Ergod. Th. & Dynam. Sys. (2011), **31**, 1043–1071 © *Cambridge University Press, 2010*
doi:10.1017/S0143385710000349

A large deviations bound for the Teichmüller flow on the moduli space of abelian differentials

VÍTOR ARAÚJO† and ALEXANDER I. BUFETOV‡§

† *Instituto de Matemática, Universidade Federal do Rio de Janeiro, C.P. 68.530, 21.945-970 Rio de Janeiro, RJ, Brazil*
(e-mail: vitor.araujo@im.ufrj.br)

‡ *Department of Mathematics, Rice University, MS 136, 6100 Main Street, Houston, Texas 77251-1892, USA*

§ *The Steklov Institute of Mathematics, Russian Academy of Sciences, Gubkina str. 8, 119991 Moscow, Russia*
(e-mail: aib1@rice.edu, bufetov@mi.ras.ru)

(*Received* 7 *December* 2008 *and accepted in revised form* 1 *February* 2010)

Abstract. Large deviation rates are obtained for suspension flows over symbolic dynamical systems with a countable alphabet. We use a method employed previously by the first author [Large deviations bound for semiflows over a non-uniformly expanding base. *Bull. Braz. Math. Soc. (N.S.)* **38**(3) (2007), 335–376], which follows that of Young [Some large deviation results for dynamical systems. *Trans. Amer. Math. Soc.* **318**(2) (1990), 525–543]. As a corollary of the main results, we obtain a large deviation bound for the Teichmüller flow on the moduli space of abelian differentials, extending earlier work of Athreya [Quantitative recurrence and large deviations for Teichmuller geodesic flow. *Geom. Dedicata* **119** (2006), 121–140].

1. *Introduction*

1.1. *The Teichmüller flow.*

Let $g \geq 2$ be an integer. Take an arbitrary integer vector $\kappa = (k_1, \ldots, k_\sigma)$ such that $k_i > 0$ for all i and $k_1 + \cdots + k_\sigma = 2g - 2$.

Let $\mathcal{M}_\kappa$ be the moduli space of abelian differentials with singularities prescribed by κ or, in other words, the moduli space of pairs (M, ω) such that M is a compact oriented Riemann surface of genus g and ω is a holomorphic one-form on M whose zeros have orders $k_1, \ldots, k_\sigma$. We impose the additional normalization requirement

$$\frac{i}{2}\int_M \omega \wedge \overline{\omega} = 1$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

(that is, the surface M has area 1 with respect to the area form induced by ω). The space $\mathcal{M}_\kappa$ need not be connected, and we denote by $\mathcal{H}$ a connected component of $\mathcal{M}_\kappa$. The Teichmüller flow g_t on $\mathcal{H}$ is defined by the formula

$$g_t(M, \omega) = (M', \omega') \quad \text{where } \omega' = e^t \Re(\omega) + i e^{-t} \Im(\omega),$$

and the complex structure on M' is uniquely determined by the requirement that the form ω' be holomorphic.

The flow g_t preserves a natural 'smooth' probability measure on $\mathcal{H}$ (see [**15, 19**]), which we denote by μ_κ. (See, e.g., [**13**] for a precise definition of the smooth measure. Informally, the construction of μ_κ can be explained as follows: the relative periods of ω with respect to its zeros yield a local system of coordinates on $\mathcal{H}$; up to a scalar multiple, the measure μ_κ is simply the Lebesgue measure in these period coordinates.)

Veech proved in [**20**] that the Teichmüller flow is a Kolmogorov flow with respect to μ_κ. Furthermore, μ_κ is the unique measure of maximal entropy for the flow g_t (see [**5**]).

In fact, the Teichmüller flow preserves a pair of infinitely smooth stable and unstable foliations on $\mathcal{H}$, the measure μ_κ admits globally defined conditional measures on the stable and unstable leaves, and the flow g_t expands and contracts the conditional measures uniformly. Informally, μ_κ is the Bowen–Margulis measure for g_t.

In [**20**] Veech also showed that the flow g_t admits no zero Lyapunov exponents with respect to the smooth measure (and all ergodic measures satisfying a certain technical condition). Forni [**9**] showed that the expansion on the unstable leaves (as well as the contraction on stable leaves) is *uniform on compact sets* (whence, in particular, absence of zero exponents for the flow follows for all ergodic measures).

Furthermore, the Teichmüller flow satisfies the following exponential estimate for visits to compact sets. Let $K \subset \mathcal{H}$ be a compact set with non-empty interior. For $X \in \mathcal{H}$, set

$$\tau_K(X) = \{\inf t : g_t X \in K\}.$$

Then there exists $\alpha > 0$ such that

$$\int_{\mathcal{H}} \exp(\alpha \tau_K(X))\, d\mu_\kappa(X) < +\infty. \tag{1.1}$$

Athreya established the estimate (1.1) for a special family of 'large' compact sets K. For arbitrary compact sets with non-empty interior (in fact, it suffices to require that for some $t_0 > 0$ the interior of the set $\bigcup_{0 \le t \le t_0} g_t K$ be non-empty), the exponential estimate was proved in [**6**] and independently by Avila *et al* in [**3**].

The uniform hyperbolicity of the Teichmüller flow on compact sets, in combination with estimate (1.1), allows us to carry over to the Teichmüller flow a number of facts known about geodesic flows on compact manifolds of negative curvature. In particular, in [**6**] it is shown that the Teichmüller flow satisfies the central limit theorem with respect to μ_κ, while Avila *et al* have shown in [**3**] that the time correlations of the Teichmüller flow decay exponentially. This paper is devoted to large deviations for the Teichmüller flow.

Take $\delta > 0$, let $\varphi : \mathcal{H} \to \mathbb{R}$ have average zero, and consider the set

$$B_{\delta,T}(\varphi) = \left\{ X \in \mathcal{H} : \left| \int_0^T \varphi(g_t X)\, dt \right| > \delta T \right\}.$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

If φ is the characteristic function of a specially chosen large compact set, then Athreya [**2**] showed that for any $\delta > 0$ the measure $\mu_\kappa(B_{\delta,T}(\varphi))$ decays exponentially as $T \to \infty$.

Our aim in this paper (see Theorem A below) is to extend the result of Athreya and establish exponential decay of $\mu_\kappa(B_{\delta,T}(\varphi))$ for a larger class of functions φ: namely, for functions which, following [**6**], we call *Hölder in the sense of Veech* (the formal definition was given in [**6**] and is repeated below).

It is essential for our proof that μ_κ be the measure of maximal entropy and that the exponential estimate (1.1) hold for μ_κ.

Unlike Athreya's proof, which relies on studying the action of the special linear group on $\mathcal{H}$, our argument uses only the symbolic coding for the Teichmüller flow on $\mathcal{H}$ or, more precisely, for its finite cover, namely the Teichmüller flow on Veech's space of zippered rectangles.

1.2. *Zippered rectangles.* Here we briefly recall the construction of the Veech space of zippered rectangles. We use the notation of [**5, 6**].

Let π be a permutation on m symbols, which will always be assumed irreducible in the sense that $\pi\{1, \ldots, k\} = \{1, \ldots, k\}$ implies $k = m$. The Rauzy operations a and b are defined by the formulas

$$a\pi(j) = \begin{cases} \pi j & \text{if } j \le \pi^{-1}m, \\ \pi m & \text{if } j = \pi^{-1}m + 1, \\ \pi(j-1) & \text{if } \pi^{-1}m + 1 < j \le m; \end{cases}$$

$$b\pi(j) = \begin{cases} \pi j & \text{if } \pi j \le \pi m, \\ \pi j + 1 & \text{if } \pi m < \pi j < m, \\ \pi m + 1 & \text{if } \pi j = m. \end{cases}$$

These operations preserve irreducibility. The *Rauzy class* $\mathcal{R}(\pi)$ is defined to be the set of all permutations that can be obtained from π by application of the transformation group generated by a and b. From now on, we fix a Rauzy class $\mathcal{R}$ and assume that it consists of irreducible permutations.

For $i, j = 1, \ldots, m$, denote by E_{ij} the $m \times m$ matrix whose (i, j)th entry is 1 while all other entries are zero. Let E denote the identity $m \times m$ matrix. Following Veech [**19**], we introduce the unimodular matrices

$$A(\pi, a) = \sum_{i=1}^{\pi^{-1}m} E_{ii} + E_{m,\pi^{-1}m+1} + \sum_{i=\pi^{-1}m}^{m-1} E_{i,i+1}, \tag{1.2}$$

$$A(\pi, b) = E + E_{m,\pi^{-1}m}. \tag{1.3}$$

Let

$$\Delta_{m-1} = \{\lambda \in \mathbb{R}^m : |\lambda| = 1,\ \lambda_i > 0 \text{ for } i = 1, \ldots, m\},$$

and write

$$\Delta_\pi^+ = \{\lambda \in \Delta_{m-1} : \lambda_{\pi^{-1}m} > \lambda_m\}, \quad \Delta_\pi^- = \{\lambda \in \Delta_{m-1} : \lambda_m > \lambda_{\pi^{-1}m}\}.$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Let $\mathcal{R}$ be a Rauzy class of irreducible permutations. A *zippered rectangle* associated to the Rauzy class $\mathcal{R}$ is a triple (λ, π, δ) where $\lambda \in \mathbb{R}^m_+$, $\pi \in \mathcal{R}$ and $\delta \in \mathbb{R}^m$, with the vector δ satisfying the following inequalities:

$$\delta_1 + \cdots + \delta_i \leq 0 \quad \text{for } i = 1, \ldots, m-1; \tag{1.4}$$

$$\delta_{\pi^{-1}1} + \cdots + \delta_{\pi^{-1}i} \geq 0 \quad \text{for } i = 1, \ldots, m-1. \tag{1.5}$$

The set of all δ satisfying the above inequalities is a cone in $\mathbb{R}^m$, which we shall denote by $K(\pi)$.

The area of a zippered rectangle is given by the expression

$$\begin{aligned}\text{Area}(\lambda, \pi, \delta) &= \sum_{r=1}^{m} \lambda_r h_r = \sum_{r=1}^{m} \lambda_r \left(-\sum_{i=1}^{r-1} \delta_i + \sum_{l=1}^{\pi(r)-1} \delta_{\pi^{-1}l} \right) \\ &= \sum_{i=1}^{m} \delta_i \left(-\sum_{r=i+1}^{m} \lambda_r + \sum_{r=\pi(i)+1}^{m} \lambda_{\pi^{-1}r} \right). \end{aligned} \tag{1.6}$$

(Again, our convention is that $\sum_{i=m+1}^{m}(\cdots) = 0$ and $\sum_{i=1}^{0}(\cdots) = 0$.)

Consider the set

$$\mathcal{V}(\mathcal{R}) = \{(\lambda, \pi, \delta) : \lambda \in \mathbb{R}^m_+,\ \pi \in \mathcal{R},\ \delta \in K(\pi)\}.$$

In other words, $\mathcal{V}(\mathcal{R})$ is the space of all possible zippered rectangles corresponding to the Rauzy class $\mathcal{R}$.

The Teichmüller flow P^t acts on $\mathcal{V}(\mathcal{R})$ via the formula

$$P^t(\lambda, \pi, \delta) = (e^t\lambda, \pi, e^{-t}\delta).$$

Veech also introduced a map $\mathcal{U}$ acting on $\mathcal{V}(\mathcal{R})$ through the formula

$$\mathcal{U}(\lambda, \pi, \delta) = \begin{cases} (A(\pi, b)^{-1}\lambda, b\pi, A(\pi, b)^{-1}\delta) & \text{if } \lambda \in \Delta^+_\pi, \\ (A(\pi, a)^{-1}\lambda), a\pi, A(\pi, a)^{-1}\delta) & \text{if } \lambda \in \Delta^-_\pi. \end{cases}$$

The map $\mathcal{U}$ and the flow P^t commute; see [**19**].

The volume form $\text{Vol} = d\lambda_1 \cdots d\lambda_m d\delta_1 \cdots d\delta_m$ on $\mathcal{V}(\mathcal{R})$ is preserved under the action of the flow P^t and of the map $\mathcal{U}$. Now consider the subset

$$\mathcal{V}^{(1)}(\mathcal{R}) = \{(\lambda, \pi, \delta) : \text{Area}(\lambda, \pi, \delta) = 1\},$$

i.e. the subset of zippered rectangles of area 1; observe that both P^t and $\mathcal{U}$ preserve the area of a zippered rectangle, and therefore the set $\mathcal{V}^{(1)}(\mathcal{R})$ is invariant under P^t and $\mathcal{U}$.

