SPATIAL LIMIT THEOREM
FOR INTERVAL EXCHANGE TRANSFORMATIONS

ALEXEY KLIMENKO

Abstract. We prove a spatial limit theorem for generic interval exchange transformations (IETs): for a generic IET the normalized ergodic sums of a sufficiently regular (e.g., Lipschitz) function have the same asymptotic behavior of distributions as the behavior of ergodic integrals for a generic translation flow on a flat surface, described by A. Bufetov.

1. Introduction

Let \( T: X \to X \) be a transformation of a probability space \((X, \mu)\) that conserves the measure \( \mu \). For a function \( \varphi \in L^1(X, \mu) \) consider its ergodic sums

\[
S_n \varphi (x) = \sum_{k=0}^{n-1} \varphi (T^k x).
\]

One can regard \( S_n \varphi \) as a random variable on the probability space \((X, \mu)\), and we are interested in asymptotical behavior of the distributions of appropriately normalized random variables \( S_n \varphi \) as \( n \to \infty \). Following Dolgopyat and Sarig [4], we refer to results of this type as spatial (distributional) limit theorems.

Our main result states spatial limit theorem for a generic interval exchange transformation. To state precisely what genericity means here, we need some notation for translation flows on flat surfaces, which are continuous-time counterparts of IETs. Here we follows Bufetov’s paper [2].

In a moduli space of Abelian differentials let \( \mathcal{M}_\kappa \) be a stratum of surfaces such that the orders of their singularities form a tuple \( \kappa = (\kappa_1, \ldots, \kappa_\sigma) \), endowed with a choice of a “horizontal” separatrix of one of the singularities. Let \( \mathcal{H} \) be a connected component of \( \mathcal{M}_k \) and let \( g^s \) be the Teichmüller flow on \( \mathcal{H} \). Finally, consider a \( g^s \)-invariant ergodic probability measure \( \mathbb{P} \) on \( \mathcal{H} \).

For every surface \( M \in \mathcal{M}_\kappa \) there is a unique decomposition of \( M \) into Markov rectangles such that the bases of these rectangles belong to a segment \( I \) of the marked separatrix starting at the singular point at the end of this separatrix, the length of this segment is at least one, while the application of the Rauzy induction makes it less than one. Therefore, the first return map to the segment \( I \) for the vertical flow \( h_t^+ \) on \( M \) yields an interval exchange transformation \( T_M \). The number of segments for \( T_M \) is equal to the number of rectangles, and hence is fixed. Moreover, permutations of segments for \( T_M, M \in \mathcal{H} \) belong to the same Rauzy class almost surely. Indeed, \( T_M \) and \( T_{g^s \kappa^* M} \) differ by a scaling and several applications of the Rauzy induction, hence their permutations belong to the
same Rauzy class, so this Rauzy class is $g$-invariant. But $\mathbb{P}$ is ergodic, so this Rauzy class should be the same almost surely.

Therefore, one can consider a set $\hat{\mathcal{H}} = \{T_M, M \in \mathcal{H}\}$, and the measure $\mathbb{P}$ induces a measure $\hat{\mathbb{P}}$ on $\hat{\mathcal{H}}$.

A. Bufetov [2] proved a limit theorem for $\mathbb{P}$-almost all translation froms provided some conditions on $\mathbb{P}$ hold. Namely, for sufficiently regular function $\psi$ on $M$ outside of a finite-codimension “degenerate” subspace, the distributions of random variables

$$\hat{I}_t \psi = \frac{I_t \psi - \mathbb{E} I_t \psi}{\sqrt{\text{Var} I_t \psi}}, \quad \text{where} \quad I_t \psi(x) = \int_0^t \psi(h_x^+(s)) \, ds,$$

have no limit as $t \to \infty$ but its set of limit points in the space of distributions can be described, see Section 2 for the detailed statement.

**Main Theorem.** Assume that a $g^s$-ergodic measure $\mathbb{P}$ on a connected component $H$ of a stratum of Abelian differentials satisfy conditions of Bufetov’s theorem. Then for $\hat{\mathbb{P}}$-almost every $T \in \hat{\mathcal{H}}$ and for a regular function $f$ on the segment $I$ where $T$ acts if $f$ lies outside of a finite-codimensional subspace, then the sequence of distributions of the random variables

$$(1) \quad \hat{S}_n f = \frac{S_n f - \mathbb{E} S_n f}{\sqrt{\text{Var} S_n f}}$$

have the same limit behaviour as the behaviour of distributions of ergodic integral for some $M \in \mathcal{H}$ with $T_M = T$ and some function $\psi$ on $M$.

The structure of the paper is the following. We discuss Bufetov’s limit theorem for translation flows in the next section, then in Section 3 we give the detailed statement of our Main Theorem, which is then proved in this section and the last Section 4.

2. **Bufetov’s limit theorem for translation flows**

In this section we mostly follows [2, Section 1].

