PDMI PREPRINT — 19/2002

A Poisson limit theorem for toral automorphisms

Manfred DENKER, Mikhail GORDIN, and Anastasya SHAROVA

Institut für Mathematische Stochastik
der Universität Göttingen
Russian Academy of Sciences
Steklov Mathematical Institute
St.Petersburg Department

November 4, 2002

ABSTRACT

We introduce a new method of proving Poisson limit laws in the theory of dynamical systems, which is based on the Chen-Stein method ([7, 20]) combined with the analysis of the homoclinic Laplace operator in [11] and some other homoclinic considerations. This is accomplished for hyperbolic toral automorphism T and the normalized Haar measure P. Let $(G_n)_{n\geq 0}$ be a sequence of measurable sets with no periodic points among its accumulation points and such that $P(G_n) \to 0$ as $n \to \infty$, and let $(s(n))_{n>0}$ be a sequence of positive integers such that $\lim_{n\to\infty} s(n)P(G_n) = \lambda$ for some $\lambda > 0$. Then, under some additional assumptions about $(G_n)_{n\geq 0}$, we prove that for every integer $k \geq 0$

$$P(\sum_{i=1}^{s(n)} 1_{G_n} \circ T^{i-1} = k) \to \lambda^k \exp(-\lambda)/k!$$

as $n \to \infty$ Of independent iterest is an upper mixing-type estimate which is one of our main tools.

Typeset by $\mathcal{AMS}$-TEX

1. Introduction

Let $(X, \mathcal{F}, P, T)$ be a dynamical system where X is a compact metrizable space with the Borel σ-field $\mathcal{F}$, P is a probability measure on $\mathcal{F}$ and T is an invertible continuous P-preserving transformation of X. Let $(G_n)_{n\geq 1}$ be a sequence of measurable subsets of X and $(s(n))_{n\in\mathbb{Z}_+}$ be a sequence of integers such that $s(n) \to \infty$ and for some $\lambda > 0$ $s(n)P(G_n) \to \lambda$ as $n \to \infty$. In this note we are interested in the distributional behaviour of the number of visits to G_n in the time span up to $s(n)$. There are various results known in the literature where under appropriate assumptions about the transformation T and the sequence $(G_n)_{n\geq 1}$ the limit distribution is a Poisson law.

A Poisson limit law is well known for arrays of independent Bernoulli random variables and variants of such processes. Some results are also known for dependent random variables (see e.g. [4J]). The first result for dynamical systems seems to be contained in Pitskel's paper [18] where finite state Markov chains and two-dimensional toral automorphisms are considered. The result for Markov chains is proved in [18] by the method of moments (using a general result in [19]), and the limit law for toral automorphisms is deduced from it via a representation by symbolic dynamics ([2]). The symbolic part was extended to general Gibbs measures in [9] and convergence in finite dimensional distributions.

Axiom A diffeomorphisms and some class of Gibbs measures are considered in Hirata's paper [14]. It contains the multidimensional Poisson theorem for the joint distribution of the number of visits for several successive time intervals. The method of proof here is symbolic dynamics and perturbation theory of transfer operators.

A Poisson limit theorem for maps of the interval has been first obtained in [8], later also for non-hyperbolic transformations in [16], while in Haydn's paper [13] the analogous result is proved for rational maps of the Riemann sphere, even in the case when critical points belong to the Julia set.

We also mention a recent paper ([6]), where some earlier references about Poisson limit theorems for dynamical systems can be found.

In all these papers the sequence of shrinking sets G_n is assumed to approach a typical point of the distribution. On the other hand, examples in [18] and [14] show that one cannot expect the Poisson limit law for sequences of sets G_n shrinking to a periodic point. Dolgopyat established convergence to the Poisson law for a sequence of balls shrinking to an arbitrary aperiodic point in [10]. This problem has been stated in [15]. We are able to extend this result to more general families of shrinking sets in case when T is a hyperbolic automorphism of the d-dimensional torus $\mathbb{T}^d$ with normalized Haar measure P. Moreover, we only need the weaker condition that the set of accummulation point of the sequence G_n is contained in the set of aperiodic points. The main achievement of the present note, however, is the new method of proof for such results.

We briefly explain this method, which is completely different from the method of moments, perturbations or other methods used before in the context of dynamical systems. All these other methods are based on some form of symbolic dynamics, while ours does not use such a representation at all.

The starting point is a difference equation (due to Chen [7], see also Arratia et al. [3]) characterizing the convergence to a Poisson law (Proposition 2.1). The sufficient condition given by this equation will be reduced to

$$\sum_{i=1}^{s(n)} E\mathbf{1}_{T^{-i+1}G_n}\psi\left(\dot{W}_n^{(i)}\right) - P(G_n)E\psi\left(\dot{W}_n^{(i)}\right) \to 0 \tag{1}$$

as $n \to \infty$, where ψ is any bounded function and where

$$\dot{W}_n^{(i)} = \sum_{j:1 \leq j \leq s(n); j \neq i} \mathbf{1}_{T^{-j+1}G_n}. \tag{2}$$

This relation is expected to hold, if for every i ($1 \leq i \leq s(n)$) some form of weak dependence between the random variables $\mathbf{1}_{T^{-i}G_n}$ and $\psi(\dot{W}_n^{(i)})$ can be shown. This weak dependence together with the "small size" of the sets G_n should imply (1). The relation (1) cannot always be true, because otherwise it implies convergence to a Poisson law, which contradicts the counterexamples mentioned above. We prove (1) for every suitable sequence $(G_n)_{n \geq 1}$ shrinking to an arbitrary aperiodic point by splitting the problem into two. To begin with, enlarge the number of missing iterates in the sum $\dot{W}_n^{(i)}$ by

$$\dot{W}_n^{(i,m(n))} = \sum_{j:1 \leq j \leq s(n),\ m(n) \leq |i-j|} \mathbf{1}_{T^{-j+1}G_n}, \tag{3}$$

where $(m(n))_{n \geq 1}$ is a sequence tending to infinity. The first problem (section 3.2) is to show (1) when $\dot{W}_n^{(i)}$ is replaced by $\dot{W}_n^{(i,m(n))}$. This step may be considered as an asymptotic decorrelation property, and requires that $m(n)$ grows rapidly enough. The proof is based on the homoclinic Laplace operator introduced in [11]. The second problem is to prove that the replacement in (1) does not affect the asymptotic relation. This will be established counting homoclinic points. It follows from our estimates that this equivalence holds true even in case $(m(n))_{n \geq 1}$ grows rather rapidly. Thus one has enough freedom to choose $(m(n))_{n \geq 1}$ so that both requirements are satisfied.

2. Poisson limit laws

There are several methods to prove Poisson limit laws in the context of dynamical systems (see e.g. [6]). Here we add another method based on the following result of Chen ([7], see also Barbour et al. [4]).

Proposition 2.1. *Let $\lambda > 0$ and $(X_n)_{n \geq 1}$ be a sequence of random variables with values in $\mathbb{Z}_+$. In order that $(X_n)_{n \geq 1}$ converges in distribution to the Poisson law with parameter λ, it is necessary and sufficient that, for every bounded function $\varphi : \mathbb{Z}_+ \to \mathbb{R}$,*

$$\lim_{n \to \infty} EX_n\varphi(X_n) - \lambda E\varphi(X_n + 1) = 0. \tag{4}$$

Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ denote the d-dimensional torus with distance function $\rho(\cdot,\cdot)$ induced by the standard Euclidean metric on $\mathbb{R}^d$. We denote by P the normalized Haar measure on $\mathbb{T}^d$, and by $B(z, r)$ the ball of radius r around z.