Let

$$\tau(\lambda, \pi) = -\log(|\lambda| - \min(\lambda_m, \lambda_{\pi^{-1}m})),$$

and for $x = (\lambda, \delta, \pi) \in \mathcal{V}(\mathcal{R})$ write

$$\tau(x) = \tau(\lambda, \pi).$$

Now define

$$\mathcal{Y}(\mathcal{R}) = \{x \in \mathcal{V}(\mathcal{R}) : |\lambda| = 1\}$$

and

$$\mathcal{V}^{(1)}_0(\mathcal{R}) = \bigcup_{x \in \mathcal{Y}(\mathcal{R}), 0 \leq t \leq \tau(x)} P^t x.$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

The set $\mathcal{V}_0^{(1)}(\mathcal{R})$ is a fundamental domain for $\mathcal{U}$, and upon identifying the points x and $\mathcal{U}x$ in $\mathcal{V}_0^{(1)}(\mathcal{R})$ we obtain a natural flow, also denoted by P^t, on $\mathcal{V}_0^{(1)}(\mathcal{R})$.

The restriction to the set $\mathcal{V}_0^{(1)}(\mathcal{R})$ of the measure given by the volume form Vol will be denoted by $\mu_{\mathcal{R}}$. By a theorem that was proved simultaneously and independently by Veech [**19**] and Masur [**15**], $\mu_{\mathcal{R}}(\mathcal{V}_0^{(1)}(\mathcal{R})) < \infty$, and we shall assume in what follows that $\mu_{\mathcal{R}}$ is normalized to have total mass 1.

Now, the vectors h and a are defined by the formulas

$$h_r = -\sum_{i=1}^{r-1} \delta_i + \sum_{l=1}^{\pi(r)-1} \delta_{\pi^{-1}l}, \tag{1.7}$$

$$a_i = -\delta_1 - \cdots - \delta_{i-1}. \tag{1.8}$$

The data (λ, h, a, π) determines the zippered rectangle (λ, π, δ) uniquely. We now metrize the space of zippered rectangles as follows.

Take two zippered rectangles $x = (\lambda, h, a, \pi)$ and $x' = (\lambda', h', a', \pi')$. Write

$$d((\lambda, h, a), (\lambda', h', a')) = \log \frac{\max_i \{\lambda_i/\lambda_i', h_i/h_i', |a_i|/|a_i'|, |h_i - a_i|/|h_i' - a_i'|\}}{\min_i \{\lambda_i/\lambda_i', h_i/h_i', |a_i|/|a_i'|, |h_i - a_i|/|h_i' - a_i'|\}},$$

and define the metric on $\Omega(\mathcal{R})$ by

$$d(x, x') = \begin{cases} d((\lambda, h, a), (\lambda', h', a')) & \text{if } \pi = \pi' \text{ and } a_m/a_m' > 0, \\ 2 + d((\lambda, h, a), (\lambda', h', a')) & \text{otherwise.} \end{cases}$$

We say that a function f on the space of zippered rectangles is *Hölder* if it is Hölder with respect to the Hilbert metric introduced above.

1.3. *Zippered rectangles and abelian differentials.* Veech [**19**] established the following connection between zippered rectangles and moduli of abelian differentials. A detailed description of this connection is given in [**14**].

A zippered rectangle naturally defines a Riemann surface endowed with a holomorphic differential. This correspondence preserves area. The orders of the singularities of ω are uniquely defined by the Rauzy class of the permutation π; see [**19**]. Thus, for any $\mathcal{R}$ we have a map

$$\pi_{\mathcal{R}} : \mathcal{V}_0^{(1)}(\mathcal{R}) \to \mathcal{M}_\kappa,$$

where κ is uniquely defined by $\mathcal{R}$.

The following theorem was proved in [**19**].

THEOREM 1.1. (Veech)

(1) *Up to a set of measure zero, $\pi_{\mathcal{R}}(\mathcal{V}_0^{(1)}(\mathcal{R}))$ is a connected component of $\mathcal{M}_\kappa$. Any connected component of any $\mathcal{M}_\kappa$ has the form $\pi_{\mathcal{R}}(\mathcal{V}_0^{(1)}(\mathcal{R}))$ for some $\mathcal{R}$.*

(2) *The map $\pi_{\mathcal{R}}$ is finite-to-one and almost everywhere locally bijective.*

(3) $\pi_{\mathcal{R}}(\mathcal{U}x) = \pi_{\mathcal{R}}(x)$.

(4) *The flow P^t on $\mathcal{V}_0^{(1)}(\mathcal{R})$ projects under $\pi_{\mathcal{R}}$ to the Teichmüller flow g_t on the corresponding connected component of $\mathcal{M}_\kappa$.*

(5) $(\pi_{\mathcal{R}})_* \mu_\kappa = \mu_{\mathcal{R}}$.

(6) $m = 2g - 1 + \sigma$.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

A function φ on $\mathcal{M}_\kappa$ is called *Hölder in the sense of Veech* if there exists a Hölder function $\theta : \mathcal{V}_0^{(1)}(\mathcal{R}) \to \mathbb{R}$ such that $\varphi \circ \pi_{\mathcal{R}} = \theta$. In particular, if a function $\varphi : \mathcal{H} \to \mathbb{R}$ is the lift of a smooth function from the underlying moduli space $\mathcal{M}_g$ of compact surfaces of genus g, then φ is Hölder in the sense of Veech; see [**6**, Remark 3, p. 587].

The main result of this paper is the following.

THEOREM A. *Let $\mathcal{H}$ be a connected component of the moduli space $\mathcal{M}_\kappa$ of abelian differentials with prescribed singularities, let g_t be the Teichmüller flow, and let μ_κ be the smooth measure. Let $\varphi : \mathcal{H} \to \mathbb{R}$ be bounded and Hölder in the sense of Veech. If $\mu_\kappa(\varphi) = 0$ and $\int_0^\tau \varphi(g_t z)\, dt \neq 0$ for some periodic point z with period $\tau > 0$, then for any $\varepsilon > 0$ the limit superior*

$$\limsup_{T \to +\infty} \frac{1}{T} \log \mu_\kappa \left\{ x \in \mathcal{H} : \left| \int_0^T \varphi(g_t x)\, dt \right| \geq T\varepsilon \right\}$$

is strictly negative.

1.4. *Symbolic coding for the Teichmüller flow.* The Teichmüller flow on Veech's space of zippered rectangles admits a representation as a suspension flow over the natural extension of the Rauzy–Veech–Zorich induction map on Veech's space of zippered rectangles [**19, 20, 24**]. The Rauzy–Veech–Zorich induction has a natural symbolic coding, and the Teichmüller flow can thus be represented as a suspension flow over a topological Markov chain with a countable alphabet. The roof function in this representation depends only on the past; on the other hand, it is neither Hölder nor bounded away from zero or infinity.

It is therefore convenient to modify the coding by considering first returns of the Teichmüller flow to an appropriately chosen subset. It turns out that the induced symbolic representation has much nicer properties; the method goes back to Veech's 1982 paper [**19**].

There is a certain freedom in choosing the subset for inducing, and thus we obtain a countable family of symbolic flows over the countable full shift which code the Teichmüller flow and whose roof functions are Hölder and bounded away from zero; for any Teichmüller-invariant probability measure, at least one of these flows codes a set of probability 1.

We summarize the above facts in the next proposition, which is essentially due to Veech [**19**]; a detailed exposition of the proof can be found in [**5**].

Let $X = \mathbb{Z}^{\mathbb{Z}}$ be the space of all bi-infinite sequences over a countable alphabet, and let $\sigma : X \to X$ be the full right shift. The Hölder structure on X is chosen in the usual way: we say that a function $\varphi : X \to \mathbb{R}_+$ is Hölder if there exists a non-negative $\alpha < 1$ such that if sequences $\omega, \tilde{\omega} \in X$ coincide at all indices not exceeding N in absolute value, then

$$|\varphi(\omega) - \varphi(\tilde{\omega})| \leq C\alpha^N.$$

If a function $r : X \to \mathbb{R}_+$ is bounded away from zero, then we denote by f_t^r the suspension flow over σ with roof function r (or just f_t when the roof function is clear from the context); we denote by X_r the phase space of the flow f_t^r.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Given a bounded measurable function φ on X_r, we define a function φ_r on X through the formula

$$\varphi_r(\omega) := \int_0^{r(\omega)} \varphi(\omega, t)\, dt. \tag{1.9}$$

We then have the following proposition (see **[5, 6]**).

PROPOSITION 1.2. *Let $\mathcal{R}$ be a Rauzy class of irreducible permutations. There exists a countable family of Hölder functions r_n, $n \in \mathbb{N}$, which are bounded away from zero and such that for each n, there exists an injective map $\mathbf{i}_n : X_{r_n} \to \mathcal{V}_0^{(1)}(\mathcal{R})$ such that:*

(1) *the diagram*

$$\begin{array}{ccc} X_{r_n} & \xrightarrow{\mathbf{i}_n} & \mathcal{V}_0^{(1)}(\mathcal{R}) \\ \downarrow{\scriptstyle f_t^{r_n}} & & \downarrow{\scriptstyle P^t} \\ X_{r_n} & \xrightarrow{\mathbf{i}_n} & \mathcal{V}_0^{(1)}(\mathcal{R}) \end{array}$$

is commutative;

(2) *for the Masur–Veech smooth measure $\mu_{\mathcal{R}}$ and all n, we have*

$$\mu_{\mathcal{R}}(\mathbf{i}_n(X_{r_n})) = 1;$$

furthermore, the measure $(\mathbf{i}_n)_*^{-1}\mu_{\mathcal{R}}$ is the unique measure of maximal entropy for the flow $f_t^{r_n}$ on X_{r_n};

(3) *for any P^t-invariant probability measure μ on $\mathcal{V}_0^{(1)}(\mathcal{R})$, there exists n such that $\mu(\mathbf{i}_n(X_{r_n})) = 1$;*

(4) *if a function $\psi : \mathcal{V}_0^{(1)}(\mathcal{R}) \to \mathbb{R}$ is Hölder in the sense of Veech, then the function $(\psi \circ \mathbf{i}_n)_r$ is Hölder on Ω.*

This proposition reduces the problem of large deviations for the Teichmüller flow to that of large deviations for suspension flows over the full countable shift. We now proceed to a study of such suspension flows. Our approach is based on the work of the first author in **[1]**, which is an adaptation of earlier work by Young **[23]**.

1.5. *Suspension flows over the countable shift.* In what follows, we present the notation for symbolic dynamics found in the papers by Buzzi and Sarig **[7, 18]** (see also the survey of Gurevich and Savchenko **[10]**), which we use in this article.

Let $\sigma : X \to X$ be the shift on the space X of bi-infinite words on a infinite countable alphabet. Denote by $\mathcal{M}_\sigma$ the family of all σ-invariant Borel probability measures on X.

We write $[x]_n$ to denote the cylinder of points in X with the same coordinates as x in the positions $0, \pm 1, \ldots, \pm(n-1)$; that is,

$$[x]_n := \{y \in X : y_i = x_i \text{ for } i \in \mathbb{Z} \text{ with } |i| < n\}.$$

We say that a function $\varphi : X \to \mathbb{R}$ is (A, α)-*Hölder-continuous* if $A > 0$ and $0 < \alpha < 1$ are such that $\operatorname{var}_k(\varphi) \le A\alpha^k$ for all $k \ge 1$, where

$$\operatorname{var}_k(\varphi) := \sup\{|\varphi(x) - \varphi(y)| : x, y \in X, y \in [x]_k\}.$$

We shall also be using the notion of *summable variation*: a function $\varphi : X \to \mathbb{R}$ is said to be of summable variation if $\sum_{k \ge 1} \operatorname{var}_k(\varphi) < \infty$.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

We say that $\varphi : X \to \mathbb{R}$ is *log-Hölder* if there exist $C, \alpha > 0$ such that for all $k \in \mathbb{N}$ and $x \in X$,

$$1 - Ce^{-\alpha k} \le \frac{\varphi(y)}{\varphi(x)} \le 1 + Ce^{-\alpha k} \quad \text{for all } y \in X \text{ with } y \in [x]_k.$$

Note that any of these conditions allows φ to be unbounded and implies the continuity of φ. For Hölder and summable variation we get uniform continuity. Moreover, writing

$$\mathrm{var}_k(\varphi, x) = \sup\{|\varphi(x) - \varphi(y)| : y \in X,\, x_i = y_i\ \forall |i| < k\},$$

we see that $\mathrm{var}_k(\varphi, x) \le \mathrm{var}_k(\varphi)$ for all $k \ge 1$ if φ is of summable variation, and that for a log-Hölder φ we get $\mathrm{var}_k(\varphi, x) \le Ce^{-\alpha k}|\varphi(x)|$, which now depends on $\varphi(x)$. Hence, a log-Hölder observable never has summable variation, unless φ is bounded. In fact, it is easy to see that the following assertions hold.