A bounded measurable function $\varphi$ on a flat surface $M$ is called weakly Lipschitz if there exists a constant $C > 0$ such that for any $x \in M$ and any $t_1, t_2 > 0$ such that the rectangle $\Pi(x, t_1, t_2) = \{h_x^+(t) : 0 \leq t \leq t_1, 0 \leq s \leq t_2\}$ is admissible, i.e. does not contain singular points in its interior, we have

$$(2) \quad \left| \int_0^{t_1} \varphi(h_x^+(s)) \, ds - \int_0^{t_2} \varphi(h_x^+(s)) \, ds \right| \leq C.$$

The set $\text{Lip}_{\psi}^w(M)$ of all weakly Lipschitz functions is a normed space with the norm defined by

$$\|\varphi\|_{\text{Lip}_{\psi}^w} = \sup_M |\varphi| + C_{\varphi},$$

where $C_{\varphi}$ is the infimum of all $C$ satisfying (2).

The main tool to describe behaviour of ergodic integrals for a translation flow is the following Bufetov cocycles.

**Definition 2.1.** A function $\Phi^+$ on arcs of the vertical flow $h_x^+$ on $M$ is called vertical Bufetov cocycle if the following holds.
1) $\Phi^+$ is additive: if $\Phi^+(x,t)$ denotes its value on the arc $h^+_{[0,t]}(x) = \{ h^+_s(x) : 0 \leq s \leq t \}$, then

$$\Phi^+(x,t_1 + t_2) = \Phi^+(x,t_1) + \Phi^+(h^+_t(x),t_2).$$

2) $\Phi^+$ is Hölder: there exists $t_0 > 0$, $\theta > 0$ such that $|\Phi^+(x,t)| < t^\theta$ for all $x \in M$, $|t| \leq t_0$.

3) $\Phi^+$ is holonomy invariant: if $\Pi(x,t_1,t_2)$ is admissible, then $\Phi^+(x,t_1) = \Phi^+(h^+_t(x),t_1)$.

The space of all such cocycles is denoted by $\mathcal{B}^+ = \mathcal{B}^+(M)$. One can also define the symmetric class of horizontal Bufetov cocycles $\mathcal{B}^-$. 

The spaces $\mathcal{B}^\pm(M)$ are isomorphic to the spaces of Kontsevich—Zorich cocycles$[6,9]$ of the surface $M$. Similar functional objects also arose earlier in the work of G. Forni$[5]$.

The base for the study of the Bufetov cocycles is matrix cocycles given by the incidence matrices of the Markov partition of the surface into rectangles and the same matrices for partitions produced from it by forward or backward Rauzy induction. These matrix cocycles $A$ on the space $(H,\mathcal{P})$ are assumed to satisfy conditions of Oseledets—Pesin reduction theorem (for details see$[2, \text{Assumption 2.3}]$).

Bufetov showed that the spaces $\mathcal{B}^\pm$ have finite dimensions and there exists a nondegenerate pairing $\langle \cdot, \cdot \rangle : \mathcal{B}^+ \times \mathcal{B}^- \to \mathbb{C}$. For a cocycle $\Phi^-$ one can define a functional $m_{\Phi^-}$ on the space of weakly Lipschitz functions as follows: we decompose $M$ into admissible rectangles, so $m_{\Phi^-}(\varphi)$ is the sum over all these rectangles $\Pi = \Pi(x,t_1,t_2)$ of the integrals $\int_{t_0}^{t_1} \int_0^{t_1} f(h^+_s(h^-_s(x))) \, dt \, d\Phi^-(s)$, where the external integral can be defined as a limit of Riemann sums. The map $\varphi \mapsto m_{\Phi^-}(\varphi)$ is linear in $\Phi^-$ so the pairing between $\mathcal{B}^+$ and $\mathcal{B}^-$ allows us to define the functional $\Phi^+_{\varphi}$ by the formula

$$m_{\Phi^-}(\varphi) = \langle \Phi^+_{\varphi}, \Phi^- \rangle \quad \text{for any} \quad \Phi^- \in \mathcal{B}^-.$$

The Teichmüller flow on $\mathcal{H}$ naturally acts on the fiber bundles of Bufetov’s cocycles over it. Hence one can decompose the space $\mathcal{B}^+(M)$ into its Oseleets subspaces:

$$\mathcal{B}^+(M) = \bigoplus_{i=1}^l \mathcal{B}^+_i(M)$$

(3)

corresponding to the Lyapunov exponents $1 = \theta_1 > \theta_2 \cdots > \theta_l > 0$. The space $\mathcal{B}^+_i(M)$ is the linear hull of the cocycle $\Phi^+_1(x,t) = t$.

Bufetov proved the following asymptotic decomposition of the ergodic integral for the generic translation flows.