Before formulating the result of the paper we need to specify the class of sequences of sets $(G_n)_{n \geq 1}$ shrinking to an aperiodic point.

Definition 2.2. *Let $\alpha \geq 0, \beta \geq 0$ and $K > 0$ be real numbers. A measurable set $S \subset \mathbb{T}^d$ is said to belong to the class $H_P(\alpha, \beta, K)$, if for every $g \in \mathbb{T}^d$ the set S satisfies the inequality*

$$P((S \setminus (S+g)) \cup ((S+g) \setminus S)) \leq K\rho^\alpha(g, 0)P^\beta(S).$$

Theorem 2.3. *Let T be a hyperbolic automorphism of the d-dimensional torus $\mathbb{T}^d$ equipped with the normalized Haar measure P. Assume that for some real number $\lambda > 0$, some sequence $(s(n))_{n\geq 1}$ of positive integers and some sequence $(r_n)_{n\geq 1}$ of positive real numbers the following conditions are satisfied:*

i) $s(n) \to \infty$ *as* $n \to \infty$;

ii) $s(n)P(G_n) \to \lambda$ *as* $n \to \infty$;

iii) $\operatorname{diam} G_n < 2r_n$ *for every* $n \geq 1$;

iv) $P(B(z_n, r_n)) \leq \xi P(G_n)$ *for some* $\xi > 0$ *and all* $n \geq 1$, *where* $G_n \subset B(z_n, r_n)$;

v) $G_n \in H(\alpha, \beta, K)$, *where* $\alpha > 0$, $\beta \geq 0$ *and* $K > 0$ *denote some constants, independent of* $n \geq 1$.

vi) *The set of accummulation points of the sequence* $(G_n)_{n\geq 1}$ *is contained in the set of aperiodic points.*

Then for every integer $k \geq 0$

$$\lim_{n\to\infty} P\left(\sum_{i=1}^{s(n)} \mathbf{1}_{G_n} \circ T^{i-1} = k\right) = \lambda^k \exp(-\lambda)/k!.$$

The proof of Theorem 2.3 is contained in the next section.

Remark 2.4 We briefly mention some examples of sequences $(G_n)_{n\geq 1}$ satisfying the assumptions in the theorem.

(1.) Balls $(B_n(z, r_n))_{n\geq 1}$ satisfy the conditions of the theorem, whenever the sequence $(s(n)r_n^d)_{n\geq 1}$ converges to a finite positive number. In this case we can take $\alpha = 1$ and $\beta = (d-1)/d$.

(2.) Slightly more generally, we can take a bounded open set satisfying v) with $\alpha = 1$, $\beta = (d-1)/d$ and use similarities to obtain the sequence $(G_n)_{n\geq 1}$ of suitable measure. It follows from the Steiner formula, that every bounded convex open set satisfies v).

(3.) We may extend the class of bounded convex open sequences replacing the similarities in (2.) by the assumption that a certain Minkowski-type functional (see (11) below for its definition) evaluated at the sequence $(G_n)_{n\geq 1}$, is uniformly bounded from above and below.

(4.) Finally, it suffices to assume that every G_n is Lipschitz homeomorphic to a ball of equal measure, so that, uniformly in $n \in \mathbb{Z}_+$, the Lipschitz constants are bounded from above.

3. Proof of Theorem 2.3

Throughout this section we use the notation as in the theorem. For the sequence of sets G_n as in Theorem 2.3 define $f_n^{(i)} = \mathbf{1}_{T^{-i+1}G_n}$, $f_n = f_n^{(1)} = \mathbf{1}_{G_n}$ and $p_n = P(G_n)$. The partial sums will be denoted by

$$W_n = \sum_{j=1}^{s(n)} f_n^{(i)} = \sum_{j=0}^{n-1} \mathbf{1}_{G_n} \circ T^j.$$

The random variables $\dot{W}^{(\cdot)}_{\cdot}$ introduced in (2) and (3) will be called *punctured sums*. In order to prove Theorem 2.3 we shall verify the assumption (4) of Proposition 2.1 for $X_n = W_n$ and for every fixed bounded measurable function $\varphi : \mathbb{Z}_+ \to \mathbb{R}$. In the sequel we use C to denote a generic constant only depending on the automorphism T. All other constants will be given explicitly.

3.1. Reduction to punctured sums.

In this subsection we prove

Proposition 3.1. *If $M := \|\varphi\|_\infty$ and if $m(n)$ is any sequence of positive integers, then*

$$|EW_n\varphi(W_n) - \lambda E\varphi(W_n+1)| \le \left|\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})\varphi(\dot W_n^{(i,m(n))}+1)\big)\right|$$

$$+2M\sum_{i=1}^{s(n)}\sum_{j:1\le j\le m(n)} E(f_n^{(i)}f_n^{(j)}) + 4Mm(n)s(n)p_n^2 + M|s(n)p_n - \lambda|.$$

Proof. As the functions $f_n^{(i)}, i = 1,\dots,s(n)$, are indicator functions, we see that

$$\begin{aligned} E(W_n\varphi(W_n)) &= \sum_{i=1}^{s(n)} Ef_n^{(i)}\varphi(W_n) = \sum_{i=1}^{s(n)} Ef_n^{(i)}\varphi(\dot W_n^{(i)} + f_n^{(i)}) \\ &= \sum_{i=1}^{s(n)} Ef_n^{(i)}\varphi(\dot W_n^{(i)}+1), \end{aligned}$$

hence we can write

$$E[W_n\varphi(W_n) - \lambda\varphi(W_n+1)] = \Sigma_{1,n} + \Sigma_{2,n} + \Sigma_{3,n},$$

where

$$\Sigma_{1,n} = \sum_{i=1}^{s(n)} Ef_n^{(i)}\varphi(\dot W_n^{(i)}+1) - \sum_{i=1}^{s(n)} Ef_n^{(i)}E\varphi(\dot W_n^{(i)}+1),$$

$$\Sigma_{2,n} = \sum_{i=1}^{s(n)} Ef_n^{(i)}E\varphi(\dot W_n^{(i)}+1) - \sum_{i=1}^{s(n)} Ef_n^{(i)}E\varphi(W_n+1)$$

and

$$\Sigma_{3,n} = \sum_{i=1}^{s(n)} \Big(Ef_n^{(i)}E\varphi(W_n+1) - \lambda E\varphi(W_n+1)\Big).$$

We first estimate $\Sigma_{2,n}$ and $\Sigma_{3,n}$.
Since

$$\varphi(\dot W_n^{(i)}+1) - \varphi(W_n+1) = f_n^{(i)}\big(\varphi(\dot W_n^{(i)}+1) - \varphi(W_n+1)\big),$$

it follows that

$$\begin{aligned} |\Sigma_{2,n}| &= \left|\sum_{i=1}^{s(n)} Ef_n^{(i)}E\big(\varphi(\dot W_n^{(i)}+1) - \varphi(W_n+1)\big)\right| \\ &= p_n\left|\sum_{i=1}^{s(n)} Ef_n^{(i)}\big(\varphi(\dot W_n^{(i)}+1) - \varphi(W_n+1)\big)\right| \\ &\le 2Mp_n\sum_{i=1}^{s(n)} Ef_n^{(i)} = 2Ms(n)p_n^2. \end{aligned}$$

It is also immediate to see that

$$|\Sigma_{3,n}| = |E\varphi(W_n+1)|\left|\sum_{i=1}^{s(n)} Ef_n^{(i)} - \lambda\right| \le M|s(n)p_n - \lambda|.$$