LEMMA 1.3. *If φ is Hölder, then φ is of summable variation. If φ is bounded and log-Hölder, then φ is Hölder.*

We also say that an observable $\varphi : X \to \mathbb{R}$ is *cohomologous to the zero function* if there exists a uniformly continuous function $\chi : X \to \mathbb{R}$ such that $\varphi = \chi \circ \sigma - \chi$.

The following standard notation for Birkhoff sums of a function $\varphi : X \to \mathbb{R}$ with respect to a transformation $f : X \circlearrowleft$ on a space X will be used: $S_n^f \varphi := \sum_{i=0}^{k-1} \varphi \circ f^i$. We shall write simply $S_k\varphi$ if the dynamics is clear from the context.

Recall that a Gibbs equilibrium state with respect to a potential $\psi : X \to \mathbb{R}$ is, according to Bowen [**4**] and Sarig [**18**], a probability measure $\mu = \mu_\psi$ on X such that there exist $P = P_\mu(\psi) \in \mathbb{R}$ and $K = K_\psi > 0$ satisfying

$$\frac{1}{K} \le \frac{\mu([x]_k)}{\exp(-Pk + S_k\psi(x))} \le K \quad \text{for every } x \in X \text{ and all } k \ge 0.$$

It is well known that, in this case, we have

$$P = \sup_{\nu \in \mathcal{M}_\sigma} \left(h_\nu(\sigma) + \int \psi \, d\nu \right) = h_\mu(\sigma) + \int \psi \, d\mu \tag{1.10}$$

so that μ achieves the supremum above.

THEOREM B. *Let $\sigma : X \to X$ be a countable full shift, and let $\psi : X \to \mathbb{R}$ be a log-Hölder function. We assume that $\mu = \mu_\psi$ is the* unique *Gibbs equilibrium state with respect to ψ. Then, for every observable $\varphi : X \to \mathbb{R}$ of summable variation and mean zero (i.e. $\mu(\varphi) = 0$) which is not cohomologous to the zero function, we have*

$$\limsup_{n \to +\infty} \frac{1}{n} \log \mu\{x \in X : |S_n\varphi(x)| \ge n\varepsilon\} \le \sup\left\{ h_\nu(\sigma) - \int \psi \, d\nu : |\nu(\varphi)| \ge \varepsilon,\, \nu \in \mathcal{M}_\sigma,\, \psi \in L^1(\nu) \right\}$$

and

$$\liminf_{n \to +\infty} \frac{1}{n} \log \mu\{x \in X : |S_n\varphi(x)| > n\varepsilon\} \ge \sup\left\{ h_\nu(\sigma) - \int \psi \, d\nu : |\nu(\varphi)| > \varepsilon,\, \nu \in \mathcal{M}_\sigma,\, \psi \in L^1(\nu) \right\}$$

for every $\varepsilon > 0$. In addition, the supremum above is strictly negative.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Based on this result, we are able to obtain a large deviation law, stated in the next theorem, for a suspension flow over a full countable shift with respect to the measure naturally induced by the Gibbs measure in the setting of Theorem B.

Let $r : X \to [r_0, +\infty)$ be a log-Hölder roof function with r_0 being a positive constant, and denote by X_r the space

$$\{(x, t) \in X \times [0, +\infty) : 0 \le t < r(x)\}.$$

Let $f_t : X_r \to X_r,\ t \ge 0$, be the special flow over the shift σ with roof function r (see, e.g., [**8**]).

We say that an observable $\varphi : X \to \mathbb{R}$ has *exponential tail* if there exists $\varepsilon_0 > 0$ such that $\int e^{\varepsilon_0|\varphi|}\, d\mu < \infty$.

It is well known that, given a σ-invariant probability μ, there exists a naturally induced f_t-invariant measure μ_r on X_r (see, e.g., [**8**]).

THEOREM C. *Let $\sigma : X \to X$ be a countable full shift and let $r : X \to [r_0, +\infty)$, with $r_0 > 0$, be a* log*-Hölder function with exponential tail. Assume that μ is the unique Gibbs equilibrium state with respect to $\psi = -h \cdot r$ for some fixed constant $h > 0$, and let $f_t : X_r \to X_r$ be the flow under r with induced f_t-invariant measure μ_r. For every* bounded *observable $\varphi : X_r \to \mathbb{R}$ of mean zero (i.e. $\mu_r(\varphi) = 0$), we define*

$$\varphi_r(x) := \int_0^{r(x)} \varphi(f_t(x, 0))\, dt$$

for $x \in X$ and assume that:

- *$\varphi_r : X \to \mathbb{R}$ is Hölder; and*
- *there exists a periodic point $z = f_\tau(z)$, with some period $\tau > 0$, such that*

$$\int_0^\tau \varphi(f_t(z))\, dt \neq 0.$$

Then we have, writing $\overline{r} = \mu(r)$ for simplicity,

$$\limsup_{T\to+\infty} \frac{1}{T} \log \mu_r\left\{z \in X_r : \left|\int_0^T \varphi(f_t(z))\, dt\right| \ge \varepsilon T\right\}$$
$$\le \sup\left\{h_\nu(\sigma) - \int \psi\, d\nu : |\nu(\varphi_r)| \ge \varepsilon\overline{r},\ \nu \in \mathcal{M}_\sigma,\ \psi \in L^1(\nu)\right\}.$$

In addition, the supremum above is strictly negative.

Moreover, if under the same conditions the observable φ additionally has compact support*, then*

$$\liminf_{T\to+\infty} \frac{1}{T} \log \mu_r\left\{z \in X_r : \left|\int_0^T \varphi(f_t(x, 0))\, dt\right| \le \varepsilon T\right\}$$
$$\ge \frac{1}{r_0} \sup\left\{h_\nu(\sigma) - \int \psi\, d\nu : |\nu(\varphi_r)| > \frac{\varepsilon\overline{r}}{r_0},\ \nu \in \mathcal{M}_\sigma,\ \psi \in L^1(\nu)\right\}.$$

The fact that the lower bound for the rate in Theorem C differs from the upper bound seems to be a limitation arising from the method of proof. The authors believe that an adaptation of the methods of Waddington [**21**] to this setting should provide sharper results.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

1.6. *Organization of the paper.* In §2 we prove Theorem B by adapting the arguments of Young in [**23**] to a full countable shift. In §3, we prove Theorem C after reducing the estimates of large deviation for the semiflow to estimates of certain sets of deviations for adequate observables on the base transformation, to which we apply Theorem B. Finally, in §4, we use Theorem C to complete the proof of Theorem A.

2. *Large deviations for a Gibbs measure on the full countable shift*

Here, the dynamics is given by $\sigma : X \to X$, the full countable shift. We assume that $\varphi : X \to \mathbb{R}$ is of summable variation and that ψ is log-Hölder with exponential tail (which ensures that $\psi \in L^1(\mu)$, in particular). Without loss of generality, we shall assume also that $\mu(\varphi) = 0$ and $\varphi \not\equiv 0$. For a given $\varepsilon > 0$, we consider

$$D_n^\varepsilon = \{x \in X : S_n\varphi(x) \geq n\varepsilon\}.$$

The following lemmas will be useful tools in the proof.

LEMMA 2.1. *Let $g : X \to \mathbb{R}$ be a function of summable variation and let $A_0 = \sum_{k\geq 1} \mathrm{var}_k(g)$. Suppose y differs from $x \in X$ in a single coordinate $0 \leq |i| < n$. Then*

$$|S_n g(x) - S_n g(y)| \leq \sum_{k=0}^{n-1} \mathrm{var}_k(g) \leq A_0.$$

Moreover, for given $\varepsilon > 0$, let n be such that $\varepsilon - A_0/n < \varepsilon/2$ and let $x \in X$ be such that $|S_n g(x)| > n\varepsilon$. Then, for any $y \in X$ with $x_i = y_i$ for all $|i| < n$, we have $|S_n g(y)| \geq n\varepsilon/2$.

Proof. Just observe that if $x, y \in X$ share the same coordinates except for the ith one, where $|i| < n$, then $\sigma^k x$ and $\sigma^k y$ share the same coordinates except for the $(i-k)$th one; thus

$$|g(\sigma^k x) - g(\sigma^k y)| \leq \mathrm{var}_{|i-k|}(\varphi)$$

and the first statement follows. For the second assertion, note that

$$\begin{aligned} |S_n g(y)| &\geq |S_n g(x)| - |S_n g(x) + S_n g(y)| \geq n\varepsilon - A_0 \\ &= n(\varepsilon - A_0/n) \geq n\varepsilon/2. \end{aligned}$$

□

From Lemma 2.1 we deduce the following: if we fix a symbol a and define $(\cdot)_j^a : X \to X$, $x \mapsto x_j^a$ such that $x_i^a = x_i$ for $i \neq j$ while $x_j^a = a$, and also SET x^a for x_0^a, we have

$$x \in D_n^\varepsilon \implies x^a \in D_n^{\varepsilon/2}.$$

LEMMA 2.2. *Let $\varphi : X \to \mathbb{R}$ be of summable variation (respectively, Hölder). Assume that $S_p\varphi(z) = 0$ for every σ-periodic point z with period $p \in \mathbb{N}$. Then there exists a uniformly continuous (respectively, Hölder) function $\chi : X \to \mathbb{R}$ such that $\varphi = \chi \circ \sigma - \chi$.*

This lemma says that if an observable of summable variation sums to zero over every periodic orbit, then this observable is cohomologous to the zero function.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Proof. We follow the usual proof of Livsic's theorem. Since X is the full countable shift, let $\omega \in X$ be a point with dense positive σ-orbit and define

$$\chi(\omega) := 0 \quad \text{and} \quad \chi(\sigma^n \omega) := \sum_{j=0}^{n-1} \varphi(\sigma^j \omega).$$

Then, for any $l \in \mathbb{Z}^+$, if $x^n = \sigma^n \omega$ and $m > n$ satisfy $x^m \in [x^n]_l$, we define

$$z := \overline{x_0^n \cdots x_{m-n-1}^n},$$

the σ-periodic point with period $m-n$ closest to x^n, i.e. z is periodic with period $m-n$ and $z \in [x]_n$. By construction, the jth coordinate of x^n and of z coincide for $j = 0, \ldots, l+m-n$, and by assumption $S_{m-n}\varphi(z) = 0$. Thus

$$\begin{aligned} \operatorname{var}_l(\chi) \le |\chi(x^m) - \chi(x^n)| &= \left| \sum_{j=n}^{m-n-1} \varphi(x^j) \right| = \left| \sum_{j=n}^{m-n-1} [\varphi(x^j) - \varphi(\sigma^j z)] \right| \\ &\le \sum_{j=n}^{m-n-1} |\varphi(x^j) - \varphi(\sigma^j z)| \le \sum_{j=n}^{m-n-1} \operatorname{var}_{j+l}(\varphi) \le \sum_{j>l} \operatorname{var}_l(\varphi). \end{aligned}$$

This shows that $\operatorname{var}_l(\chi) \xrightarrow[l\to+\infty]{} 0$, so χ is a uniformly continuous function. For each $n \in \mathbb{Z}^+$, it is easy to see that $\varphi(x^n) = \chi(x^{n+1}) - \chi(x^n)$; and, since $\{x^n\}_{n\in\mathbb{Z}^+}$ is dense in X and φ and χ are continuous, we get that $\varphi = \chi \circ \sigma - \chi$ as asserted. □

It follows from Lemma 2.2 that if we assume φ is not cohomologous to the zero function, then the following is true.

(G) There exists a periodic point $z \in X$ such that $S_p\varphi(z) > 0$, where p is the (minimal) period of z. Then *there exists* $\varepsilon_1 > 0$ *such that for all* $0 < \varepsilon < \varepsilon_1$ *and for all big enough* $n > 0$ *we have* $|S_n\varphi(z)| > 2\varepsilon n$.