**Theorem 2.2** (2 Theorem 1). Let $\mathbb{P}$ be an ergodic $g$-invariant probability measure on $\mathcal{H}$. Then for any $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that for $\mathbb{P}$-almost every $M \in \mathcal{H}$, any $\varphi \in \text{Lip}_+^\infty(M)$, any $x \in M$, and any $T > 0$ we have

$$\left| \int_0^T \varphi(h^+_s(x)) \, dt - \Phi^+_{\varphi}(x,T) \right| \leq C_\varepsilon \| \varphi \|_{\text{Lip}_+^\infty}(1 + T^\varepsilon).$$

Later we will need the following truncated version of this formula: let $\Phi^+_{\varphi,\leq r}$ be a projection of $\Phi^+_{\varphi}$ onto the first $r$ terms $\mathcal{B}^+_1 \oplus \cdots \oplus \mathcal{B}^+_r$ in the Oseleets decomposition.
Then
\begin{equation}
\int_0^T \varphi(h_t^+(x)) \, dt - \Phi_{\varphi,\leq t}^+(x, T) \leq C_\varepsilon \|\varphi\|_{1\text{Lip}_+^+}(1 + T^{\theta_r - \varepsilon}).
\end{equation}
for some $\varepsilon > 0$.

We now pass to Bufetov’s limit theorem. Informally speaking, it states that the distribution of the normalized ergodic integral for $\varphi$ tends to the distribution of the first nonzero component of $\Phi_{\varphi}^+$. Let $\Phi_{\varphi}^+ = \sum_{i=1}^l \Phi_{\varphi,i}^+$, where $\Phi_{\varphi,i}^+ \in \mathcal{B}_i^+$, then denote $i(\varphi) = \min\{i = 2, \ldots, l : \Phi_{\varphi,i}^+ \neq 0\}$.

The Teichmüller flow on the moduli space induces natural action on the bundle $\{\mathcal{B}^+(M), M \in \mathcal{H}\}$. Namely, for $\Phi^+ \in \mathcal{B}^+(M)$ and $x \in M$ let $y = g^s x$ be the image of $x$ and denote
\begin{equation}
(g^s \Phi^+)(y, t) := \Phi^+(x, e^st),
\end{equation}
then we have $g^s \Phi^+ \in \mathcal{B}^+(g^s M)$. Moreover, the decomposition (3) is invariant under this action, and we denote the action of $g^s$ on the $\mathcal{B}_i^+$ by $g^s_i$.

As one can see from the approximation theorem, the distribution of $\hat{I}_t \psi$ is close to the distribution of $\Phi^+_\psi(\cdot, t)$, and more precisely, to the distribution $\Phi^+_\psi,i(\psi)(\cdot, t)$. The scaling (3) allows us to identify the last distribution with the one of $(g^{\log_2 T} \Phi^+_\psi,i(\psi))(\cdot, 1)$. This yields the following theorem.

**Theorem 2.3** ([2, Theorem 2]). Let the above-mentioned matrix cocycles satisfy the conditions of Oseledets—Pesin theorem with respect to an ergodic $g^s$-invariant probability measure $\mathbb{P}$. Then there exists a constant $\alpha > 0$ and a measurable function $C: \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+$ such that for $\mathbb{P}$-almost every $M \in \mathcal{H}$ and every $\psi \in \text{Lip}^w_+(M)$ such that $\Phi^+_\psi \notin \mathcal{B}_1^+(M)$ (and hence $i(\psi)$ is well defined), we have
\begin{equation}
d(\text{Law}(\hat{I}_T(\psi)), \text{Law}((g^{\log_2 T} \Phi^+_\psi,i(\psi))(\cdot, 1))) \leq C(M, g^{\log_2 T} M) T^{-\alpha},
\end{equation}
where $d$ is either Kantorovich—Rubinstein or Lévy—Prohorov distance on the space of distributions.

Note that in the case of simple Lyapunov exponents the formula (6) has a simpler form: $\text{Law}((g^{\log_2 T} \Phi^+_\psi,i(\psi))(\cdot, 1))$ is the normalized distribution of any nonzero cocycle in $\mathcal{B}_i^+(g^{\log_2 T} M)$ since the normalization depends only on the projectivization of $g^{\log_2 T} \Phi^+_\psi,i(\psi)$, and $\mathcal{B}_i^+$ is one-dimensional.

**Remark 2.4.** In particular, Bufetov’s theorems applies to the Masur—Veech “smooth” measure $\mathbb{P}$, [7, 8]. In this case the corresponding measure $\hat{\mathbb{P}}$ on the space of interval exchange transformations is equivalent to the Lebesgue measure. As Avila and Viana showed in [1], the second Lyapunov exponent for the Masur—Veech measure is simple.

Let us also recall the recent result by Chaika and Eskin [3]. It implies that for any given flat surface the set of directions such that the corresponding translation flow satisfies assumptions of Bufetov’s theorems has full Lebesgue measure on the circle.
3. Main theorem and approximation of ergodic sums via integrals

Let us give the precise statement for the main theorem of this paper.