It remains to estimate $\Sigma_{1,n}$. Letting $\psi(\cdot) = \varphi(\cdot+1)$, we can rewrite $\Sigma_{1,n}$ as follows:

$$\begin{aligned}
&\sum_{i=1}^{s(n)} \Big(Ef_n^{(i)}\psi(\dot W_n^{(i)}) - Ef_n^{(i)}E\psi(\dot W_n^{(i)})\Big)\\
=&\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})\psi(\dot W_n^{(i)})\big)\\
=&\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})\psi(\dot W_n^{(i,m(n))})\big)\\
&+\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})(\psi(\dot W_n^{(i)}) - \psi(\dot W_n^{(i,m(n))}))\big)\\
=&\Sigma_{1,n}^{(1)} + \Sigma_{1,n}^{(2)}.
\end{aligned}$$

Now

$$\begin{aligned}
|\Sigma_{1,n}^{(2)}| \le &\sum_{i=1}^{s(n)} E\big(f_n^{(i)}|\psi(\dot W_n^{(i)}) - \psi(\dot W_n^{(i,m(n))}|\big)\\
&+\sum_{i=1}^{s(n)} Ef_n^{(i)}E|\psi(\dot W_n^{(i)}) - \psi(\dot W_n^{(i,m(n))})|\\
\le& 2M\sum_{i=1}^{s(n)} E\big(f_n^{(i)}\mathbf{1}_{\dot W_n^{(i)} \ne \dot W_n^{(i,m(n))}}\big) + 2M\sum_{i=1}^{s(n)} Ef_n^{(i)}P(\dot W_n^{(i)} \ne \dot W_n^{(i,m(n))})\\
\le& 2M\sum_{i=1}^{s(n)} E\big(f_n^{(i)}\sum_{j:1\le|j-i|<m(n)} f_n^{(j)}\big) + 2M\sum_{i=1}^{s(n)}\sum_{j:1\le|j-i|<m(n)} Ef_n^{(i)}Ef_n^{(j)}\\
=& 2M\sum_{i=1}^{s(n)}\sum_{j:1\le|j-i|\le m(n)} E(f_n^{(i)}f_n^{(j)}) + 2M(2m(n)-1)s(n)p_n^2,
\end{aligned}$$

and we obtain the inequality

$$\begin{aligned}
|\Sigma_{1,n}| \le &\left|\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})\psi(\dot W_n^{(i,m(n))})\big)\right|\\
&+ 2M\sum_{i=1}^{s(n)}\sum_{j:1\le|j-i|\le m(n)} E(f_n^{(i)}f_n^{(j)})\\
&+ 2M(2m(n)-1)s(n)p_n^2.
\end{aligned}$$

This finishes the proof of the proposition. □

From now on we assume that the sequence $(m(n))_{n\geq 1}$ tends to ∞ at a rate $o(s(n))$. Some further conditions on the sequence $(m(n))_{n\geq 1}$ will be imposed below.

This condition and assumption ii) in Theorem 2.3 imply that $\lim_{n\to\infty} m(n)s(n)p_n^2 = 0$ and $\lim_{n\to\infty} s(n)p_n = \lambda$, hence (4) holds if the two summands

$$\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})\phi(\dot{W}_n^{(i,m(n))} + 1)\big) \tag{5}$$

and

$$\sum_{i=1}^{s(n)} \sum_{j:1\leq j\leq m(n)} E(f_n^{(i)} f_n^{(j)}) \tag{6}$$

tend to zero.

3.2. Punctured sums.

As explained in the introduction the second step in the proof of Theorem 2.3 is the estimation of enlarged punctured sums. This is the content of the proposition proved in this subsection.

Let L_2 be the space of all complex-valued functions on the torus $\mathbb{T}^d$, square integrable with respect to the Haar measure P, and let L_2^0 consists of all functions from L_2 with vanishing integral with respect to P. We denote by $\|\cdot\|_2$, $(\cdot,\cdot)$ and I the norm, the inner product and the identity operator in L_2, respectively. For an operator S defined on L_2 let S^* denote its conjugate, and let U denotes the unitary operator defined by $Uf = f\circ T, f\in L_2$. Set

$$\hat{f}_n^{(i)} = f_n^{(i)} - Ef_n^{(i)}.$$

First note that the group $\Gamma = \{\gamma \in \mathbb{T}^d : T^n(\gamma) \to 0 \text{ as } |n|\to\infty\}$ can be described as the intersection of the stable subgroup $\Gamma_s = \{\gamma\in\mathbb{T}^d : T^n(\gamma)\to 0 \text{ as } n\to\infty\}$ and the unstable subgroup $\Gamma_u = \{\gamma\in\mathbb{T}^d : T^n(\gamma)\to 0 \text{ as } n\to -\infty\}$ of $\mathbb{T}^d$. It is known ([10]) that for a hyperbolic automorphism T Γ is a dense T-invariant subgroup of $\mathbb{T}^d$. Since Γ is an abelian group, it is a free group on d generators. Fix generators $\gamma_1,\ldots,\gamma_d$ and observe, that there exist constants $A>0, \kappa>0$ such that for every $p\in\mathbb{Z}, 1\leq l\leq d$ we have

$$\rho(T^p\gamma_l, 0) \leq A\exp(-\kappa|p|). \tag{7}$$

Proposition 3.2. *Let κ be as in (7). If $\sigma > 2(1-\beta)(\alpha\kappa)^{-1}$ and if $(m(n))_{n\geq 1}$ satisfies $m(n)\geq [\sigma\log s(n)] + 1$ for $n\in\mathbb{Z}_+$, then*

$$\sum_{i=1}^{s(n)} E\big((f_n^{(i)} - Ef_n^{(i)})\psi(\dot{W}_n^{(i,m(n))})\big) \to 0 \qquad \text{as } n\to\infty.$$

In the remaining part of this subsection we prove the proposition.

For any $g\in\mathbb{T}^d$ we denote by H_g the unitary operator defined by the translation by g: $(H_g f)(\cdot) = f(g+\cdot)$. We have $U^n H_g U^{-n} \to I$ as $|n|\to\infty$ (in the strong operator topology) if and only if $g\in\Gamma$ (note that $U^n H_g U^{-n} = H_{T^n g}$). Any operator of the form H_γ with $\gamma\in\Gamma$ is called a *homoclinic translation operator* or simply a *homoclinic operator*. Let us denote by $H_{p,l}$

the homoclinic operator corresponding to $T^p\gamma_l$, $p\in\mathbb{Z}$, $l\in\{1,...,d\}$, so that $U^nH_{p,l}U^{-n}=H_{p+n,l}$ for every $p,n\in\mathbb{Z}$, $l\in\{1,...,d\}$. Then we set

$$\Delta=\sum_{l=1}^{d}\sum_{p\in\mathbb{Z}}(I-H_{p,l}^*)(I-H_{p,l}).$$

The operator Δ is called the *homoclinic Laplace operator* in [11]. More precisely, the above expression defines an unbounded symmetric operator on a dense subset of L_2, which commutes with U. As has been established in [11], there exists a constant $c>0$ such that for any f from a dense subset of L_2^0 we have

$$(\Delta f,f)\geq c\|f\|_2^2. \tag{8}$$

By this property Δ is Friedrichs closable, and from now on Δ denotes this closure.