Indeed, there exists a periodic point z with period $p \in \mathbb{Z}^+$ such that $|S_p\varphi(z)| \neq 0$; so we can find $\varepsilon_0 > 1$ such that $|S_{kp}\varphi(z)| > 3\varepsilon kp$ for all $k \in \mathbb{Z}^+$ and $0 < \varepsilon < \varepsilon_1$. Therefore, for every $0 \le l < p$ and $k > p\varepsilon \max\{|S_i\varphi(z)| : 0 \le i < p\}$,

$$|S_{kp+l}\varphi(z)| = |S_{kp}\varphi(z) + S_l(z)| \ge (kp+l)\left(\frac{3\varepsilon kp}{kp+l} - \frac{S_l\varphi(z)}{kp+l}\right) \ge 2\varepsilon(kp+l),$$

which proves property (G).

The following lemma enables us to choose a good cover for D_n^ε.

LEMMA 2.3. *Fix a finite subset* $\mathcal{A}_0$ *of the alphabet* $\mathcal{A}$. *Given a finite family of functions* $\varphi_1, \ldots, \varphi_k : X \to \mathbb{R}$ *of summable variation and real numbers* $\alpha_1, \ldots, \alpha_k$, *consider*

$$D = \{x \in X : \varphi_i(x) > \alpha_i \text{ for } i = 1, \ldots, k\}$$

and assume that D *has positive* μ*-measure.*

Then there exists a periodic point $z \in D$, *and for any given large integer* $n > 0$ *there is an integer* $m > n$ *and a* finite *family* $\mathcal{C}_n$ *of* m*-separated points in* D *such that for*

$$\mathcal{A}_n = \{a \in \mathcal{A} : a \text{ is a letter in the first } n \text{ coordinates of some element } x \in \mathcal{C}_n\}$$

the following hold:

(1) *for all* $x \in \mathcal{C}_n$, $[x]_m \subset D$;

(2) $\sum_{x\in\mathcal{C}_n} \mu([x]_m) \ge ((n-1)/n) \cdot \mu(D)$;

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

(3) *the projection* $\pi_{n,m} : X \to \mathcal{A}^{m-n}$ *onto the coordinates* $n, \dots, m-1$ *of* $\mathcal{C}_n$ *contains only letters from* $\mathcal{A}_0$, *i.e.* $\pi_{n,m}(\mathcal{C}_n) \subset \mathcal{A}_0^{m-n}$;
(4) $z_j^a \in \mathcal{C}_n$ *for all* $0 \le j \le n$ *and* $a \in \mathcal{A}_n$.

Remark 2.4. The periodic point z from property (G) belongs to $D_n^{2\varepsilon}$ for all sufficiently small $\varepsilon > 0$ and large enough $n \in \mathbb{Z}^+$.

In addition, for n such that $2A_0/n < \varepsilon$, we have that D_n^ε contains z_j^a for every symbol a in $\mathcal{A}_n$ and each $0 \le j \le n$, by Lemma 2.3.

Proof. Let $\widetilde{\mathcal{C}}_n$ be a maximal n-separated set in D; that is, we choose one point in each non-empty intersection $[a_0, a_1, \dots, a_{n-1}] \cap D$ for $a_0, \dots, a_{n-1} \in \mathcal{A}$. Observe that this set may be infinite and that $\{[x]_n : x \in \widetilde{\mathcal{C}}_n\}$ forms a disjoint open cover of D.

Now we choose a convenient finite approximation: let $\mathcal{C}_n$ be a *finite subset of* $\widetilde{\mathcal{C}}_n$ such that

$$\sum_{x \in \widetilde{\mathcal{C}}_n \setminus \mathcal{C}_n} \mu([x]_n) \le \frac{1}{n}\mu(D). \tag{2.1}$$

In this way we obtain

$$\mu(D) \le \mu\Big(\bigcup_{x \in \widetilde{\mathcal{C}}_n \setminus \mathcal{C}_n} [x]_n\Big) + \mu\Big(\bigcup_{x \in \mathcal{C}_n} [x]_n\Big) \le \frac{1}{n}\mu(D) + \mu\Big(\bigcup_{x \in \mathcal{C}_n} [x]_n\Big),$$

which implies statement (2) of the lemma for any $m > n$.

At this point, we can add finitely many elements of D to $\mathcal{C}_n$, according to convenience. We first define $\mathcal{A}_n$ as the set of all letters at the first n coordinates of the points of $\mathcal{C}_n$. Then we take the periodic point $z \in D$ given by property (G). Finally, we redefine $\mathcal{C}_n$ to be the union

$$\mathcal{C}_n \cup \{z_j^a : a \in \mathcal{A}_n, 0 \le j \le n\}.$$

This retains the above properties, and the new set $\mathcal{C}_n$ further satisfies statement (4). Since φ_i is of summable variation, for each $x \in \mathcal{C}_n$ we can find $y \in X$ and $m = m(x) > n$ such that:
(a) $y_j = x_j$ for $|j| \le m$ and, in particular, $y \in [x]_n$;
(b) $|\varphi_i(y) - \varphi_i(x)| \le \sum_{k>m-n} \mathrm{var}_k(\varphi) < \alpha_i - \varphi_i(x)$ so that $\varphi_i(y) > \alpha_i$, for all $i = 1, \dots, k$ and $y \in D$.

Now let $L_0 = \#\mathcal{A}_0$. Since $\mathcal{C}_n^\varepsilon$ is finite, we can consider $m_n = \max\{m(x) : x \in \mathcal{C}_n\}$ and then take an integer $l_n \ge \log \#\mathcal{C}_n^\varepsilon / \log L_0$. For $M_n = m_n + l_n$, replace each $x \in \mathcal{C}_n^\varepsilon$ by a y that satisfies (a)–(b) above and also the property that:
(c) $(y_{m_n+1}, \dots, y_{m_n+l_n})$ are distinct points in $\mathcal{A}_0^{l_n}$.

This ensures that the new elements of $\mathcal{C}_n$ will still be distinct points but can be separated in the ℓ_n coordinates following m_n. Note also that the choice of ℓ_n was made so as to give 'enough room' in ℓ_n coordinates to write $\#\mathcal{C}_n$ distinct words in $\mathcal{A}_0$ letters. The proof is thus complete. □

2.1. *The upper bound.* Here we give the main step of the proof of the upper bound for the limit superior in the statement of Theorem B. From now on, we take $\mathcal{C}_n$ to be the cover of D_n^ε provided by Lemma 2.3, with $i = 1$ and $\alpha_1 = n\varepsilon - \omega = n(\varepsilon - \omega/n)$ for some small $\omega > 0$. We also set $\hat{\psi} := P - \psi$, where $P = P_\mu(\psi)$ from (1.10).

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

2.1.1. *Choosing a good sequence of probability measures from the covering.* We consider the families of probability measures

$$\eta_n := \frac{1}{Z_n} \sum_{x \in \mathcal{C}_n} e^{-S_n \hat{\psi}(x)} \cdot \delta_x \quad \text{where } Z_n := \sum_{x \in \mathcal{C}_n} e^{-S_n \hat{\psi}(x)}$$

and

$$\nu_n := \frac{1}{n} \sum_{j=0}^{n-1} \sigma_*^j(\eta_n).$$

From the assumption that μ is a Gibbs equilibrium measure for $\hat{\psi}$, we get

$$Z_n \le \frac{1}{K} \sum_{x \in \mathcal{C}_n} \mu([x]_n) \le \frac{1}{K} \tag{2.2}$$

since, by the definition of $\mathcal{C}_n$, the cylinders $[x]_n$ and $[y]_n$ with distinct $x, y \in \mathcal{C}_n$ must be disjoint.

2.1.2. *Tightness of the sequence ν_n.* The following simple argument shows that we can assume $\eta_n(\sigma^{-j}[a]) > 0$ for every letter a in $\mathcal{A}_n$.

Remark 2.5. The probability measure η_n, defined above for the set D_n^ε, satisfies $\eta_n([a]) \ge e^{-S_n \hat{\psi}(z^a)}/Z_n > 0$ since, by Remark 2.4, D_n^ε contains z^a for every symbol a in $\mathcal{A}_n$. The same argument with z_j^a $(0 \le j \le n)$ in the place of z^a shows that $\eta_n(\sigma^{-j}[a]) > 0$ for every letter a in $\mathcal{A}_n$.

LEMMA 2.6. *For each letter b of $\mathcal{A}_n$ and each $0 \le j < n$, define*

$$\zeta_b^n(j) := \sum_{x \in \mathcal{C}_n \cap \sigma^{-j}[b]} \exp(-S_n \hat{\psi}(x) + \hat{\psi}(\sigma^j x)).$$

There exists a constant $L > 0$ such that $\zeta_a^n(j) \ge L$ for every $a \in \mathcal{A}_n$, all $n > 0$ and each $0 \le j < n$.

Proof. Fix some symbol $a \in \mathcal{A}_n$. For $n \in \mathbb{Z}^+$ large enough so that property (G) holds and for $0 \le j < n$, write

$$\begin{aligned}
\zeta_a^n(j) &= \sum_{b_0, \ldots, \hat{b_j}, \ldots, b_{n-1}} \sum_{\substack{x \in \mathcal{C}_n \\ x_0 = b_\ell, \ldots, x_j = a, \ldots, x_{n-1} = b_{n-1}}} \exp(-S_j \hat{\psi}(x) - S_{n-j-1} \hat{\psi}(\sigma^{j+1} x)) \\
&\ge \sum_{b_0, \ldots, \hat{b_j}, \ldots, b_{n-1}} K^2 \mu([b_0, \ldots, b_{j-1}]) \mu([b_{j+1}, \ldots, b_{n-1}]) \\
&\quad \times \sum_{\substack{x \in \mathcal{C}_n \\ x_0 = b_0, \ldots, x_j = a, \ldots, x_{n-1} = b_{n-1}}} \exp(-\hat{\psi}(\sigma^j x)),
\end{aligned}$$

where we have used only the Gibbs property and $\hat{b_j}$ denotes *absence of* b_j from the index of the sum above. Now, using the fact that z_j^a belongs to $\mathcal{C}_n \cap \sigma^{-j}[a]$ and that $\hat{\psi}$ is

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

log-Hölder, we bound the last summand as follows:

$$\sum_{\substack{x\in\mathcal{C}_n\\ x_\ell=b_\ell,\dots,x_j=a,\dots,x_{n-1}=b_{n-1}}} \exp(-\hat{\psi}(\sigma^j x)) \geq \exp(-\hat{\psi}(\sigma^j z_j^a)) \geq \exp(-\hat{\psi}(\sigma^j z) - \operatorname{var}_j(\hat{\psi}, z)).$$

Since this bound does not depend on the choice of $b_\ell, \dots, \hat{b_j}, \dots, b_{n-1}$, we conclude that

$$\zeta_a^n(j) \geq K^2 \exp(-\hat{\psi}(\sigma^j z) - \operatorname{var}_j(\hat{\psi}, z)).$$

To establish the statement in the lemma, we set

$$\begin{aligned} L &= \min\{K^2 \exp(-\hat{\psi}(\sigma^j z) - \operatorname{var}_j(\psi, z)) : 0 \leq j < n, a \in \mathcal{A}_n\} \\ &= \min\{K^2 \exp(-\hat{\psi}(\sigma^j z) - \operatorname{var}_1(\psi, z)) : 0 \leq j < p\}, \end{aligned}$$

since, for all sufficiently large n, the period p of z is smaller than n and

$$\operatorname{var}_j(\psi, z) \leq C|\psi(z)|e^{-\alpha j} \xrightarrow[j\to+\infty]{} 0.$$

The lower bound does not depend on n or on $\mathcal{A}_n$. □

Now consider the sequences of measures ν_n and η_n defined above for D_n^ε.

PROPOSITION 2.7. *There exists a constant $C_2 > 0$ such that for every symbol a in $\mathcal{A}_n$, we have $\nu_n([a]) \leq C_2\mu([a])$ for all sufficiently large n.*

This shows, in particular, that the sequence $(\nu_n)_{n\geq 1}$ is tight.

Proof. We need the following lemma.