For \( \hat{P} \)-almost every interval exchange transformation \( T \) there exists a surface \( M \in \mathcal{H} \) with \( T = T_M \) such that Bufetov’s theorems hold for \( M \). The surface \( M \) is decomposed into several rectangles of heights \( h_i \) with bases on the continuity segments \( I_i \) of \( T \). Introduce a coordinate system \((x, y)\) on the union of these rectangles so that \( x \) is a “horizontal” coordinate and \( y \) is a “vertical” one, i.e. the flow \( h_i^+ \) corresponds to the differential equations \( \dot{x} = 0, \dot{y} = 1 \). In other words, the flow \( h_i^+ \) on \( M \) is the special flow over \( T \) with the roof function

\[
    h(x) = h_i \text{ for } x \in I_i.
\]

For any function \( f \) on \( I \) define the following function \( \psi_f \) on \( M \).

\[
    (7) \quad \psi_f(x, y) = \frac{f(x)}{h(x)}.
\]

**Theorem 3.1** (Main theorem restated). Let \( \hat{P} \) be an ergodic \( g \)-invariant probability measure on a connected component \( \mathcal{H} \) of the space of Abelian differentials and assume that \( \hat{P} \) satisfies the conditions of Bufetov’s limit theorem. Then for any \( \varepsilon > 0 \) there exists a measurable function \( C: \mathcal{H} \times \mathcal{H} \to \mathbb{R}_+ \) such that the following holds.

Take any \( T \in \hat{H} \) such that there exists \( M \in \mathcal{H} \) with \( T = T_M \) such that the conditions of Bufetov’s approximation theorem hold for \( M \) (recall that this takes place for \( \hat{P} \)-almost all \( T \)). Then for all functions \( f \) on \( I \) such that \( \psi_f \in \text{Lip}^u(M) \) and \( \Phi^+_{\psi_f} \neq 0 \) we have

\[
    d_{KR}(\text{Law}(\hat{S}_n f), \text{Law}((g_2^{\log n} \Phi^+_{\psi_f})(\cdot, 1))) \leq C(M, g^{\log n M}) n^{\theta^2_2 - \theta_2 + \epsilon},
\]

\[
    d_{LP}(\text{Law}(\hat{S}_n f), \text{Law}((g_2^{\log n} \Phi^+_{\psi_f})(\cdot, 1))) \leq C(M, g^{\log n M}) n^{\frac{2}{3}(\theta^2_2 - \theta_2) + \epsilon}.
\]

**Remark 3.2.** Note that if \( f \) is Lipschitz on each continuity segment \( I_i \) of the interval exchange transformation \( T \), then \( \psi_f \in \text{Lip}^u(M) \). Indeed, the function \( \psi_f \) is Lipschitz continuous inside any admissible rectangle with the same constant \( L_f \cdot (1/ \min(h_j)) \), where \( L_f \) is the Lipschitz constant for \( f \). It remains to observe that the product of the height and the width of any admissible rectangle does not exceed the area of \( M \), which equals 1.

The first step of the proof is to approximate an ergodic sum for the interval exchange transformation \( T: I \to I \) by an ergodic integral for translation flow on some surface \( M \) such that \( T = T_M \). Until the end of the paper we assume that the setting and the conditions of Theorem 3.1 hold.

**Proposition 3.3.** 1. Let \( t_n(x) \) be the time when the point \((x, 0)\) under the action of the flow \( h_i^+ \) makes its \( n \)-th return to the transversal \( I = \{y = 0\} \). Then one have

\[
    (8) \quad S_n f(x) = t_n(x) \psi_f(x, 0).
\]

2. \( \int_I f(x) \, dx = \int_M \psi_f(x, y) \, dxdy \).
Proof. 1. Observe that
\[ \int_0^{t_1(x)} \psi_f(h^+_t(x,0)) \, dt = \int_0^{h(x)} \frac{f(x)}{h(x)} \, dt = f(x). \]
Now
\[ I_{t_n(x)} \psi_f(x,0) = \sum_{k=0}^{n-1} \int_{t_k(x)}^{t_{k+1}(x)} \psi_f(h^+_t(x,0)) \, dt = \sum_{k=0}^{n-1} \int_0^{h(T^k(x))} \frac{f(T^k x)}{h(T^k x)} \, dt = S_n f(x). \]
2. The integrand in the right-hand side of the equality does not depend on \( y \), and its integration in \( y \) yields the left-hand side.

Proposition 3.4. Let \( \psi_1 \) be the function on \( M \) constructed by (4) from the constant function 1 on \( I \). Then for any \( \delta > 0 \) there exists \( A_\delta \) such that
\[ \left| t_n(x) - \left[ n - \Phi^+_{\psi_1,2}((x,0),n) \right] \right| \leq A_\delta (1 + n^{\theta_2+\delta}). \]

Remark 3.5. Later we will use the \( O \)-notation for such estimates. The constants in \( O(\cdot) \) can depend on the surface \( M \) and small parameters \( \delta, \varepsilon \), etc.

Proof. Applying the formula (8) to the function \( f \equiv 1 \), we get the equation
\[ n = I_{t_n(x)} \psi_1(x,0), \]
and we will find asymptotics for its solution.