Notice, that $T^n\gamma\to 0$ exponentially fast as $|n|\to\infty$, if $\gamma\in\Gamma$; hence the rate of convergence in

$$\|(I-H_{T^n\gamma})f\|_2\to 0 \text{ as } |n|\to\infty \tag{9}$$

(which holds for every $f\in L_2$) can be made specific under mild assumptions on f. For instance, if f is Hölder continuous in L_2-sense, we have exponential rate in (9) (in particular, this is why the operator Δ is densely defined). Given such a function f, and $p,l\in\mathbb{Z}$ $(1\leq l\leq d)$ define

$$r_l(f,p)=\|(I-H_{p,l})f\|_2$$

and

$$w(f,p)=\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}r_l(f,q)r_l(f,q+p).$$

Lemma 3.3. *With the above notation we have*

$$|E\hat{f}_n^{(i)}\psi(\dot{W}_n^{(i,m(n))})|\leq 4c^{-1}M\sum_{|p|\geq m(n)}w(f_n,p).$$

Proof. In view of the property (8) the operator Δ has a bounded right inverse Δ^{-1} on L_2^0 whose norm does not exceed c^{-1}. Note that for every $g\in\mathbb{T}$ and ψ with $|\psi(\cdot)|\leq M$ we may write

$$\begin{aligned}
&\|(I-H_g)\psi(\dot{W}_n^{(i,m(n))})\|_2\\
\leq&2M(P(H_g\dot{W}_n^{(i,m(n))}\neq\dot{W}_n^{(i,m(n))}))^{1/2}\\
\leq&2M\left(\sum_{j:|i-j|\geq m(n);1\leq j\leq n}P(H_gf_n^{(j)}\neq f_n^{(j)})\right)^{1/2}\\
=&2M\sum_{j:|i-j|\geq m(n);1\leq j\leq n}\|(I-H_g)f_n^{(j)}\|_2.
\end{aligned}$$

Therefore we have

$$
\begin{aligned}
&|E(\hat{f}_n^{(i)}\psi(\dot{W}_n^{(i,m(n))}))|\\
=&|(\hat{f}_n^{(i)},\psi(\dot{W}_n^{(i,m(n))}))| = |\big(\Delta\Delta^{-1}\hat{f}_n^{(i)},\psi(\dot{W}_n^{(i,m(n))})\big)|\\
\le&\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}|\big((I-H_{q,l}^*)(I-H_{q,l})\Delta^{-1}\hat{f}_n^{(i)},\psi(\dot{W}_n^{(i,m(n))})\big)|\\
=&\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}|\big((I-H_{q,l})\Delta^{-1}\hat{f}_n^{(i)},(I-H_{q,l})\psi(\dot{W}_n^{(i,m(n))})\big)|\\
=&\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}|\big(\Delta^{-1}(I-H_{q,l})\hat{f}_n^{(i)},(I-H_{q,l})\psi(\dot{W}_n^{(i,m(n))})\big)|\\
\le&\frac{1}{c}\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}\|(I-H_{q,l})\hat{f}_n^{(i)}\|_2\|(I-H_{q,l})\psi(\dot{W}_n^{(i,m(n))})\|_2\\
\le&\frac{2}{c}M\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}\sum_{j:|i-j|\ge m(n)}\|(I-H_{q,l})f_n^{(i)}\|_2\|(I-H_{q,l})f_n^{(j)}\|_2\\
=&\frac{2}{c}M\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}\sum_{j:|i-j|\ge m(n)}\|(I-H_{q-i+1,l})f_n\|_2\|(I-H_{q-j+1,l})f_n\|_2\\
=&\frac{2}{c}M\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}\sum_{|p|\ge m(n)}\|(I-H_{p+q,l})f_n\|_2\|(I-H_{q,l})f_n\|_2\\
=&\frac{2}{c}M\sum_{l=1}^{d}\sum_{q\in\mathbb{Z}}\sum_{|p|\ge m(n)}r_l(f_n,p+q)r_l(f_n,q)\\
=&\frac{2}{c}M\sum_{|p|\ge m(n)}w(f_n,p).
\end{aligned}
$$

□

Recall that $G_n\in H_P(\alpha,\beta,K)$ for some constants α, β and K by assumption *v)* of Theorem 2.3.

Lemma 3.4. *There exists a constant $C_0=C(T,\alpha,\beta,K)$ such that*

$$\left|\sum_{i=1}^{s(n)}E\big(\widehat{f}_n^{(i)}\psi(\dot{W}_n^{(i,m(n))})\big)\right|\le C_0Ms(n)m(n)e^{(-\alpha\kappa m(n)/2)}P^{\beta}(G_n).$$

Proof. Since $G_n\in H_P(\alpha,\beta,K)$, we have for $f_n=f_n^{(1)}$ and $|p|$ large enough that

$$
\begin{aligned}
&r_l(f_n,p)=\|(I-H_{p,l})f_n\|_2\\
=&P^{1/2}((G_n\setminus(G_n+T^p\gamma_l))\cup((G_n+T^p\gamma_l)\setminus G_n))\\
\le&K^{1/2}\rho^{\alpha/2}(T^p\gamma_l,0)P^{\beta/2}(G_n)\le K^{1/2}A^{\alpha/2}\mathrm{e}^{(-(\alpha\kappa|p|/2)}P^{\beta/2}(G_n)
\end{aligned}
$$

and that

$$
\begin{aligned}
w(f_n,p) =& \sum_{l=1}^{d}\sum_{q\in\mathbb{Z}} r_l(f_n,q) r_l(f_n,p+q)\\
\leq & dKA^{\alpha}P^{\beta}(G_n)\sum_{q\in\mathbb{Z}}\exp\left(-\alpha\kappa(|q|+|p+q|)/2\right)\\
\leq & 2dKA^{\alpha}P^{\beta}(G_n)\sum_{p\geq m(n)}\left(2\sum_{q\geq 0}\mathrm{e}^{-\frac{\alpha\kappa}{2}(2q+|p|)}+\sum_{0<q<|p|}\mathrm{e}^{-\frac{\alpha\kappa|p|}{2}}\right.\\
\leq & 2dKA^{\alpha}P^{\beta}(G_n)\left(|p|-2+\frac{2}{1-\mathrm{e}^{-\alpha\kappa}}\right)\mathrm{e}^{-(\alpha\kappa|p|)/2}.
\end{aligned}
$$

Then, by Lemma 3.3

$$
\begin{aligned}
|E\hat{f}_n^{(i)}\psi(\dot{W}_n^{(i,m(n))})| \leq & 4c^{-1}M\sum_{|p|\geq m(n)} w(f_n,p)\\
\leq & \frac{8dKM}{c}A^{\frac{\alpha}{2}}P^{\beta}(G_n)\sum_{p\geq m(n)}\left(p-2+\frac{2}{1-\mathrm{e}^{-\alpha\kappa}}\right)\mathrm{e}^{-\frac{\alpha\kappa|p|}{2}}\\
\leq & C_0 M m(n)\mathrm{e}^{-\alpha\kappa m(n)/2}P^{\beta}(G_n),
\end{aligned}
$$

where $C_0 = C(T,\alpha,\beta,K)$ denotes some constant.

The lemma follows by summation over $i = 1,...,s(n)$. □

Proof of Proposition 3.2. In view of the assumptions in Theorem 2.3 $P(G_n) = p_n = O(s(n)^{-1})$ as $n\to\infty$. By the choice of $m(n)$ it follows that

$$
\begin{aligned}
& s(n)m(n)\exp\left(-\alpha\kappa m(n)/2\right)P^{\beta}(G_n)\\
=& O\big((\log s(n))\exp\left[(1-\beta-(\sigma\alpha\kappa)/2)\log s(n)\right]\big) = o(1),
\end{aligned}
$$

and the proposition follows from Lemma 3.4. □

3.3 Estimating the puncturing effect.

In this subsection we show that the second summand (6) in Proposition 3.1 converges to zero, if $m(n)$ tends to infinity at a rate $o(s(n))$. This implies that the punctured and enlarged punctured sums are stochastically equivalent. We shall prove this relation by embedding the sets G_n into parallelograms R_n in the sense of [5] or [1] (the precise definition of a parallelogram can be found below as well). In order to explain the statement and the proof of our main proposition we need more details about hyperbolic toral automorphisms.