LEMMA 2.8. *There exists $C_2 > 0$ such that $\eta_n(\sigma^{-j}[a]) \leq C_2\mu([a])$ for every $n \in \mathbb{Z}^+$, each $0 \leq j < n$ and every symbol $a \in \mathcal{A}_n$.*

Proof. Fix a symbol $a \in \mathcal{A}_n$ and an integer $0 \leq j < n$. We have

$$\begin{aligned} \eta_n(\sigma^{-j}[a]) &= \frac{1}{Z_n} \sum_{x\in\mathcal{C}_n\cap\sigma^{-j}[a]} \exp(-S_j\hat{\psi}(x) - \hat{\psi}(\sigma^j x) - S_{n-j-1}\hat{\psi}(\sigma^{j+1}x)) \\ &\leq K\mu([a]) \cdot \frac{\zeta_a^n(j)}{Z_n}, \end{aligned}$$

since

$$\exp(-\hat{\psi}(\sigma^j x)) \leq \exp(-\inf\hat{\psi}([a])) \leq K\mu([a])$$

by the Gibbs property of μ. By using Lemma 2.6, we can bound Z_n as follows:

$$\begin{aligned} Z_n &= \sum_b \sum_{x\in\mathcal{C}_n\cap[b]} e^{-\hat{\psi}(x)} e^{-S_{n-1}\hat{\psi}(\sigma x)} \\ &\geq \sum_b \frac{\mu([b])}{K} \sum_{x\in\mathcal{C}_n\cap[b]} e^{-S_{n-1}\hat{\psi}(\sigma x)} \\ &\geq \sum_b \frac{\mu([b])}{K} \cdot \zeta_b^n \geq \frac{L}{K}. \end{aligned}$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Finally, we find an upper bound on ζ_a^n by using again the Gibbs property of μ:

$$\zeta_a^n(j) \le \sum_{x \in \mathcal{C}_n \cap \sigma^{-j}[a]} K\mu([x]_j) \cdot K\mu([\sigma^{j+1}x]_{n-j-1}) \le K^2,$$

since $\mathcal{C}_n$ is a n-separated subset.

This shows that

$$\eta_n(\sigma^{-j}[a]) \le K\mu([a]) \cdot K^2/(L/K) = (K^4/L) \cdot \mu([a])$$

and concludes the proof of Lemma 2.8. □

Now, since the bounds in Lemmas 2.6 and 2.8 do not depend on $0 \le j < n$ for all large enough n, it follows that for any given $a \in \mathcal{A}_n$ and sufficiently large n we have

$$\frac{1}{n}\sum_{j=0}^{n-1} \eta_n(\sigma^{-j}[a]) = \nu_n([a]) \le C_2\mu([a]),$$

which completes the proof of Proposition 2.7. □

2.1.3. *Upper bound for large deviations on the base dynamics.* Using the definitions of ν_n and Z_n and observing that

$$\nu_n(\varphi) = \frac{1}{n}\sum_{j=0}^{n-1} \eta_n(\varphi \circ \sigma^j) = \frac{1}{Z_n}\sum_{x \in \mathcal{C}_n} e^{-S_n\hat{\psi}(x)} \cdot \frac{1}{n}\sum_{j=0}^{n-1} \varphi(\sigma^j x) > \varepsilon - \frac{\omega}{n}$$

for all $n > 0$, we see that any weak* accumulation point ν of the sequence ν_n must satisfy $\nu(\varphi) \ge \varepsilon$. In what follows, we assume without loss of generality that ν_n converges to ν as $n \to \infty$ in the weak* topology.

On the one hand, since $\{[x]_n : x \in \mathcal{C}_n\}$ is an approximate cover of D_n^ε, from Lemma 2.3 and the Gibbs property we obtain

$$\begin{aligned}\limsup_{n\to+\infty} \frac{1}{n}\log \mu(D_n^\varepsilon) &\le \limsup_{n\to+\infty} \frac{1}{n}\log K\frac{n}{n-1}\sum_{x\in\mathcal{C}_n} e^{-S_n\hat{\psi}(x)} \\ &= \limsup_{n\to+\infty} \frac{1}{n}\log Z_n.\end{aligned}$$

On the other hand, by considering the partition† of X given by

$$\mathcal{P} = \{[a] : a \in \mathcal{A}_0\} \cup \{x \in X : x_0 \notin \mathcal{A}_0\},$$

we note that by the choice of the points in $\mathcal{C}_n$ the refined partition

$$\mathcal{P}^{M_n} := \bigvee_{|i|<M_n} \sigma^i\mathcal{P}$$

separates the elements of $\mathcal{C}_n$, i.e. there is at most one element of $\mathcal{C}_n$ in each atom of $\mathcal{P}^n$. From [**22**, Lemma 9.9], we have

$$H_{\nu_n}(\mathcal{P}^{M_n}) - \int S_n\hat{\psi}(x)\, d\nu_n(x) = \log \sum_{x\in\mathcal{C}_n} e^{-S_n\hat{\psi}(x)}.$$

† The complicated choice of covering in Lemma 2.3 was made to enable us to choose a finite partition here.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Following standard arguments (see, e.g., [**22**, p. 220]), we then deduce that for every $1 < q < n$,

$$\frac{1}{n}\log Z_n \le \frac{1}{q}H_{\nu_n}(\mathcal{P}^q) + \frac{2q}{n}\log \#\mathcal{P} - \int \hat{\psi}\, d\nu_n, \tag{2.3}$$

where $\#\mathcal{P}$ denotes the number of elements of the partition $\mathcal{P}$.

For the final step we need the following lemma.

LEMMA 2.9. *We have $\nu_n(\hat{\psi}) \to \nu(\hat{\psi})$ as $n \to \infty$.*

Proof. From the log-Hölder property and μ-integrability of $\hat{\psi}$ we get† that for any given fixed $x \in X$,

$$\infty > \mu(|\hat{\psi}|) = \sum_a \mu(|\hat{\psi}|\chi_{[a]}) \ge \sum_a (|\hat{\psi}(x^a)| - \operatorname{var}_1(\hat{\psi}, x^a))\mu([a])$$

and thus $\sum_a |\hat{\psi}(x^a)|\mu([a]) \le \mu(|\hat{\psi}|) + \operatorname{var}_1(\hat{\psi}, x^a) < \infty$.

Given a function $g : X \to \mathbb{R}^+$, define for each $L > 0$ the function g_L to equal g if $g > L$ and 0 otherwise.

Now, from the log-Hölder property of $\hat{\psi}$ and the μ-integrability of $\hat{\psi}$, together with Proposition 2.7, we obtain that for any large enough n and positive L,

$$\begin{aligned}
\nu_n(|\hat{\psi}|_L) &= \sum_a \nu_n(|\hat{\psi}|_L \cdot \chi_{[a]}) \le \sum_{a:\sup \hat{\psi}|[a] > L} (|\hat{\psi}(x^a)| + \operatorname{var}_1(\hat{\psi}, x^a)) \cdot \nu_n([a]) \\
&\le C_2 \sum_{a:\sup \hat{\psi}|[a] > L} (|\hat{\psi}(x^a)| + \operatorname{var}_1(\hat{\psi}, x^a))\mu([a]) \\
&\le \int_{|\psi|>L} (|\psi(x)| + 2\operatorname{var}_1(\psi, x))\, d\mu(x) \\
&\le \int_{|\psi|>L} |\psi(x)|(1 + 2Ce^{-\alpha})\, d\mu(x) \xrightarrow[L\to+\infty]{} 0,
\end{aligned}$$

which shows that $\nu_n(\hat{\psi})$ is a uniformly convergent sequence of integrals. □

From inequality (2.3) and Lemma 2.9, we conclude that

$$\begin{aligned}
\limsup_{n\to+\infty} \frac{1}{n}\log Z_n &\le \frac{1}{q}\limsup_{n\to+\infty} H_{\nu_n}(\mathcal{P}^q) + \limsup_{n\to+\infty} \int -\hat{\psi}\, d\nu_n \\
&\le h_\nu(\sigma, \mathcal{P}) - \int \hat{\psi}\, d\nu \le h_\nu(\sigma) - \nu(\hat{\psi}).
\end{aligned} \tag{2.4}$$

Finally, we note that as a consequence of the assumption that μ is the unique Gibbs measure associated to ψ, for all $\nu \in \mathcal{M}_\sigma \setminus \{\mu\}$ we have

$$h_\mu(\sigma) - \mu(P - \psi) = 0 > h_\nu(\sigma) - \nu(P - \psi).$$

This shows that (2.4) is negative.

† The same argument shows that, in fact, $\psi \in L^1(\mu) \iff \sum_{a\in\mathcal{A}} |\psi(x^a)| < \infty$ for any given fixed $a \in \mathcal{A}$.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

2.2. *The lower bound.* Let ν be a σ-invariant probability measure such that $\varphi, \psi \in L^1(\nu)$ and $|\nu(\varphi)| > \varepsilon$ for a fixed small $\varepsilon > 0$. We define

$$\widehat{D}_n^{\varepsilon} = \{x \in X : S_n\varphi(x) > n\varepsilon\}.$$

We will find a sequence ν_{n_k} of invariant measures converging to ν such that

$$\mu(\widehat{D}_{n_k}^{\varepsilon}) \geq n_k \cdot \exp(h_{\nu_{n_k}} - \nu_{n_k}(\hat{\psi}) - 2\delta)$$

for small $\delta > 0$ with

$$h_{\nu_{n_k}} \geq h_\nu - \delta \quad \text{and} \quad \nu_{n_k}(\hat{\psi}) \leq \nu(\hat{\psi}) + \delta.$$

Then

$$\lim_{k\to+\infty} \frac{1}{n_k} \log \mu(\widehat{D}_{n_k}^{\varepsilon}) \geq h_\nu - \int \hat{\psi}\, d\nu - 4\delta. \tag{2.5}$$

Following the ideas in [**23**], we approximate ν by a finite convex combination of σ-ergodic measures and then use their ergodicity and a weak form of specification to build the separated set which will provide the estimates for $\mu(\widehat{D}_n^{\varepsilon})$.

2.2.1. *Approximating by ergodic measures.* We use the ergodic decomposition theorem from [**16, 17**] for the measure-preserving endomorphism σ of the Lebesgue space $(X, \mathcal{B}, \nu)$, where $\mathcal{B}$ is the Borel σ-algebra of X.

THEOREM 2.10. *There exists a smallest σ-invariant measurable partition $\mathcal{I}$ of X excluding a set of ν-null measure. Let $\{\nu_\xi\}_{\xi\in\mathcal{I}}$ be the disintegration of ν into conditional probability measures, and let $\hat{\nu}$ be the probability measure induced on the quotient space $X/\mathcal{I}$. Then:*

(1) *the ν_ξ are σ-invariant ergodic probability measures for $\hat{\nu}$-a.e. $\xi \in \mathcal{I}$;*
(2) *for each $n \geq 1$ and every ν-integrable function $g : X \to \mathbb{R}^n$:*
   (a) *$\xi \in \mathcal{B} \mapsto \nu_\xi(g)$ is $\hat{\nu}$-integrable;*
   (b) *$\nu(g) = \int \nu_\xi(g)\, d\hat{\nu}(\xi)$;*
(3) *$h_\nu(\sigma) = \int h_{\nu_\xi}(\sigma|\xi)\, d\hat{\nu}(\xi)$.*

We use this theorem to build a finite linear convex combination of ergodic measures that approximates ν.

LEMMA 2.11. *Define $g : X/\mathcal{I} \to \mathbb{R}^3$ by*

$$g(\xi) = (\nu_\xi(\varphi), \nu_\xi(\hat{\psi}), h_{\nu_\xi})$$

(which is $\hat{\nu}$-integrable), and let $0 < \delta < (\nu(\varphi) - \varepsilon)/4$.

Then there exists a finite linear convex combination ν_0 of ergodic measures such that $\|\nu(g) - \nu_0(g)\| < \delta$, where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^3$.

Proof. Choose $\zeta > 0$ so that

$$\zeta/(1-\zeta) < \delta/(2 + \|\nu(g)\|).$$

Let $\mathcal{Q}$ be a denumerable partition of $\mathbb{R}^3$ into cubes with diameter smaller than ζ. Also, let $\mathcal{Q}_0 \subset \mathcal{Q}$ be the family of such cubes bounded by a cube $[-L, L]^3$, where $L > 0$ is big enough so that:

- $q = \hat{\nu}(g^{-1}(\bigcup \mathcal{Q}_0)) > 1 - \zeta$;
- $s = \sum_{R\in\mathcal{Q}\setminus\mathcal{Q}_0} (|\nu_\xi(\varphi)| + |\nu_\xi(\hat{\psi})| + h_{\nu_\xi}) \cdot \hat{\nu}(g^{-1}R) < \zeta$.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

We can now define a probability measure

$$\nu_0 = \frac{1}{q} \sum_{R \in \mathcal{Q}_0} \hat{\nu}(g^{-1}R) \cdot \nu_{\xi_R}$$

where, for each $R \in \mathcal{Q}_0$, ν_{ξ_R} is an ergodic measure chosen in $g^{-1}(R)$. Hence ν_0 is a finite convex linear combination of σ-ergodic measures.