First of all, (3) with \( r = 1 \) gives that for any \( \alpha > 0 \) there exists \( C > 0 \) such that for all \( t > 0 \) we have
\[ |I_t \psi_1 - \overline{\psi_1}t| \leq \alpha t + C, \quad \text{where} \quad \overline{\psi_1} = \int_M \psi_1 \, dxdy = 1. \]
Therefore,
\[ (1 - \alpha)t_n(x) - C \leq n \leq (1 + \alpha)t_n(x) + C, \]
hence, conversely, for any \( \beta > 0 \) and some \( D > 0 \) we have
\[ (1 - \beta)n - D \leq t_n(x) \leq (1 + \beta)n + D. \]

Further, the formula (4) with \( r = 1 \) gives us that
\[ |n - t_n(x)| = |n - \Phi^+_{\psi_1,1}((x,0),t_n(x))| \leq C_\varepsilon (1 + n^{\theta_2+\varepsilon}) \leq C'_\varepsilon (1 + n^{\theta_2+\varepsilon}). \]

Combining this estimate with (4) for \( r = 2 \) we have
\[ |n - t_n(x) - \Phi^+_{\psi_1,2}((x,0),n)| \leq |n - \Phi^+_{\psi_1,2}((x,0),t_n(x))| + |\Phi^+_{\psi_1,2}(h^+_n((x,0),t_n(x)) - n)| \]
\[ \leq C''_\varepsilon (1 + n^{\theta_2+\varepsilon}) + \tilde{C}_\varepsilon (1 + |t_n(x) - n|^{\theta_2+\varepsilon}) \leq C'_\varepsilon (1 + n^{(\theta_2+\varepsilon)^2}). \]

Hence we have obtained that for any \( \delta > 0 \) there exists \( A_\delta > 0 \) such that
\[ |t_n(x) - (n - \Phi^+_{\psi_1,2}((x,0),n))| \leq A_\delta (1 + n^{(\theta_2+\varepsilon)^2}). \]
Now choosing \( \varepsilon \) such that \( (\theta_2 + \varepsilon)^2 \leq \theta_2^2 + \delta \) we obtain the desired estimate. □
Lemma 3.6. Denote $\tilde{f} = (1/|I|) \int_I f(x) \, dx$. Then the following asymptotic estimate holds for $S_n(f)$: for any $\delta > 0$

$$S_n(f) = \tilde{f} \cdot n + \Phi^+_{(\psi_f - \tilde{f} \psi_1),2}((x,0),n) + O(n^{\theta_2 + \delta})\|\psi_f\|_{\text{Lip}_+^w}.$$

Proof. Applying Proposition 3.4 and formula (4) with $r = 2$ to the identity (8) we have

$$S_n f(x) = \tilde{f} \cdot n - \tilde{f} \cdot \Phi^+_{\psi_1,2}((x,0),n) + \Phi^+_{\psi_f,2}((x,0),t_n(x)) + O(n^{\theta_2 + \delta}).$$

The third term in the right-hand side of this formula is approximated as follows:

$$\Phi^+_{\psi_f,2}((x,0),t_n(x)) - \Phi^+_{\psi_f,2}((x,0),n) = \Phi^+_{\psi_f,2}(h_n^+((x,0),t_n(x) - n) =$$

$$= O((t_n(x) - n)^{\theta_2 + \delta})\|\psi_f\|_{\text{Lip}_+^w} = O(n^{\theta_2 + \delta})\|\psi_f\|_{\text{Lip}_+^w}.$$ 

Substituting this estimate with $\epsilon$ such that $(\theta_2 + \epsilon)^2 < \theta_2^2 + \delta$ into (10), we obtain the statement of the lemma.

Later we will use the following extension of this lemma.

Corollary 3.7. Under the assumptions of Lemma 3.6, for any $\epsilon > 0$ one has that for any $\gamma \in [0,1]$

$$S_n(f) = \tilde{f} \cdot n + \Phi^+_{(\psi_f - \tilde{f} \psi_1),2}(h_n^+(-\gamma n^{\theta_2 - \epsilon}(x,0),n) + O(n^{\theta_2 + \delta})\|\psi_f\|_{\text{Lip}_+^w}.$$

where the constant in $O(\cdot)$ does not depend on $\gamma$.

Proof. The difference between asymptotics in (10) and (11) equals

$$\Phi^+_{(\psi_f - \tilde{f} \psi_1),2}(h_n^+((x,0),-\gamma n^{\theta_2 - \epsilon}) - \Phi^+_{(\psi_f - \tilde{f} \psi_1),2}((x,0),-\gamma n^{\theta_2 - \epsilon}) = O((\gamma n^{\theta_2 - \epsilon})^{\theta_2 + \delta})\|\psi_f\|_{\text{Lip}_+^w},$$

since $\|\psi_f\|_{\text{Lip}_+^w} \leq \|\psi_f - \tilde{f} \psi_1\|_{\text{Lip}_+^w} = \|\tilde{f}\|_{\text{Lip}_+^w}$, where $\tilde{f}(x) = f(x) - \tilde{f}$.