Let $T:\mathbb{T}^d\to\mathbb{T}^d$ be an algebraic automorphism of $\mathbb{T}^d$. The covering map of T is an invertible linear map $\widetilde{T}:\mathbb{R}^d\to\mathbb{R}^d$, leaving the lattice $\mathbb{Z}^d$ invariant, hence $\pi\ \widetilde{T} = T\ \pi$, where $\pi:\mathbb{R}^d\to\mathbb{T}^d$ denotes the canonical map onto the quotient group. The spectrum spec$(\widetilde{T})$ of the operator $\widetilde{T}$ splits into three disjoint components

$$
\mathrm{spec}(\widetilde{T}) = \sigma_s(\widetilde{T})\cup\sigma_u(\widetilde{T})\cup\sigma_n(\widetilde{T}),
$$

located outside, inside and on the unit circle $\{z:|z|=1\}\subseteq\mathbb{C}$, respectively. Since we assume that T is a hyperbolic automorphism, $\sigma_n(\widetilde{T}) = \emptyset$. The decomposition of the spectrum spec$(\widetilde{T}) =$

$\sigma_s(\widetilde{T}) \cup \sigma_u(\widetilde{T})$ induces a $\widetilde{T}$-invariant decomposition $\mathbb{R}^d = L_s \oplus L_u$ into the direct sum of its *stable* and *unstable* subspaces L_s and L_u. Denote by d_s and d_u their dimensions (so that $d = d_s + d_u$), and let p_s and p_u be the corresponding projections (with kernels L_u and L_s, respectively). Set $R_s^* = \max_{\lambda \in \sigma_s} |\lambda|, r_s^* = \min_{\lambda \in \sigma_s} |\lambda|$, $R_u^* = \max_{\lambda \in \sigma_u} |\lambda|, r_u^* = \min_{\lambda \in \sigma_u} |\lambda|$, and fix some $Q \in (\max(R_u^*, {r_s^*}^{-1}), \infty)$ and $q \in (1, \min(r_u^*, {R_s^*}^{-1})$. Note that $Q \geq q > 1$, and that for some $A \geq 1$ and every $n \in \mathbb{Z}_+$ we have the following estimates

$$\|\widetilde{T}^n|_{L_s}\| \leq Aq^{-n}, \quad \|\widetilde{T}^{-n}|_{L_u}\| \leq Aq^{-n}, \tag{10}$$

and

$$\|\widetilde{T}^{-n}|_{L_s}\| \leq AQ^n, \quad \|\widetilde{T}^n|_{L_u}\| \leq AQ^n,$$

where the operator norm $\|\cdot\|$ is derived from the Euclidean norm in $L_s \subset \mathbb{R}^d$ and $L_u \subset \mathbb{R}^d$, respectively. The map π is injective when restricted to either L_s or L_u, and we have $\pi(L_s) = \Gamma_s \subset \mathbb{T}^d$ and $\pi(L_u) = \Gamma_u \subset \mathbb{T}^d$. For $t \in \mathbb{T}^d$ we set $\Gamma_s(t) = \Gamma_s + t$ and $\Gamma_u(t) = \Gamma_u + t$, so that $\Gamma_s(t)$ and $\Gamma_u(t)$ are the stable and the unstable leaves (or cosets) of the point t. The restriction of the standard Riemannian metric of $\mathbb{T}^d$ to every $\Gamma_s(t)$ defines the inner distance $\rho_s(\cdot,\cdot)$ and the measure μ_s (d_s-dimensional Riemannian volume) on $\Gamma_s(t)$. Denote by $\mathrm{diam}_s(A)$ the diameter of a set $A \in \Gamma_s(t)$ relative to ρ_s. Quite analogously we introduce the distance ρ_u, the measure μ_u and the diameter function diam_u on $\Gamma_u(t)$. The transformation T^n ($n \in \mathbb{Z}$) maps $\Gamma_s(t)$ onto $\Gamma_s(T^n(t))$ and $\Gamma_u(t)$ onto $\Gamma_u(T^n(t))$ scaling μ_s and μ_u by $\exp(-hn)$ and $\exp(hn)$, respectively, where $h = \sum_{\lambda \in \sigma_u} \log |\lambda|$ denotes the topological entropy of T.

Let $R \subset \mathbb{T}^d$ be an open set with the following property: for each $t \in R$ relatively open bounded sets $R_s(t) \subset \Gamma_s(t)$ and $R_u(t) \subset \Gamma_u(t)$ are specified such that $R_s(t) \subset R$, $R_u(t) \subset R$ and for every $t_1, t_2 \in R$ the intersection $R_s(t_1) \cap R_u(t_2)$ consists of exactly one point denoted by $[t_1, t_2]_R$. The set R is called a (open) *parallelogram*, if for every $t_1, t_2 \in R$ the map $[\cdot,\cdot]_R : R_s(t_1) \times R_u(t_2) \to R$ is a homeomorphism onto R. In this case we write $R = [R_s(t_1), R_u(t_2)]_R$. If $t = t_1 = t_2$, we call $R_s(t)$ and $R_u(t)$ the s-section and the u-section of R through t. For every $t \in R$ the map $R \ni x \mapsto [t, x]_R \in R_s(t)$ is a continuous open map which projects R onto $R_u(t)$ and sends the Riemannian measure on R to the measure $c_T \mu_u(R_u(y))\mu_s$ (here $y \in R$ is arbitrary). The map $R \ni x \mapsto [x, t]_R \in R_u(t)$ is a projection of R onto $R_s(t)$ with analogous properties. Moreover, the map $R_u(t_1) \ni x \mapsto [x, t_2]_R \in R_u(t_2)$ is a homeomorphism and preserves the measure μ_u. For a parallelogram R we denote by R_s and R_u the isomorphism classes (as topological and measure spaces) of its s- and u-sections. We call R_s and R_u the *edges* of R.

Topologically and measure-theoretically every parallelogram R is isomorphic to the direct product $R_s \times R_u$ of its edges.

More precisely, we have the following relation between the restriction of the Haar measure (proportional to the $d-$dimensional Riemannian measure) and the direct product of the d_s- and d_u-dimensional Riemannian measures μ_s and μ_u:

$$P|_R = c_T \mu_s|_{R_s} \times \mu_u|_{R_u},$$

where c_T is a constant only depending of the position of L_s and L_u in $\mathbb{R}^d$ (if $d = 2$ and φ is the angle between L_s and L_u, then $c_T = |\sin \varphi|$).

The projections described above send the Riemannian measure on R to the measure $c_T \mu_u(R_u(x))\mu_s$ and $c_T \mu_s(R_s(x))\mu_u$, respectively.