Analogously, we define a tail measure

$$\nu_1 = \sum_{R \in \mathcal{Q} \setminus \mathcal{Q}_0} \hat{\nu}(g^{-1}R) \cdot \nu_{\xi_R}$$

and note that $\|\nu_1(g)\| \le s < \zeta$.

Now we check that ν_0 is an approximation of ν:

$$\begin{aligned}
\|\nu(g) - \nu_0(g)\| &= q^{-1}\|q\nu(g) - q\nu_0(g)\| \\
&= q^{-1}\|q\nu(g) - (q\nu_0 + \nu_1)g + \nu_1(g)\| \\
&\le \frac{1}{q}\|(q-1)\nu(g)\| + \frac{1}{q}\|\nu(g) - (q\nu_0 + \nu_1)g\| + \frac{\|\nu_1(g)\|}{q} \\
&\le \frac{1-q}{q}\|\nu(g)\| + \frac{\zeta}{q} + \frac{\zeta}{q} \le (2 + \|\nu(g)\|)\frac{\zeta}{1-\zeta} \le \delta.
\end{aligned}$$

So the proof of the lemma is complete. □

Let us write $\nu_0 = \sum_{i=1}^{k} a_i \eta_i$, where $a_i > 0$, $\sum_i a_i = 1$ and the η_i are σ-ergodic probability measures.

2.2.2. *Building a good cover using ergodicity and a form of specification.* Now we will make heavy use of the fact that we have approximated ν by a combination of *ergodic* measures. As in Lemma 2.6, let $A_0 = \sum_{k \ge 1} \mathrm{var}_k(\psi)$. Let $N > 1$ be such that

$$\frac{A_0}{N-k} \le \frac{\delta}{4} \quad \text{and} \quad \left|\frac{1}{N}\eta_i(\varphi)\right| + \frac{\delta}{N} \le \frac{\delta}{8k}.$$

In addition, choose N big enough so that for any $n > N$ and each $i = 1, \dots, k$, the subset

$$Y_n^i = \left\{ \frac{1}{[a_i n]} S_{[a_i n]}\hat{\psi} \le \eta_i(\hat{\psi}) + \delta \text{ and } \frac{1}{[a_i n]} S_{[a_i n]}\varphi \ge \eta_i(\varphi) - \delta \right\}$$

of X satisfies $\eta_i(Y_n^i) > 1 - \delta$ (where $[a] = \max\{j \in \mathbb{Z} : j \le a\}$ denotes the biggest integer that is less than or equal to $a \in \mathbb{R}$). Assume also that N is large enough so that $\mathrm{var}_{[a_i n]}(\varphi) < \delta/k$ for all $i = 1, \dots, k$ and $n > N$.

Using a result of Katok, [**12**, Theorem 1.1]†, we have that there exists an $[a_i n]$-separated set $E_n^i \subset Y_n^i$ with at least $\exp([a_i n](h_{\eta_i} - \delta))$ points. We number the elements of E_n^i as $x_1^i, \dots, x_{m_i}^i$.

† The theorem was stated only for homeomorphisms of compact spaces, but the proof does not use this assumption!

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Consider a k-tuple $(j_1, \dots, j_k)$ with $1 \le j_i \le m_i$ for $i = 1, \dots, k$. There is a corresponding point $y = y_{j_1,\dots,j_k} \in X$ (not uniquely defined) whose positive orbit shadows the orbit segments

$$(x^1_{j_1}, \sigma x^1_{j_1}, \dots, \sigma^{[a_1 n]} x^1_{j_1}), \dots, (x^k_{j_k}, \sigma x^k_{j_k}, \dots, \sigma^{[a_k n]} x^k_{j_k}).$$

Let $\mathcal{E}$ be the family of points obtained in this manner, and fix $y \in \mathcal{E}$. By the summable variation of φ, for $m = \sum_i [a_i n]$ and $n_0 = \min_i [a_i n]$ we have

$$\left| S_m \varphi(y) - \sum_{i=1}^{k} S_{[a_i n]} \varphi(x^i_{j_i}) \right| \le \sum_{i=1}^{k} \mathrm{var}_{[a_i n]}(\varphi) \le \delta.$$

Now, because $a_i n - 1 \le [a_i n] \le a_i n$, we can write

$$\begin{aligned} \frac{1}{m} S_m \varphi(y) &\ge \frac{1}{m} \sum_{i=1}^{k} S_{[a_i n]} \varphi(x^i_{j_i}) - \delta \ge \sum_{i=1}^{k} \frac{[a_i n]}{m} (\eta_i(\varphi) - \delta) - \delta \\ &\ge \frac{1}{m} \sum_{i=1}^{k} a_i n \cdot (\eta_i(\varphi) - \delta) - \frac{1}{m} \sum_i (\eta_i(\varphi) - \delta)^+ - \delta \end{aligned}$$

where $a^+ = \max\{0, a\}$, since we must take the sign of $\eta_i(\varphi) - \delta$ into account. Note that by the choice of δ in Lemma 2.11 and the fact that $m \le \sum_i a_i n = n$, we have

$$\sum_{i=1}^{k} a_i n \cdot (\eta_i(\varphi) - \delta) = n \cdot (\nu_0(\varphi) - \delta) \ge n \cdot (\nu(\varphi) - 2\delta) > 0.$$

This, together with the choice of N, gives

$$\frac{1}{m} S_m \varphi(y) \ge \frac{n}{m} \cdot (\nu(\varphi) - 2\delta) - \frac{\delta}{8} - \delta \ge \nu(\varphi) - \frac{25}{8}\delta > \varepsilon,$$

which means that $y \in \widehat{D}^{\varepsilon}_m$.

In addition, note that for different choices of the k-tuples we get distinct points $y, y' \in \mathcal{E}$ which are m-separated; that is, $[y]_m \cap [y']_m = \emptyset$ by construction.

Finally, observe that for $w \in [y]_m$ we have, by Lemma 2.1,

$$\begin{aligned} S_m \varphi(w) &\ge S_m \varphi(y) - 2A_0 \ge \left(\nu(\varphi) - \frac{25}{8}\delta - \frac{\delta}{4} \right) \cdot m - 2A_0 \\ &\ge \left(\nu(\varphi) - \frac{27\delta}{8} - \frac{2A_0}{m} \right) \cdot m \ge \left(\nu(\varphi) - \frac{31}{8}\delta \right) \cdot m > m \cdot \varepsilon, \end{aligned}$$

where we have used the fact that

$$m = \sum_{i=1}^{k} [a_i n] \ge \sum_{i=1}^{k} (a_i n - 1) = n - k.$$

Thus $\{[y]_m\}_{y \in \mathcal{E}}$ is a family of m-separated subsets inside D^{ε}_m.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

2.2.3. *Estimating the measure of* $\widehat{D}_m^\varepsilon$. Using the previous arguments, we can finally bound the measure of $\widehat{D}_m^\varepsilon$ from below. Since $\mathcal{E} \subset \bigcup_i Y_n^i$ and μ is a Gibbs measure, we have

$$\begin{aligned}\mu(\widehat{D}_m^\varepsilon) &\geq \sum_{y\in\mathcal{E}} \mu([y]_m) \geq \sum_{y\in\mathcal{E}} \frac{1}{K}\cdot \exp(-S_m\hat{\psi}(y)) \\ &\geq \frac{1}{K}\sum_{y\in\mathcal{E}} \exp\biggl(-\sum_i [a_i n]\cdot(\eta_i(\hat{\psi})+\delta)\biggr).\end{aligned}$$

We also know that

$$\#E_n^i \geq \exp([a_i n](h_{\eta_i}-\delta)),$$

and from this we obtain

$$\mu(\widehat{D}_m^\varepsilon) \geq \frac{1}{K}\cdot\exp\biggl(\sum_i [a_i n]\cdot(h_{\eta_i}-\eta_i(\hat{\psi})-2\delta)\biggr).$$

Hence, for any given $\delta>0$, there exists a large N such that for all $n>N$ we can find $m \geq n-k$ satisfying

$$\frac{1}{m}\log\mu(\widehat{D}_m^\varepsilon) \geq -\frac{1}{m}\log K + \frac{1}{m}\sum_{i=1}^{k}[a_i n]\cdot(h_{\eta_i}-\eta_i(\hat{\psi})-2\delta).$$

By the upper bound on large deviations already obtained, we know that $h_{\eta_i}-\eta_i(\hat{\psi})-2\delta\leq 0$; hence

$$\begin{aligned}\frac{1}{m}\log\mu(\widehat{D}_m^\varepsilon) &\geq -\frac{1}{m}\log K + \frac{n}{m}\sum_{i=1}^{k} a_i\cdot(h_{\eta_i}-\eta_i(\hat{\psi})-2\delta) \\ &\geq -\frac{1}{m}\log K + (h_\nu-\delta)-(\nu(\varphi)+\delta)-2\delta.\end{aligned}$$

This completes the proof of (2.5).

2.3. *The rates.* Now we obtain explicit expressions for the rates of decay of the measure of the deviation set. On the one hand, in §2.1 we showed that there exists a σ-invariant probability ν such that $|\nu(\varphi)|\geq\varepsilon$, ψ is ν-integrable and inequality (2.4) holds; that is,

$$\limsup_{n\to+\infty}\frac{1}{n}\log\mu(D_n^\varepsilon) \leq h_\nu(\sigma)-\int\hat{\psi}\,d\nu<0. \tag{2.6}$$

On the other hand, in §2.2 it was proved that given any $\delta>0$ and σ-invariant probability ν such that $|\nu(\varphi)|>\varepsilon$ and ψ is ν-integrable, there exists a sequence n_k tending to $+\infty$ such that (2.5) holds; that is,

$$\liminf_{n\to+\infty}\frac{1}{n}\log\mu(\widehat{D}_n^\varepsilon) \geq \sup_{\nu\in\mathcal{M}_\sigma}\biggl\{h_\nu(\sigma)-\int\hat{\psi}\,d\nu : |\nu(\varphi)|>\varepsilon,\ \nu(\hat{\psi})<\infty\biggr\}. \tag{2.7}$$

From (2.6) and (2.7) we deduce that the supremum above is also an upper bound for the limit superior and is strictly negative. This completes the proof of Theorem B.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

3. *Large deviations for maximal entropy measures for special flows over a full countable shift*

In this section we prove Theorem C. We assume that μ is a σ-ergodic probability on the full countable shift X which is a Gibbs measure and the unique equilibrium state with respect to $\psi = -h \cdot r$, where h is the topological entropy of the flow $f_t : X_r \circlearrowleft$ built over σ with roof function $r : X \to [r_0, +\infty)$, for some $r_0 > 0$. In particular, r (as well as ψ) is μ-integrable.

This means that the induced f_t-invariant probability measure μ_r on X_r is the measure of maximal entropy of the flow.

We assume further that r is log-Hölder with exponential tail.

3.1. *Reduction to the base dynamics.*

Here we describe how to pass from the deviation set for the suspension flow with respect to a bounded observable of summable variation to another deviation set for the base dynamics with respect to another observable, now unbounded.

Consider a continuous observable $\varphi : X_r \to \mathbb{R}$, and note that the time average of φ under the action of the semiflow on the point $z = (x, s) \in X_r$ can be written as

$$\begin{aligned}\int_0^T \varphi(f_t(z))\,dt = \sum_{j=1}^{n-1} \int_0^{r(\sigma^j(x))} \varphi(f_t(\sigma^j(x), 0))\,dt \\ + \int_s^{r(x)} \varphi(f_t(x, 0))\,dt \\ + \int_0^{T+s-S_n r(x)} \varphi(f_t(\sigma^n(x), 0))\,dt,\end{aligned}$$

where $n = n(x, s, T) \in \mathbb{N}$ is such that

$$S_n^\sigma r(x) \le s + T < S_{n+1}^\sigma r(x).$$

Recalling that

$$\varphi_r(x) := \int_0^{r(x)} \varphi(f_t(x, 0))\,dt$$

for $x \in X$, we obtain

$$\int_0^T \varphi(f_t(z))\,dt = S_n^\sigma \varphi_r(x) + I_T(x, s), \tag{3.1}$$

where

$$I_T(x, s) = \int_0^{T+s-S_n r(x)} \varphi(f_t(\sigma^n(x), 0))\,dt - \int_0^s \varphi(f_t(x, 0))\,dt.$$

Assume now that $\varphi : X_r \to \mathbb{R}$ is bounded and that $\varphi_r : X \to \mathbb{R}$ is Hölder.