4. END OF PROOF: APPROXIMATION OF UNIFORM DISTRIBUTION OF INITIAL POINTS

Lemma 3.6 yields that the distribution of $S_n[f](x)$ with $x$ uniformly distributed on $I$ (below we denote this as $x \sim \text{Unif}(I)$) is close to that of $\Phi^+_{\psi_f,\leq 2}(x,n)$ with $x \sim \text{Unif}(I)$.

However, Theorem 2.3 deals with the distribution $\Phi^+_{\psi_f,\leq 2}(p,n)$, where $p$ is uniformly distributed on the whole surface $M$. To relate these two distributions we consider the intermediate one, that of $\Phi^+_{\psi_f,\leq 2}(q,n)$ with $q = h_n^+(-\gamma n^{\theta_2 - \delta})$, where $x \sim \text{Unif}(I)$ and $\gamma \sim \text{Unif}[0,1]$ are independent. Then such $q$ will be distributed almost uniformly on $M$.

We proceed to the formal considerations. Recall the following property of Kantorovich—Rubinstein and Lévy—Prokhorov distances. (Here and below we write $d(\xi,\eta)$ instead of $d(\text{Law}(\xi),\text{Law}(\eta))$ for brevity.)

Proposition 4.1. 1. $d_{K.R, LP}(\xi, \xi + \epsilon) \leq |\epsilon|$.

2a. $d_{K.R}(\xi, (1 + \epsilon)\xi) \leq |\epsilon|E[\xi].$

2b. $d_{LP}(\xi, (1 + \epsilon)\xi) \leq |\epsilon|^{2/3}(\text{Var} \xi)^{1/3} \left(1 + |\epsilon| \right)^{2/3} \left(1 - |\epsilon| \right)^{1/3}$.
Proof. The first item is clear. To prove item 2a we use “mass transportation” interpretation and shift each piece of mass \(dm\) from a point \(x\) to the point \((1+\varepsilon)x\) yielding work equal to \(|\varepsilon x|\ dm\) and its integral over the whole mass is \(|\varepsilon|E[\varepsilon]\). The last item is proved as follows. Take any Borel set \(A \subset \mathbb{R}\) and split it into \(A_{\leq \mu} = \{x \in A, |x| \leq \mu\}\) and \(A_{> \mu} = A \setminus A_{\leq \mu}\), with the parameter \(\mu\) being chosen below. Then if \(\xi \in A_{\leq \mu}\) then \((1+\varepsilon)\xi \in A^\delta_{\leq \mu}\) assuming \(\delta \leq |\varepsilon|\mu\). Similarly, \((1+\varepsilon)\xi \in A_{\leq \mu}\) yields \(\xi \in A^\delta_{\leq \mu}\) if \(\delta \leq |\varepsilon|\mu/(1-|\varepsilon|)\). Further,

\[
\mathbb{P}(\xi \in A_{> \mu}) \leq \frac{\text{Var} \xi}{\mu^2}, \quad \mathbb{P}((1+\varepsilon)\xi \in A_{> \mu}) \leq \frac{(1+|\varepsilon|)^2 \text{Var} \xi}{\mu^2}.
\]

Therefore, \(d_{LP}(\xi, (1+\varepsilon)\xi) \leq \delta\) if

\[
\frac{(1+|\varepsilon|)^2 \text{Var} \xi}{\mu^2} \leq \delta, \quad \delta \leq \frac{|\varepsilon|\mu}{(1-|\varepsilon|)},
\]

and these conditions are satisfied with \(\delta\) equal to the right-hand side of the inequality in item 2b and \(\mu = \delta(1-|\varepsilon|)/|\varepsilon|\). \(\square\)

**Proposition 4.2.** Let \(q\) be a random point on the surface \(M\) with

\[
q = h^+_{-\gamma T}(x),
\]

where \(x \sim \text{Unif}(I)\) and \(\gamma \sim \text{Unif}[0,1]\) are independent. Then the distribution of \(q\) has density \(\rho_T\) with respect to the uniform measure on \(M\) and we have

\[
|\rho_T - 1| = O(T^{\theta_2+\varepsilon-1}) \text{ for any } \varepsilon > 0.
\]

**Proof.** Considering a small rectangle near a point \((x, y) \in M\) one can see that the density \(\rho_T(x, y)\) equals \(N_T((x, y))/T\), where \(N_T((x, y))\) is the number of intersections of the arc segment \(h^+_{[0, T]}((x, y))\) with the segment \(I\). Hence we have

\[
\left| N_T((x, y)) - \int_0^T \psi_1(h^+_{\gamma T}((x, y))) \, dt \right| \leq 1,
\]

as if we split the integral by the times when the arc crosses \(I\), every integral over the internal interval of the partition equals 1, and the value of the integral over the first or the last interval belongs to \([0, 1]\). Now \([1]\) with \(r = 1\) and \(\varphi = \psi_1\) concludes the proof. \(\square\)

Without loss of generality we may assume that \(\overline{T} = 0\). Consider the following random variables:

\[
\eta_0 = \Phi^+_{\psi, 1, 2}(p, n), \quad p \sim \text{Unif}(M),
\]

\[
\eta_1 = \Phi^+_{\psi, 2}(h^+_{-\gamma \theta_2-\varepsilon}, n), \quad \begin{cases} x \sim \text{Unif}(I) \text{ and } \gamma \sim \text{Unif}[0,1] \\ \eta_2 = S_n f(x), \end{cases}
\]

are independent.