Set $M' = \max(||p_s||, ||p_u||)$, where $||\cdot||$ is the operator norm relative to the Euclidean norm in $\mathbb{R}^d$. If $\rho(x, y) < 1/M'$ for some $x, y \in \mathbb{T}^d$, then there exists a unique point $[x, y] \in \Gamma_s(x) \cap \Gamma_u(y)$ such that $\rho_s([x, y], x) < 1$ and $\rho_u([x, y], y) < 1$ (indeed, consider a pair $\widetilde{x}, \widetilde{y} \in \mathbb{R}^d$ with $\pi(\widetilde{x}) =$

$x, \pi(\widetilde{y}) = y$ and $dist_{\mathbb{R}^d}(\widetilde{x}, \widetilde{y}) = \rho(x, y)$ and set $[x, y] = x + \pi(p_s(\widetilde{y} - \widetilde{x}))$. A parallelogram R is said to be *small* whenever $\mathrm{diam}(R) < 1/M'$ and the map $(x, y) \mapsto [x, y]_R$ agrees with $(x, y) \mapsto [x, y]$ restricted to $R \times R$. For a small parallelogram R $\mathrm{diam}_s(R_s(x)) < 1$ and $\mathrm{diam}_u(R_u(x)) < 1$ for every $x \in R$. Let S be a bounded convex subset of a $p-$dimensional Euclidean space E. For $l = 0, \ldots, p$ we define numbers $V_l(S)$ by the relation

$$V_l(S) = \frac{b_p}{b_{p-l}} \sup_{L \in GR(l)} \mu_{p-l}(\Pi_{L'} S),$$

where $GR(l)$ is the set of all l-dimensional subspaces of E, L' is the orthogonal complement of the subspace $L \subset E$, $\Pi_{L'}$ denotes the orthogonal projection onto L', μ_{p-l} is the $(p-l)$-dimensional Riemannian measure and $b_l = \pi^{l/2}/\Gamma(1 + l/2)$. We also set

$$e(S) = \max_{l=0,\ldots p} V_l(S)^{1/(p-l)}/\mu_p(S)^{1/p}. \tag{11}$$

We shall apply the quantities just defined to convex sets contained in the subspaces L_s, L_u and their translates, and, moreover, to those contained in the stable and the unstable leaves of $\mathbb{T}^d$, since they are immersions. Denote the corresponding functionals by $V_{s,l}$ $(l = 0, \ldots, d_s)$, $V_{u,l}$ $(l = 0, \ldots, d_u)$, e_s and e_u. For a parallelogram $R = [R_s, R_u]$ we set

$$E(R) = \max\left(\frac{\mu_s^{1/d_s}(R_s)}{\mu_u^{1/d_u}(R_u)}, \frac{\mu_u^{1/d_u}(R_u)}{\mu_s^{1/d_s}(R_s)}\right). \tag{12}$$

We now state the main proposition of this section.

Proposition 3.5. *Let $R = [R_s, R_u]$ be a small convex parallelogram; then for every $l > 0$, $s, m \in \mathbb{Z}_+$*

$$\sum_{i=1}^{s} \sum_{j: l < |i-j| \le m} P(T^i(R) \cap T^j(R))$$
$$\le C s P(R) e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R) \big(q^{-l} + m P^{1/d}(R)\big),$$

where the quantities on the right hand side are defined in (10), (11) *and* (12).

The proof of this proposition is based on three lemmas.

Lemma 3.6. *There exists a constant $C > 0$ such that for every $k \in \mathbb{Z}_+$, every $x \in \mathbb{T}^d$ and all bounded convex sets $S_s \subset \Gamma_s(x)$, $S_u \subset \Gamma_u(x)$ we have*

$$V_{s,l}(T^{-k}(S_s)) \le C q^{-lk} \exp(kh) V_{s,l}(S_s), \quad l = 0, \ldots, d_s, \tag{13}$$

$$V_{u,l}(T^k(S_u)) \le C q^{-lk} \exp(kh) V_{u,l}(S_u), \quad l = 0, \ldots, d_u. \tag{14}$$

Proof. Standard. □

For two parallelograms $R^{(i)} = [R_s^{(i)}, R_u^{(i)}]$, $(i = 1, 2)$ the following representation holds:

$$P(R^{(1)} \cap R^{(2)}) = c_T \int_{R_u^{(1)}} \int_{R_s^{(2)}} \#\{R_s^{(1)}(t_1) \cap R_u^{(2)}(t_2)\} d\mu_u(t_1) d\mu_s(t_2), \tag{15}$$

hence we need to estimate $\#\{R_s(x) \cap T^k(R_u(y))$.

Lemma 3.7. *Let $x, y \in \mathbb{T}^d$ and $S_s \in \Gamma_s(x), S_u \in \Gamma_u(y)$ be bounded convex sets; then for every integer $k \geq 0$ we have*

$$\begin{aligned}&\#\{S_s \cap T^k(S_u)\}\\ \leq& C \exp(kh) e_s^{d_s}(S_s) e_u^{d_u}(S_u) E^d(S_s, S_u) \big(q^{-k/2} + (\mu_s(S_s)\mu_u(S_u))^{1/d}\big)^d,\end{aligned}$$

where

$$E(S_s, S_u) = \max\left(\frac{\mu_s^{1/d_s}(S_s)}{\mu_u^{1/d_u}(S_u)}, \frac{\mu_u^{1/d_u}(S_u)}{\mu_s^{1/d_s}(S_s)}\right).$$

Proof. Choose some $\widetilde{x}, \widetilde{y}$ such that $\pi(\widetilde{x}) = x, \pi(\widetilde{y}) = y$. Then π is a bijective map from $L_s + \widetilde{x}$ onto $\Gamma_s(x)$ and from $L_u + \widetilde{x}$ onto $\Gamma_u(x)$ (observe that π is compatible with the Riemannian metrics and measures on these submanifolds). Thus, the sets $\widetilde{S}_s \subset L_s + \widetilde{x}$ and $\widetilde{S}_u \subset L_u + \widetilde{y}$ are defined uniquely by $\pi(\widetilde{S}_s) = S_s$ and $\pi(\widetilde{S}_u) = S_u$. Note that $\widetilde{S}_s$ and $\widetilde{S}_u$ are determined up to a translation by an element of $\mathbb{Z}^d$, but such ambiguity plays no role in the sequel. Because of $\pi\widetilde{T} = T\pi$, π also maps $\widetilde{T}^k(\widetilde{S}_u)$ isomorphically onto $T^k(S_u)$ for any $k \in \mathbb{Z}^d$. Analyzing π-preimages, we see that points $t \in S_s \cap T^k(S_u)$ are in one-to-one correspondence with pairs $(t_s, t_u) \in \widetilde{S}_s \times \widetilde{T}^k(\widetilde{S}_u)$ such that $t_s - t_u \in \mathbb{Z}^d$, or, because of translation invariance, with pairs $(t'_s, t'_u) \in (\widetilde{S}_s - p_u(\widetilde{x})) \times (\widetilde{T}^k(\widetilde{S}_u) - p_s(\widetilde{y}))$ such that $t'_s - t'_u \in \mathbb{Z}^d$. Observe that $\widetilde{S}_s - p_u(\widetilde{x}) \subset L_s$ and $\widetilde{T}^k(\widetilde{S}_u) - p_s(\widetilde{y}) \in L_u$. Finally, substituting the latter set by its opposite, we arrive at

$$\#\{\widetilde{S}_s \cap \widetilde{T}^k(\widetilde{S}_u)\} = \#\{a \in \mathbb{Z}^d : a = t_s + t_u, t_s \in \widetilde{S}'_s, t_u \in \widetilde{T}^k(\widetilde{S}'_u)\}, \tag{16}$$

where $\widetilde{S}'_s \subset L_s$ and $\widetilde{S}'_u \subset L_u$ are the sets isometric to S_s and S_u. Note that we used the representation of $\widetilde{S}_s \cap \widetilde{T}^k(\widetilde{S}_u)$ as a Minkowski sum $\widetilde{S}'_s \oplus \widetilde{T}^k(\widetilde{S}'_u)$. Thus we reduced the problem of counting homoclinic points to that of counting integer lattice points in a "parallelogram" with edges $\widetilde{S}'_s$ and $\widetilde{T}^k(\widetilde{S}'_u)$. Denoting by $N_{\mathbb{Z}^d}(G)$ the number of integer lattice points in a set G, we can rewrite (16) as

$$\#\{\widetilde{S}_s \cap \widetilde{T}^k(\widetilde{S}_u)\} = N_{\mathbb{Z}^d}(\widetilde{S}'_s \oplus \widetilde{T}^k(\widetilde{S}'_u)).$$

