d} \right\rfloor \right) \left(\left\lfloor \frac{n}{d} \right\rfloor - \left\lfloor \frac{n}{2d} \right\rfloor + 1 \right) \\ &= \frac{n^2}{12} \sum_{d \leq n} \frac{\mu_{\text{ar}}(d)}{d^2} + O \left(\sum_{d \leq n} \left(\frac{n}{d} + 1 \right) \right) = \frac{n^2}{2\pi^2} + O(n \log n). \end{aligned}$$

Here $\sum_{d \geq 1} \mu_{\text{ar}}(d) d^{-2} = \zeta(2)^{-1}$ and its tail is $O(n^{-1})$. In particular, there are at least $n^2/20$ such pairs for all sufficiently large n .

For each pair in (23), consider

$$\frac{ps + r}{q} < t \leq \frac{ps + 1}{q}, \quad 6/7 \leq s \leq r.$$

This strip is contained in $\mathcal{A}_r$ for large n . Its upper boundary is at most $2/3 + 2/n \leq 1$ for $n \geq 6$, while $p/q > 1/6$ ensures $s + t > 1$. Its vertical width is ε/q , and its horizontal length is eventually at least $1/8$. Moreover, $q/s \leq 2n$, so every point of this strip contributes the primitive vector $(-ps + qt, q/s)$ to F_1 . The normalized area is at least $\varepsilon/(4n)$, and hence

$$\int_{\mathcal{A}_r} F_1 \, dm \geq \frac{n\varepsilon}{80} = \frac{\alpha}{80}. \quad (24)$$

For the second moment, every vector counted by F_2 is of the form $(-ps + qt, q/s)$ with

$$1 \leq q \leq 4n, \quad 0 \leq p < 2q, \quad \gcd(p, q) = 1.$$

Indeed, $p \leq -1$ would give $-ps + qt \geq s + t > 1$, while the other bounds follow from $s \geq 1/2, t \leq 1$, and $q/s \leq 4n$. Let $A_{p, q}$ be the intersection of $\mathcal{A}_r$ with the strip

$$(ps + r)/q < t \leq (ps + 1)/q.$$

Dropping the further height restriction $q/s \leq 4n$ only increases the count, so intersections of these strips bound the factorial second moment of F_2 from above.

For a fixed q , distinct values of p give disjoint strips, since $s \geq 1/2 > \varepsilon$. For distinct primitive indices with $q < q'$, put $\Delta = q' - q$ and $D = p'q - pq'$. Coprimality implies $D \neq 0$. The two full strips intersect in a parallelogram of area $\varepsilon^2/|D|$ and horizontal width

$$\frac{\varepsilon(q + q')}{|D|} \leq 8n\varepsilon = 8\alpha < 1/4.$$

The intersection of their upper boundaries has horizontal coordinate $s_0 = \Delta/D$. If the parallelogram meets $\mathcal{A}_r$, then $s_0 \in [1/4, 5/4]$. Therefore

$$\frac{4}{5}\Delta \leq D \leq 4\Delta, \quad m(A_{p,q} \cap A_{p',q'}) \leq \frac{5\varepsilon^2}{2\Delta}. \quad (25)$$

Fix $q < q'$ and let $d = \gcd(q, q')$. The integer D is a positive multiple of d not exceeding 4Δ , leaving at most $4\Delta/d$ possibilities. For each fixed D , all integer solutions are of the form

$$p = p_0 + (q/d)k, \quad p' = p'_0 + (q'/d)k.$$

The interval $0 \leq p < 2q$ permits at most $2d$ integer values of k . Thus at most 8Δ pairs (p, p') can contribute for these q, q' . At this point we may drop coprimality in the upper bound: it has already excluded parallel strips, so division by D is justified. Equation (25) gives

$$\sum_{p, p'} m(A_{p,q} \cap A_{p',q'}) \leq 20\varepsilon^2,$$

where the sum is over the possible contributing primitive pairs. Summing over $q < q'$ and accounting for both orders yields

$$\int_{\mathcal{A}_r} F_2(F_2 - 1) dm \leq 2 \binom{4n}{2} 20\varepsilon^2 \leq 320\alpha^2. \quad (26)$$

Since $F_1 \leq F_2$, pointwise we have

$$F_1 \leq \mathbf{1}_{\{F_1=F_2=1\}} + F_2 \mathbf{1}_{\{F_2>1\}}, \quad F_2 \mathbf{1}_{\{F_2>1\}} \leq F_2(F_2 - 1).$$

Thus (24) and (26) imply

$$m\{(s, t) \in \mathcal{A}_r : F_1 = F_2 = 1\} \geq \alpha/80 - 320\alpha^2.$$

Choose $\alpha_0 = 1/51200$, so that this lower bound is at least $\alpha_0/160$. By (4) and $3 < \pi^2/3 < 4$, the complement of

$$\{y \in Y_{r_n} : 3n < H_{r_n, n}(y) < 4n\}$$

has m -measure tending to zero. Restrict to points satisfying both $F_1 = F_2 = 1$ and $3n < H_{r_n, n} < 4n$. The unique point counted by F_1 has slope less than $3n$ for large n . Conversely, every outer-strip primitive vector whose slope is below $H_{r_n, n} < 4n$ has vertical coordinate less than $4n$ and is counted by F_2 . There is therefore exactly one extra ordinary return. This proves (22), for example with $\eta = \alpha_0/320$.

To finish the proof of Proposition 3, suppose that $f \circ B = \lambda f$ is a nonzero $L^2(m)$ eigenfunction with $\lambda \neq 1$. Measure preservation gives $|\lambda| = 1$, and ergodicity permits the normalization $|f| = 1$ almost everywhere. Set $f_r = f \circ \phi_r^{-1}$ on Y_r . The conjugacy gives

$$f_r \circ B_r = \lambda f_r, \quad \|f_r - f|_{Y_r}\|_{L^2(m_r)} \longrightarrow 0.$$

For the second assertion, change variables by ϕ_r and approximate f in $L^2(m)$ by continuous functions. The error after composition with ϕ_r is controlled by $(\phi_r)_*m = m_r \leq r^{-2}m$, and the continuous part converges by $\phi_r \rightarrow \text{id}$ in measure. Since B_r preserves m_r ,

$$\|f \circ B_r^n - \lambda^n f\|_{L^2(m_r)} \leq 2\|f_r - f|_{Y_r}\|_{L^2(m_r)} \longrightarrow 0$$

along $r = r_n$, independently of the growing exponent n . On the set in (22), however, $B_r^n = B^{n+1}$, so the expression inside this norm has absolute value $|1 - \lambda| > 0$. That set has m_r -measure at least η , a contradiction. Ergodicity already excludes nonconstant eigenfunctions with eigenvalue one. The absence of nonconstant L^2 eigenfunctions proves weak mixing.