Note that \(\eta_2\) does not depend on \(\gamma\) but it is convenient to regard \(\eta_1\) and \(\eta_2\) as random variables on the same probability space. Let us also denote

\[
\hat{\eta}_k = \frac{\eta_k - \mathbb{E}\eta_k}{\sqrt{\text{Var} \eta_k}}.
\]
Lemma 4.3. For any $\beta > 0$ we have
\[
d_{KR}(\tilde{\eta}_0, \tilde{\eta}_1) = O(n^{\theta_2-\delta}), \\
d_{LP}(\tilde{\eta}_0, \tilde{\eta}_1) = O(n^{\frac{1}{2}(\theta_2-1)+\beta}),
\]

Proof. In the proof we denote $T = n^{\theta_2-\delta}$. Then the lemma statement takes the form
\[
d_{KR}(\tilde{\eta}_0, \tilde{\eta}_1) = O(T^{\theta_2-1+\epsilon}), \\
d_{LP}(\tilde{\eta}_0, \tilde{\eta}_1) = O(n^{\frac{1}{2}(\theta_2-1)+\epsilon})
\]
for any $\epsilon > 0$.

1. Consider the following random variable
\[
\tilde{\eta}_1 = \frac{\eta_1 - \mathbb{E}\eta_0}{\sqrt{\text{Var} \eta_0}}
\]
Then $\tilde{\eta}_0$ and $\tilde{\eta}_1$ can be regarded as the same function on $M$ but with different measures on $M$, namely, the measures given in Proposition 4.2. Therefore the distributions of $\tilde{\eta}_0$ and $\tilde{\eta}_1$ are equivalent: $dF_{\tilde{\eta}_0} = \rho dF_{\tilde{\eta}_1}$, and $\rho = 1 + O(T^{\theta_2-\epsilon})$. Hence
\[
|\mathbb{E}\tilde{\eta}_1| = |\mathbb{E}\tilde{\eta}_1 - \mathbb{E}\tilde{\eta}_0| \leq \int |x| |\rho(x) - 1| dF_{\tilde{\eta}_0} \leq \mathbb{E}|\tilde{\eta}_0| \cdot O(T^{\theta_2-1+\epsilon}) = O(T^{\theta_2-1+\epsilon}),
\]
since $\mathbb{E}|\tilde{\eta}_0| \leq \sqrt{\text{Var} \tilde{\eta}_0} = 1$. Similarly,
\[
|\mathbb{E}\tilde{\eta}_1^2 - \mathbb{E}\tilde{\eta}_0^2| \leq \int x^2 |\rho(x) - 1| dF_{\tilde{\eta}_0} \leq \mathbb{E}\tilde{\eta}_0^2 \cdot O(T^{\theta_2-1+\epsilon}) = O(T^{\theta_2-1+\epsilon}),
\]
and thus $\text{Var} \tilde{\eta}_1 = 1 + O(T^{\theta_2-1+\epsilon})$.

Further, let us estimate the distances between $\tilde{\eta}_0$ and $\tilde{\eta}_1$. For Lévy—Prokhorov distance observe that for any Borel set $A$
\[
\frac{\mathbb{P}(\tilde{\eta}_0 \in A)}{\mathbb{P}(\tilde{\eta}_1 \in A)} = 1 + O(T^{\theta_2-1+\epsilon}).
\]
Hence $|\mathbb{P}(\tilde{\eta}_0 \in A) - \mathbb{P}(\tilde{\eta}_1 \in A)| = O(T^{\theta_2-1+\epsilon})$, thus $d_{LP}(\tilde{\eta}_0, \tilde{\eta}_1) = O(T^{\theta_2-1+\epsilon})$. For Kantorovich—Rubinstein distance we can (in mass transportation interpretation) move all excessive mass in $\text{Law}(\tilde{\eta}_0)$ to the origin and then back to the points where $\text{Law}(\tilde{\eta}_1)$ has excessive mass. The total work for this transportation is
\[
\int |x| |\rho(x) - 1| dF_{\tilde{\eta}_0}(x) = O(T^{\theta_2-1+\epsilon}).
\]
It remains to estimate the distances $d(\tilde{\eta}_0, \tilde{\eta}_1)$. We use the identity
\[
\tilde{\eta}_1 = \frac{\tilde{\eta}_1 - \mathbb{E}\tilde{\eta}_1}{\sqrt{\text{Var} \tilde{\eta}_1}}
\]
and the estimates
\[
|\mathbb{E}\tilde{\eta}_1| = O(T^{\theta_2-1+\epsilon}), \quad \text{Var} \tilde{\eta}_1 = 1 + O(T^{\theta_2-1+\epsilon})
\]
obtained above. This yields
\begin{equation}
\begin{aligned}
&d_{KR}(\tilde{\eta}_1, \hat{\eta}_1) = O(T^{\theta_2-1+\epsilon}), \\
d_{LP}(\tilde{\eta}_1, \hat{\eta}_1) = O(T^{\frac{2}{3}(\theta_2-1)+\epsilon}).
\end{aligned}
\end{equation}