The action of $\widetilde{T}$ preserves the righthand side of this equation, and for every $l \in \mathbb{Z}$ we can write

$$\#\{\widetilde{S}_s \cap \widetilde{T}^k(\widetilde{S}_u)\} = N_{\mathbb{Z}^d}(\widetilde{T}^{-l}(\widetilde{S}'_s) \oplus \widetilde{T}^{k-l}(\widetilde{S}'_u)).$$

For any bounded set $G \in \mathbb{R}^d$ we estimate the number $N_{\mathbb{Z}^d}(G)$ of integer lattice points in G in the following way: assign to any lattice point in G the translate of the unit cube at the origin translated to the lattice point; observe that $N_{\mathbb{Z}^d}(G)$ equals the volume of the union of these cubes; the latter is estimated from above by the volume of the $d^{1/2}$-neighbourhood (with respect to the Euclidean distance) of the set G. Denote by $B(r)$, $B_s(r)$, $B_u(r)$ the balls of radius r centered at the origin in the Euclidean spaces $\mathbb{R}^d$, L_s and L_u, respectively. Then we take a convex parallelogram $[G_s, G_u]$ as G and use the inclusion of a neighborhood of the Minkowski sum of two linearly independent sets into the Minkowski sum of their neighborhoods, taken in corresponding lower dimensional affine subspaces; after this we apply the Steiner formula to obtain

$$
\begin{aligned}
&N_{\mathbb{Z}^d}(\widetilde{T}^{-l}(\widetilde{S}'_s)\oplus\widetilde{T}^{k-l}(\widetilde{S}'_u))\\
\le&\mu(\widetilde{T}^{-l}(\widetilde{S}'_s)\oplus\widetilde{T}^{k-l}(\widetilde{S}'_u)\oplus B(d^{1/2}))\\
\le&\mu(\widetilde{T}^{-l}(\widetilde{S}'_s)\oplus B_s(M'd^{1/2}))\oplus(\widetilde{T}^{k-l}(\widetilde{S}'_u)\oplus B_u(M'd^{1/2})))\\
=&\mu_s(\widetilde{T}^{-l}(\widetilde{S}_s)\oplus B_s(M'd^{1/2}))\mu_u(\widetilde{T}^{k-l}(\widetilde{S}_u)\oplus B_u(M'd^{1/2}))\\
=&\sum_{a=0}^{d_s}C_{d_s}^a W_a(\widetilde{T}^{-l}(\widetilde{S}_s))(M'd^{1/2})^a\ \sum_{b=0}^{d_u}C_{d_u}^b W_b(\widetilde{T}^{k-l}(\widetilde{S}_u))(M'd^{1/2})^b\\
\le&\sum_{a=0}^{d_s}C_{d_s}^a V_{s,a}(\widetilde{T}^{-l}(\widetilde{S}_s))(M'd^{1/2})^a\ \sum_{b=0}^{d_u}C_{d_u}^b V_{u,b}(\widetilde{T}^{k-l}(\widetilde{S}_u))(M'd^{1/2})^b.
\end{aligned}
$$

Setting $l=[k/2]$ and applying (13), (14) and the definition of the functionals e_s and e_u, we obtain

$$
\begin{aligned}
&N_{\mathbb{Z}^d}(\widetilde{T}^{-l}(\widetilde{S}'_s)\oplus\widetilde{T}^{k-l}(\widetilde{S}'_u))\\
\le&C\sum_{a=0}^{d_s}C_{d_s}^a q^{-\frac{ak}{2}}\mathrm{e}^{\frac{kh}{2}}V_{s,a}(S_s)\ \sum_{b=0}^{d_u}C_{d_u}^b q^{-\frac{bk}{2}}\mathrm{e}^{\frac{kh}{2}}V_{u,b}(S_u)\\
\le&C\exp(kh)\sum_{a=0}^{d_s}C_{d_s}^a q^{-\frac{ak}{2}}e_s^{d_s-a}(S_s)\mu_s^{\frac{d_s-a}{d_s}}(S_s)\\
&\qquad\sum_{b=0}^{d_u}C_{d_u}^b q^{-\frac{bk}{2}}e_u^{d_u-b}(S_s)\mu_u^{\frac{d_u-b}{d_u}}(S_u)\\
\le&C\exp(kh)e_s(S_s)^{d_s}e_u(S_u)^{d_u}\left(q^{-k/2}+\mu_s^{1/d_s}(S_s)\right)^{d_s}\\
&\quad\left(q^{-k/2}+\mu_u^{1/d_u}(S_u)\right)^{d_u}\\
\le&C\exp(kh)e_s^{d_s}(S_s)e_u^{d_u}(S_u)E^{2d_sd_u/d}(S_s,S_u)\\
&\quad\left(q^{-k/2}+(\mu_s(S_s)\mu_u(S_u))^{1/d}\right)^{d_s}\left(q^{-k/2}+(\mu_s(S_s)\mu_u(S_u))^{1/d}\right)^{d_u}\\
\le&C\mathrm{e}^{kh}e_s^{d_s}(S_s)e_u^{d_u}(S_u)E^d(S_s,S_u)\left(q^{-k/2}+(\mu_s(S_s)\mu_u(S_u))^{1/d}\right)^d,
\end{aligned}
$$

since

$$\mu_s^{1/d_s}(S_s)=\mu_s^{d_u/(d_sd)}(S_s)\mu_s^{1/d}(S_s)\le E^{d_u/d}(S_s,S_u)(\mu_s(S_s)\mu_u(S_u))^{1/d},$$

$$\mu_u^{1/d_u}(S_u)=\mu_u^{d_s/(d_ud)}(S_u)\mu_u^{1/d}(S_u)\le E^{d_s/d}(S_s,S_u)(\mu_u(S_u)\mu_s(S_s))^{1/d}$$

and $2d_ud_s/d\le d$. □

Lemma 3.8. *Let $R=[R_s,R_u]$ be a small convex parallelogram. Then for every $k\in Z$*

$$P(R\cap T^k(R))\le Ce_s^{d_s}(S_s)e_u^{d_u}(S_u)E^d(S_s,S_u)P(R)(q^{-|k|/2}+P^{1/d}(R))^d. \tag{17}$$

Proof. Both sides of (17) do not change under a sign change of k. Hence, it suffices to consider the case $k \geq 0$. By formula (15) and Lemma 3.7 we have

$$
\begin{aligned}
& P(R \cap T^k(R)) \\
=& c_T \int_{R_u} \int_{T^k(R_s)} \#\{R_s(t_1) \cap T^k(R_u(t_2))\} d\mu_u(t_1) d\mu_s(t_2) \\
\leq& C \exp(kh) e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R_s, R_u) \\
& \quad \left(q^{-k/2} + (\mu_s(R_s)\mu_u(R_u))^{1/d}\right)^d \int_{R_u} \int_{T^k(R_s)} d\mu_u(t_1) d\mu_s(t_2) \\
\leq& C \mathrm{e}^{kh} e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R_s, R_u) \\
& \quad \left(q^{-k/2} + (\mu_s(R_s)\mu_u(R_u))^{1/d}\right)^d \mu_s(R_s)\mu_u(R_u) \\
=& C e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R_s, R_u) \mu_s(R_s)\mu_u(R_u) \\
& \quad \left(q^{-k/2} + (\mu_s(R_s)\mu_u(R_u))^{1/d}\right)^d \\
\leq& C e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R_s, R_u) P(R) (q^{-k/2} + P(R)^{1/d})^d.
\end{aligned}
$$