Note that φ_r is not necessarily bounded. Recall also that $\mu_r(\varphi) = \mu(\varphi_r)/\mu(r)$. We assume without loss of generality that $\mu(\varphi_r) = 0$. Moreover, we assume that there exists some σ-periodic point $z \in X$, with period $p \in \mathbb{Z}^+$, such that

$$S_p^\sigma \varphi_r(z) = \int_0^\tau \varphi(f_t(z, 0))\,dt \neq 0 \quad \text{where } \tau := S_p r(z). \tag{3.2}$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

3.2. *The limit superior.* From now on, all Birkhoff sums will be taken with respect to σ. The previous discussion showed that for $\varepsilon > 0$,

$$\left\{z \in X_r : \left|\int_0^T \varphi(f_t(z))\,dt\right| \geq \varepsilon T\right\} = \{(x, s) \in X_r : |S_n\varphi_r(x) + I_T(x, s)| \geq \varepsilon T\},$$

where $n = n(x, s, T)$ as before. Because

$$|S_n\varphi_r(x)| + |I_T(x, s)| \geq |S_n\varphi_r(x) + I_T(x, s)| \geq \varepsilon T,$$

we therefore have that for every $0 < \xi < 1$, the deviation set is contained in

$$\{(x, s) \in X_r : |S_n\varphi_r(x)| \geq \varepsilon(1 - \xi)T\} \cup \{(x, s) \in X_r : |I_T(x, s)| \geq \varepsilon\xi T\}. \tag{3.3}$$

Observe first that, owing to the exponential tail of r, the subset $R_L := \{x \in X : r(x) > L\}$ for $L > 0$ satisfies

$$C_0 := \int e^{\varepsilon_0 r}\,d\mu \geq \int_{R_L} e^{\varepsilon_0 r}\,d\mu \geq e^{\varepsilon_0 L}\mu(R_L),$$

and thus $\mu(R_L) \leq C_0 e^{-\varepsilon_0 L}$.

Upon taking $L > 0$ big enough so that $(n + 1)e^{-\varepsilon_0 n/2} < 1$ for all $n > L$, we have

$$\int_{R_L} r\,d\mu \leq \sum_{i \geq L} \int_i^{i+1} r\,d\mu \leq C_0 \sum_{i \geq L} (i + 1)e^{-\varepsilon_0 i} \leq C_0 \sum_{i \geq L} e^{-\varepsilon_0 i/2} \leq C_0 \frac{e^{-\varepsilon_0 L/2}}{1 - e^{-\varepsilon_0/2}}. \tag{3.4}$$

Now we deduce an upper bound for the measure of each set in (3.3). We shall write $\|\varphi\|$ for $\sup|\varphi|$. On the one hand, since

$$|I_T(x, s)| \leq (s + S_{n+1}r(x) - S_n r(x)) \cdot \|\varphi\| = (s + (r \circ \sigma^n)(x)) \cdot \|\varphi\|,$$

using (3.4) and the fact that μ is σ-invariant yields

$$\begin{aligned}
&\mu_r\{(x, s) \in X_r : |I_T(x, s)| \geq \varepsilon\xi T\} \\
&\quad \leq \mu_r\left\{(x, s) \in X_r : s \geq \frac{\varepsilon\xi T}{2\|\varphi\|}\right\} + \mu_r\left\{(x, s) \in X_r : (r \circ \sigma^n)(x) \geq \frac{\varepsilon\xi T}{2\|\varphi\|}\right\} \\
&\quad = \frac{1}{\bar{r}}\left(\int_{\{x \in X : r(x) \geq \varepsilon\xi T/(2\|\varphi\|)\}} r\,d\mu + \int_{\{x \in X : (r \circ \sigma^n)(x) \geq \varepsilon\xi T/(2\|\varphi\|)\}} r \circ \sigma^n\,d\mu\right) \\
&\quad = \frac{1}{\bar{r}}\left(\int_{R_{\varepsilon\xi T/(2\|\varphi\|)}} r\,d\mu + \int_{\sigma^{-n}R_{\varepsilon\xi T/(2\|\varphi\|)}} r \circ \sigma^n\,d\mu\right) = \frac{2}{\bar{r}}\int_{R_{\varepsilon\xi T/(2\|\varphi\|)}} r\,d\mu \\
&\quad \leq 2\frac{C_0}{\bar{r}} \cdot \frac{\exp(-\varepsilon_0\varepsilon\xi T/(2\|\varphi\|))}{1 - \exp(-\varepsilon_0/2)}.
\end{aligned} \tag{3.5}$$

On the other hand, there is a relation between $n(x, s, T)$ and T for μ_r-almost all points: writing $\bar{r}$ for $\mu(r) = \int r\,d\mu$, we have

$$\frac{S_n r(x)}{n} \leq \frac{T + s}{n} < \frac{S_{n+1}r(x)}{n} \quad \text{so that } \frac{n}{T} = \frac{n(x, s, T)}{T} \xrightarrow[T\to\infty]{} \frac{1}{\bar{r}}. \tag{3.6}$$

Note that, for all sufficiently small $a > 0$, the first subset in (3.3) is contained in the following union:

$$\left\{(x, s) \in X_r : \frac{T}{n} \leq (1 - a)\bar{r}\right\} \cup \{(x, s) \in X_r : |S_n\varphi_r(x)| \geq n\varepsilon(1 - \xi)(1 - a)\bar{r}\}, \tag{3.7}$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

where we have omitted the dependence of n on (x, s, T) for simplicity. Again, given $\omega > 0$, the second subset in (3.7) is contained in

$$(X\backslash R_{\omega T}) \cap \left\{(x, s) \in X_r : |S_n\varphi_r(x)| \geq n\varepsilon(1-\xi)(1-a)\overline{r} \text{ and } \frac{T}{n} \leq (1+a)\overline{r}\right\}$$
$$\cup\, R_{\omega T} \cup \left\{(x, s) \in X_r : \frac{T}{n} > (1+a)\overline{r}\right\}. \tag{3.8}$$

For the first subset in (3.8), we can use Theorem B (because we have a σ-periodic point z such that $S_p\varphi_r(z) \neq 0$ by condition (3.2) and, from Lemma 2.2, we know that φ_r is not cohomologous to the zero function) to obtain an exponentially small upper bound that depends on T. In what follows, for any $L > 0$, we write R_L^c for $X\backslash R_L$; then

$$\mu_r\left(R_{\omega T}^c \cap \left\{|S_n\varphi_r| \geq n\varepsilon(1-\xi)(1-a)\overline{r} \text{ and } \frac{T}{n} \leq (1+a)\overline{r}\right\}\right)$$
$$\leq \frac{\omega T}{\overline{r}}\mu\left\{x \in X : |S_n\varphi_r(x)| \geq n\varepsilon(1-\xi)(1-a)\overline{r} \text{ and } n \geq \frac{T}{(1+a)\overline{r}}\right\}$$
$$\leq \frac{\omega T}{\overline{r}} \cdot \exp((\beta+\delta)T/((1+a)\overline{r})) \tag{3.9}$$

for some small $\delta > 0$, where $\beta = \beta(a, \xi) < 0$ is given by Theorem B:

$$\beta = \sup_{\nu\in\mathcal{M}_\sigma}\left\{h_\nu(\sigma) - \int \psi\, d\nu : |\nu(\varphi_r)| \geq \varepsilon(1-\xi)(1-a)\overline{r},\ \psi \in L^1(\nu)\right\}.$$

For the middle subset in (3.8), we can use the bound (3.4) to get

$$\mu_r(R_{\omega T}) \leq \frac{C_0}{\overline{r}}\frac{e^{-\varepsilon_0\omega T/2}}{1-e^{-\varepsilon_0/2}}. \tag{3.10}$$

Now we only need an upper large deviation estimate on $n(x, s, T)/T$ to finish.

3.2.1. *The lap number versus flow time.* From (3.6), we consider the measure of the following subsets of X_r, for any given $0 < \zeta < 1/\overline{r}$:

$$\mu_r\left\{\left|\frac{n(x, s, T)}{T} - \frac{1}{\overline{r}}\right| \geq \zeta\right\} = \mu_r\left\{\frac{n}{T} - \frac{1}{\overline{r}} \geq \zeta\right\} + \mu_r\left\{\frac{n}{T} - \frac{1}{\overline{r}} \leq -\zeta\right\}$$
$$\overset{\text{by (3.6)}}{=} \mu_r\left\{T \leq \frac{n\overline{r}}{1+\zeta\overline{r}} \text{ and } \frac{1}{n}S_nr \cdot \left(1 - \frac{s}{S_nr}\right) \leq \frac{\overline{r}}{1+\zeta\overline{r}}\right\}$$
$$+ \mu_r\left\{\frac{T}{n} \geq \frac{\overline{r}}{1-\zeta\overline{r}}\right\}. \tag{3.11}$$

Since r itself can be taken as an observable in Theorem B, when n is so large that

$$1 - \frac{s}{S_nr(x)} \geq 1 - \frac{s}{nr_0} \geq 1 - \xi > 0 \quad \text{with } \frac{\overline{r}}{(1-\xi)(1+\zeta\overline{r})} < \overline{r},$$

we can bound the first summand in (3.11) by

$$\mu_r\left\{T \leq \frac{n\overline{r}}{1+\zeta\overline{r}} \text{ and } \frac{1}{n}S_nr \cdot \left(1 - \frac{s}{S_nr}\right) \leq \frac{\overline{r}}{1+\zeta\overline{r}}\right\}$$
$$\leq \underbrace{\mu_r\left\{n \geq T\frac{1+\zeta\overline{r}}{\overline{r}} \text{ and } \frac{1}{n}S_nr \leq \frac{\overline{r}}{(1-\xi)(1+\zeta\overline{r})}\right\}}_{A_n}.$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

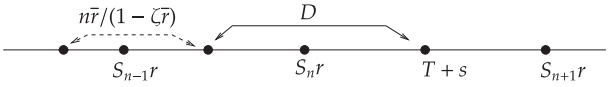


FIGURE 1. Relative positions on the real line of $T+s$ and $n\bar{r}/(1-\zeta\bar{r})$.

Next, we split A_n into pieces that are easier to estimate: for $\omega > 0$ small and T big we have, from (3.4) and Theorem B,

$$\begin{aligned}\mu_r(A_n) &= \mu_r(A_n \cap R_{\omega T}) + \mu_r(A_n \backslash R_{\omega T}) \\ &\le \mu_r(R_{\omega T}) + \frac{\omega T}{\bar{r}}\mu\left\{x \in X : n \ge T\frac{1+\zeta\bar{r}}{\bar{r}} \text{ and } \frac{1}{n}S_n r \le \frac{\bar{r}}{(1-\xi)(1+\zeta\bar{r})}\right\} \\ &\le \frac{C_0}{\bar{r}}\frac{e^{-\varepsilon_0\omega T/2}}{1-e^{-\varepsilon_0/2}} + \frac{\omega T}{\bar{r}}\exp((\gamma+\delta)(1+\zeta\bar{r})T/\bar{r}) \qquad (3.12)\end{aligned}$$

because

$$(\gamma+\delta)n < (\gamma+\delta)(1+\zeta\bar{r})T/\bar{r},$$

where $\delta > 0$ is small and $\gamma = \gamma(\xi, \zeta) < 0$ is given by

$$\sup_{\nu\in\mathcal{M}_\sigma}\left\{h_\nu(\sigma) - \int \psi\, d\nu : |\nu(r) - \bar{r}| \ge \bar{r}\left(1 - \frac{1}{(1-\xi)(1+\zeta\bar{r})}\right), \psi \in L^1(\nu)\right\}.$$

For the second summand in (3.11) observe that, by using the relation (3.6) and considering the position of $n\bar{r}/(1-\zeta\bar{r})$ on the real line with respect to $S_nr(x)$ (see Figure 1), we have either

$$r(\sigma^n(x)) = S_{n+1}r(x) - S_nr(x) \ge D/2$$

or

$$r(\sigma^{n-1}(x)) = S_nr(x) - S_{n-1}r(x) \ge D/2,$$

where $D = T + s - \bar{r}n/(1-\zeta\bar{r})$.