\textbf{Lemma 4.4.} For any \(\beta > 0\) we have
\begin{equation}
\begin{aligned}
&d_{KR}(\tilde{\eta}_1, \hat{\eta}_2) = O(n^{\theta_2-\theta_2+\beta}), \\
&d_{LP}(\tilde{\eta}_1, \hat{\eta}_2) = O(n^{\frac{2}{3}(\theta_2-\theta_2)+\beta}).
\end{aligned}
\end{equation}

\textbf{Proof.} As in the previous lemma, consider
\[ \tilde{\eta}_2 = \frac{\eta_2 - \mathbb{E}\eta_1}{\sqrt{\text{Var } \eta_1}}. \]
Then Corollary 3.7 yields
\[ |\tilde{\eta}_2 - \hat{\eta}_1| \leq \frac{|\eta_1 - \eta_2|}{\sqrt{\text{Var } \eta_1}} = O(n^{\theta_2-\theta_2+\delta}), \]

\text{hence}
\[ d_{KR,LP}(\tilde{\eta}_2, \hat{\eta}_1) = O(n^{\theta_2-\theta_2+\delta}), \]
since for \(d_{KR}\) we move every piece of mass by the distance at most \(\alpha = \sup |\tilde{\eta}_2 - \hat{\eta}_1|\), and for \(d_{LP}\) we see that \(\tilde{\eta}_1 \in A\) implies \(\tilde{\eta}_2 \in A^\alpha\) and vice versa. Further,
\[ |\mathbb{E}\tilde{\eta}_2| \leq |\tilde{\eta}_1| + \mathbb{E}|\tilde{\eta}_2 - \tilde{\eta}_1| = O(n^{\theta_2-\theta_2+\delta}), \]
\[ |\mathbb{E}\tilde{\eta}_2^2 - \mathbb{E}\tilde{\eta}_1^2| \leq 2\mathbb{E}|\tilde{\eta}_1| \cdot \alpha + \alpha^2 = O(n^{\theta_2-\theta_2+\delta}), \]
whence \(\text{Var } \tilde{\eta}_2 = 1 + O(n^{\theta_2-\theta_2+\delta})\). As in the previous lemma, these estimates yields the same formulas as (12) for \(d(\tilde{\eta}_2, \hat{\eta}_2)\), and this concludes the proof. \(\square\)

Theorem 3.1 now follows from Bufetov’s limit theorem and Lemmas 4.3 and 4.4.

\textbf{Acknowledgments.} The work was partially supported by RFBR–CNRS grant No. 18-51-15010 and RFBR grant No. 18-31-20031. The author deeply thanks Alexander Bufetov for fruitful discussions on this matter.

\textbf{References}

[1] A. Avila and M. Viana, Simplicity of Lyapunov spectra: proof of the Zorich-Kontsevich conjecture, \textit{Acta Math.}, \textbf{198} (2007), 1–56.
[2] A. Bufetov, Limit Theorems for Translation Flows, \textit{Ann. of Math. (2)}, \textbf{179}:2 (2014), 431499.
[3] J. Chaika, A. Eskin, Every flat surface is Birkhoff and Oseledets generic in almost every direction. \textit{J. Modern Dynam.}, 2015, \textbf{9}, 1-23.
[4] D. Dolgopyat, O. Sarig, Temporal Distributional Limit Theorems for Dynamical Systems. \textit{J. Stat. Phys.} \textbf{166} (2017) 680–713.
[5] G. Forni, Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. \textit{Ann. of Math.} \textbf{155} (2002), 1–103.
[6] M. Kontsevich and A. Zorich, Lyapunov exponents and Hodge theory, 2007. \texttt{arXiv:hep-th/9701164}
[7] H. Masur, Interval exchange transformations and measured foliations, \textit{Ann. of Math.} \textbf{115} (1982) 169-200.
[8] W. A. Veech, Gauss measures for transformations on the space of interval exchange maps, *Ann. of Math.* 115 (1982) 201-242.

[9] A. Zorich, Flat surfaces, Frontiers in number theory, physics, and geometry. I, Springer, Berlin, 2006, pp. 437-583.

**Alexey Klimenko**
Steklov Mathematical Institute of Russian Academy of Sciences,
Gubkina str. 8, 119991, Moscow, Russia;
and National Research University Higher School of Economics,
Usacheva str. 6, 119048, Moscow, Russia
klimenko@mi.ras.ru