□

Corollary 3.9. *For a small convex parallelogram $R = [R_s, R_u]$ and every $k \in \mathbb{Z}$ we have*

$$
P(R \cap T^k(R)) \leq C e_s(R_s)^{d_s} e_u(R_u)^{d_u} E^d(R) P(R) (q^{-|k|/2} + P^{1/d}(R)). \tag{18}
$$

Proof. Obvious, because $E(R) = E(R_s, R_u)$. □

Proof of Proposition 3.5. In view of Corollary 3.9 we have

$$
\begin{aligned}
& \sum_{i=1}^{s} \sum_{j: l < |i-j| \leq m, j \neq i} P(T^i(R) \cap T^j(R)) \\
\leq& 2CsP(R) e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R) \sum_{j: l < j \leq m} (q^{-j/2} + P^{1/d}(R)) \\
\leq& CsP(R) e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R) ((1 - q^{-1})^{-1} q^{-l} + mP^{1/d}(R)) \\
\leq& CsP(R) e_s^{d_s}(R_s) e_u^{d_u}(R_u) E^d(R) (q^{-l} + mP^{1/d}(R)).
\end{aligned}
$$

□

3.4. Proof of the theorem.

Let $(G_n)_{n \geq 1}$ and $(r_n)_{n \geq 1}$ be the sequences satisfying the conditions of Theorem 2.3. We begin the proof of the theorem with the following simple lemma whose proof is an easy consequence of the continuity of the automorphism T and condition vi) in the theorem.

Lemma 3.10. *Let*

$$
l(n) = \min\{l \geq 1 : \exists i \in \{j : 1 \leq |j| \leq l\} \ \ni \ G_n \cap T^i(G_n) \neq \emptyset\} \to \infty.
$$

By conditions i) and ii) of the theorem $P(G_n) \to 0$. This implies, in view of conditions iii) and iv), that $r_n \to 0$, and we may assume that the numbers r_n are small enough. Then every ball

$B(z_n, r_n)$ can be inscribed into a parallelogram R_n in the following way. Choose some $\tilde{z}_n \in \mathbb{R}^d$ such that $\pi(\tilde{z}_n) = z_n$ and take a ball $\tilde{B}(\tilde{z}_n, r_n)$ in $\mathbb{R}^d$ of radius r_n around $\tilde{z}_n$. The latter ball can be inscribed into a parallelogram in $\tilde{R}_n = [p_s(\tilde{B}(\tilde{z}_n, r_n)), p_u(\tilde{B}(\tilde{z}_n, r_n))]$ (p_s and p_u denote projections onto the stable and unstable subspaces) which projects by π to a parallelogram R_n such that $B(z_n, r_n) \subset R_n$. Note that all R_n are similar and, up to similarity, only depend on the geometry of the pair of subspaces $L_s, L_u \subset \mathbb{R}^d$. In particular, the characteristics e_s,e_u and E of such parallelograms are the same and depend on T only. In view of this and condition iv) we have

$$P(G_n) \geq C\xi^{-1} P(R_n).$$

It follows from the similarity of all $\{R_n\}$ and from $r_n \to 0$ that $\mathrm{diam}(R_n) \to 0$.

Let $l(n)$ be as in Lemma 3.10. Choose $(m(n))_{n\geq 1}$ so that $m(n) \geq [\sigma \log s(n)] + 1$ for $n \in \mathbb{Z}_+$ and $m(n) = O(s^{1/d}(n))$ as $n \to \infty$, where σ is the number from the statement of Proposition 3.2. Then by Proposition 3.5 we have

$$\begin{aligned}
&\sum_{i=1}^{s(n)} \sum_{j:1\leq|i-j|\leq m(n)} E(f_n^{(i)} f_n^{(j)}) \\
=&\sum_{i=1}^{s(n)} \sum_{j:l(n)<|i-j|\leq m(n)} P(T^i(G_n) \cap T^j(G_n)) \\
\leq&\sum_{i=1}^{s(n)} \sum_{j:l(n)<|i-j|\leq m(n)} P(T^i(R_n) \cap T^j(R_n)) \\
\leq& C s(n) P(R_n)\big(q^{-l(n)} + m(n) P^{1/d}(R_n)\big) \\
\leq& C s(n) P(G_n)\big(q^{-l(n)} + m(n) P^{1/d}(G_n)\big) \\
=&O\big(q^{-l(n)} + m(n) s^{-1/d}(n)\big) = o(1)
\end{aligned}$$

as $n \to \infty$. Therefore the sum (6) tends to zero. With this choice of $m(n)$, by Proposition 3.2, also the summand (5) tends to zero. The proof is completed using Proposition 3.1.

References

1. Adler R.L.: Symbolic dynamics and Markov partitions. Bull. Amer. Soc. (N.S.) **35** (1998), 1–56.
2. Adler R.L., Weiss B.: Similarity of automorphisms of the torus. Memoirs American Math. Soc. **98** (1970).
3. Arratia R., Goldstein L.; Gordon L.: Poisson approximation and the Chen-Stein method. Stat. Science 4 (1990), 403–434.
4. Barbour A.D., Holst L., Janson, S: Poisson Approximation. Oxford Univ. Press, Oxford 1992.
5. Bowen R.: Equilibrium states and the ergodic theory of Anosov diffeomorphisms. Lecture Notes in Mathematics **470**, Springer-Verlag, Berlin-Heidelberg-New York 1975.
6. Bruin H., Saussol B., Troubetzkoy S., Vaienti S.: Return times statistics via inducing. Preprint. Inst. Math. de Luminy, 2001-42, 1–24.
7. Chen L.H.Y: Poisson approximation for dependent trials. Ann. Prob. **3** (1975), 534–545.
8. Coelho Z., Collet P.: Asymptotic limit law for the close approach of two trajectories in expanding maps of the circle. Probab. Theory Relat. Fields **99** (1994), 237–250.
9. Denker M.: Remarks on weak limit laws for fractal sets. In: Fractal geometry and Stochastics. Eds. Bandt, Graf, Zähle. Progress in Probab. **37**, Birkhäuser 1995, 167–178.
10. Dolgopyat D.: Limit theorems for partially hyperbolic systems. Preprint.
11. Gordin M.: Homoclinic approach to the central limit theorem for dynamical systems. Contemporary Mathematics **146** (1993), 149–192.

12. Haydn N.: The distribution of the first return time for rational maps. J.Stat.Phys. **94** (1999), 1027–1036.
13. Haydn N.: Statistical properties of equilibrium states for rational maps. Ergodic Theory and Dynamical Systems **20** (2000), 1371 - 1390.
14. Hirata M.: Poisson law for Axiom A diffeomorphisms. Ergodic Theory Dynamical Systems **13** (1993), 533–556.
15. Hirata M.: Poisson law for Axiom A system. Takahashi, Yoichiro (ed.), Algorithms, fractals, and dynamics. Proceedings of the Hayashibara Forum '92: International symposium on new bases for engineering science, algorithms, dynamics, and fractals, Okayama, and a symposium on algorithms, fractals, and dynamics, Kyoto. New York, NY: Plenum Press (1995), 79-87.
16. Hirata M., Benoit S., Vaienti S.: Statistics of return times: a general framework and new applications. Commun.Math.Phys. **206** (1999), 33–55.
17. Matheron G.: Random sets and integral geometry. John Wiley & Sons, Inc. 1975.
18. Pitskel B.: Poisson limit law for Markov chains. Ergodic Theory Dynamical Systems **11** (1991), 501–513.
19. Sevastyanov B.A.: Poisson limit law for a scheme of sums of independent random variables. (Russian) Th. Prob. Appl. **17** (1972), 695–699.
20. Stein C. M.: Approximate computation of expectation. Statistics Lecture Notes — Monograph Series, **7** (1986), Insitute of Mathematical Statistics, Hayward, CA.