Then, upon setting

$$\tau := T/n > \bar{r}/(1-\zeta\bar{r}) > \bar{r} > r_0$$

and using the σ-invariance of μ together with the tail estimate (3.4) and the bound $T \ge r_0 n$ (recall that $n = n(x, s, T)$), we find that

$$\begin{aligned}&\mu_r\left\{(x,s) \in X_r : \frac{T}{n(x,s,T)} \ge \frac{\bar{r}}{1-\zeta\bar{r}}\right\} \\ &\le \mu_r\left\{r\circ\sigma^{n-i} \ge T\left(1 + \frac{s}{T} - \frac{\bar{r}/\tau}{1-\zeta\bar{r}}\right) \text{ and } \frac{T}{n} = \tau, \text{ for } i = 0 \text{ or } 1\right\} \\ &\le \mu_r\left\{r\circ\sigma^{n-i} \ge \underbrace{T\left(1 - \frac{\bar{r}/r_0}{1-\zeta\bar{r}}\right)}_{T(\zeta)} \text{ for } i = 0 \text{ or } 1\right\} \\ &= \mu_r\{(x,s) \in X_r : x \in \sigma^{-n}R_{T(\zeta)}\} + \mu_r\{(x,s) \in X_r : x \in \sigma^{-n+1}R_{T(\zeta)}\} \\ &\le \sum_{k=0}^{[T/r_0]+1} \mu_r\underbrace{\{(x,s) \in X_r : x \in \sigma^{-k}R_{T(\zeta)}\}}_{B_k}.\end{aligned}$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Now we split the set into two parts, as in (3.12), and use the σ-invariance of μ to obtain

$$\begin{aligned}
&\sum_{k=0}^{[T/r_0]+1} (\mu_r(B_k \cap R_{\omega T}) + \mu_r(B_k \backslash R_{\omega T})) \\
&\leq \frac{C_0}{\bar{r}} \left(\left[\frac{T}{r_0} \right] + 2 \right) \frac{\exp(-\varepsilon_0 \omega T/2)}{1 - \exp(-\varepsilon_0/2)} + \frac{\omega T}{\bar{r}} \sum_{k=0}^{[T/r_0]+1} \mu(R_{T(\zeta)}) \\
&\leq \frac{C_0 + \omega T}{\bar{r}} \left(\left[\frac{T}{r_0} \right] + 2 \right) \left(\frac{\exp(-\varepsilon_0 \omega T/2) + \exp(-T\varepsilon_0/2(1 - \bar{r}/(r_0 - r_0 \zeta \bar{r})))}{1 - \exp(-\varepsilon_0/2)} \right).
\end{aligned} \tag{3.13}$$

Putting (3.12) and (3.13) together and taking $\omega, \delta > 0$ arbitrarily small, we get

$$\limsup_{T \to +\infty} \frac{1}{T} \log \mu_r \left\{ \left| \frac{n}{T} - \frac{1}{\bar{r}} \right| \geq \zeta \right\} \leq \max \left\{ \gamma \frac{1 + \zeta \bar{r}}{\bar{r}}, -\frac{\varepsilon_0}{2} \left(1 - \frac{\bar{r}}{r_0(1 - \zeta \bar{r})} \right) \right\}. \tag{3.14}$$

3.2.2. *Exponentially small tail.* Finally, comparing the last subset in (3.8) with the usage of ζ in (3.11), we see that $a + 1 = (1 - \zeta \bar{r})^{-1}$ and hence $\zeta = a/((1 + a)\bar{r})$. So, upon putting (3.5), (3.8), (3.9), (3.10) and (3.14) together we arrive at (again letting $\omega, \delta > 0$ be arbitrarily small)

$$\begin{aligned}
&\limsup_{T \to +\infty} \frac{1}{T} \log \mu_r \left\{ z \in X_r : \left| \int_0^T \varphi(f_t(z))\, dt \right| \geq \varepsilon T \right\} \\
&\quad \leq \max \left\{ \frac{\beta}{(1 + a)\bar{r}}, \frac{\gamma}{\bar{r}} \cdot \frac{2 + a}{1 + a}, -\frac{\varepsilon_0}{2} \left(1 - \frac{\bar{r}}{r_0(1 - \zeta \bar{r})} \right), -\frac{\varepsilon_0 \varepsilon \xi}{2\|\varphi\|}, -\frac{\varepsilon_0 \omega}{2\|\varphi\|} \right\}
\end{aligned}$$

for all $a, \zeta > 0$ sufficiently small and $\xi, \varepsilon, \omega > 0$. Observe that ε_0 does not depend on ε and, by the assumptions on μ (i.e. that it is the unique equilibrium state for the potential ψ), we have

$$\gamma(\xi, \zeta) \xrightarrow[\xi, \zeta \to 0]{} 0.$$

Thus we can take $\varepsilon, \xi, \zeta > 0$ so small that $\beta/((1 + a)\bar{r})$ is the maximum value above. Then, by letting a be very small, we obtain the statement of Theorem C.

3.3. *The limit inferior.* For the limit inferior we need to restrict the class of observables under consideration. We assume that $\varphi : X_r \to \mathbb{R}$ is continuous and bounded, with $\mu_r(\varphi) = 0$ and $\varphi_r : X \to \mathbb{R}$ a Hölder function, and we assume now, in addition, that φ has *compact support*, i.e. there exists a compact subset $K \subset X_r$ such that $\varphi \equiv 0$ on $X_r \backslash K$. Let $r_1 = \max_K r \geq r_0$ in what follows. We show that any deviation set for φ under the flow f_t can be related to a specific deviation set for φ_r under the shift map, in such a way that we can apply the lower bound for the rate of large deviations provided by Theorem B.

We start by noting that the function

$$\varrho(x, s) := \varphi(x, s) - \varphi_r(x)$$

is bounded and satisfies

$$\varrho_r(x) = \int_0^{r(x)} (\varphi(x, t) - \varphi_r(x))\, dt = \int_0^{r_1} (\varphi(x, t) - \varphi_r(x))\, dt = 0$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

and

$$\|\varrho\| := \sup_{(x,s)\in X_r} \left|\varphi(x,s) - \int_0^{r(x)} \varphi(x,t)\,dt\right| \le \|\varphi\| + r_1\|\varphi\| = (1+r_1)\|\varphi\|.$$

Now, from relation (3.1) applied with ϱ in place of φ, for all $(x,s)\in X_r$ and $T>0$ we have

$$\int_0^T \varrho(f_t(x,s))\,dt = S_n\varrho_r(x) + I_T(x,s) = I_T(x,s) \quad \text{and}$$

$$\begin{aligned} |I_T(x,s)| &\le \left|\int_0^{T+s-S_nr(x)} \varrho(f_t(\sigma^n(x),0))\,dt\right| + \left|\int_0^s \varrho(f_t(x,0))\,dt\right| \\ &\le C_1 := 2r_1\|\varrho\| \le 2r_1(1+r_1)\|\varphi\|, \end{aligned} \tag{3.15}$$

with $n = n(x,s,T)$. Therefore, by the definition of ϱ, for each $(x,s)\in X_r$ and all $T>0$ we have

$$\left|\int_0^T \varphi(f_t(x,s))\,dt\right| \ge \left|\int_0^T \varphi_r(\pi\circ f_t(x,s))\,dt\right| - C_1, \tag{3.16}$$

where $\pi : X_r \to X$ is the projection onto the first coordinate.

Because φ has compact support, the relation (3.1) tells us that *for the purpose of calculating* $\int_0^T \varphi(f_t(x,s))\,dt$ *with given* $(x,s)\in X_r$ *and* $T>0$*, we may assume without loss of generality that both* $s<r_1$ *and* $T+s-S_nr(x)<r_1$ *hold, since* $\varphi(y,t)=0$ *for all* $y\in X$ *and* $t\ge r_1$. In other words, any value of the Birkhoff integral of φ for the flow f_t always coincides with the value of the Birkhoff integral for some $(x,s)\in X_r$ and $T>0$ satisfying the conditions stated above.

Now we again use the relation (3.1) with φ_r in place of φ to get

$$\begin{aligned} \left|\int_0^T \varphi_r(\pi\circ f_t(x,s))\,dt\right| &= \left|\sum_{i=0}^{n-1}\int_0^{r(\sigma^i(x))} \varphi_r(\sigma^i(x))\,dt \right. \\ &\qquad \left. - s\varphi_r(x) + (T - r(\sigma^{n-1}(x)))\varphi_r(\sigma^{n-1}(x))\right| \\ &\ge |S_n(r\cdot\varphi_r)(x)| - 2r_1\|\varphi_r\| \\ &\ge r_0\cdot|S_n\varphi_r(x)| - 2r_1^2\|\varphi\|. \end{aligned} \tag{3.17}$$

This implies that if

$$|S_n\varphi_r(x)| > \varepsilon(1+\xi)T/r_0,$$

then

$$\begin{aligned} \left|\int_0^T \varphi(f_t(x,s))\,dt\right| &\ge r_0\cdot\frac{\varepsilon(1+\xi)T}{r_0} - 2r_1^2\|\varphi\| - C_1 \\ &= \left(\varepsilon(1+\xi) - \frac{2r_1^2 + 2r_1(1+r_1)}{T}\|\varphi\|\right)T > \varepsilon T \end{aligned}$$

for all $\xi,\varepsilon>0$ and $T > (4r_1^2+2r_1)\|\varphi\|/(\xi\varepsilon)$. Therefore, for $\varepsilon,\xi,\zeta>0$ we can write

$$\begin{aligned} &\mu_r\left\{(x,s) : \left|\int_0^T \varphi(f_t(x,s))\,dt\right| > \varepsilon T\right\} \\ &\ge \mu_r\{(x,s) : |S_n\varphi_r(x)| > \varepsilon(1+\xi)T/r_0\} \\ &\ge \mu_r\left\{(x,s)\in X_r : |S_n\varphi_r(x)| > \varepsilon(1+\xi)\frac{T}{r_0} \text{ and } \frac{T}{n} \le \frac{\bar r}{1-\zeta\bar r}\right\}. \end{aligned}$$

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Finally, since $r \geq r_0$, we have the following (crude) lower bound for the last expression:

$$r_0 \cdot \mu\left\{x \in X : |S_n\varphi_r(x)| > \frac{\varepsilon(1+\xi)\overline{r}}{r_0(1-\zeta\overline{r})}n\right\}.$$

Theorem B gives that for all large enough n and $T > 0$ (recall that $T \geq nr_0$),

$$\mu_r\left\{(x, s) \in X_r : \left|\int_0^T \varphi(f_t(x, s))\, dt\right| > \varepsilon T\right\} \geq r_0 \cdot e^{(\omega+\delta)n} \geq r_0 \cdot e^{(\omega+\delta)T/r_0},$$

where $\delta > 0$ can be taken arbitrarily small and $\omega = \omega(\varepsilon, \xi, \zeta) < 0$ is given by Theorem B:

$$\omega = \sup\left\{h_\nu(\sigma) - \int \psi\, d\nu : |\nu(\varphi_r)| > \frac{\varepsilon(1+\xi)\overline{r}}{r_0(1-\zeta\overline{r})},\ \nu \in \mathcal{M}_\sigma,\ \psi \in L^1(\nu)\right\}.$$

Since $\xi, \zeta, \delta > 0$ are arbitrary, we see that the exponential decay rate of the measure is bounded below by

$$\liminf_{T\to+\infty} \frac{1}{T} \log \mu_r\left\{\left|\int_0^T \varphi(f_t(x, s))\, dt\right| > \varepsilon T\right\} \geq \frac{\omega(\varepsilon, 0, 0)}{r_0}.$$

The proof of Theorem C is thus complete.

4. *Application to the Teichmüller flow*

In this short final section, we apply Theorem C to the coding of the Teichmüller flow on the moduli space of abelian differentials.

The applications of these results to systems that admit a coding through flows over countable full shifts are consequences of the following simple observation.

Recall that a measure-preserving dynamical system $(Y, g_t, \mathcal{B}, \nu)$ (where g_t is a $\mathcal{B}$-measurable flow) is a *factor* of the system $(X, f_t, \mathcal{A}, \mu)$ (where f_t is a $\mathcal{A}$-measurable flow) if there exists a measurable map $\mathbf{i} : Y \to X$ such that:

- $\mathbf{i}$ commutes with the actions of the dynamical systems, i.e. $\mathbf{i}(g_t y) = f_t(\mathbf{i}y)$ for all $y \in Y$ and all t;
- $\mathbf{i}(Y) = X$ and the induced measure $\nu(\mathbf{i}^{-1}A)$, for $A \in \mathcal{A}$, equals μ.

LEMMA 4.1. *Assume that $(Y, g_t, \mathcal{A}, \nu)$ is a factor of $(X, f_t, \mathcal{B}, \mu)$ with a factor map $\mathbf{i} : Y \to X$.*

Then for any observable $\varphi : X \to \mathbb{R}$ with $\mu(\varphi) = 0$, the deviation sets

$$D_X(\varphi, \varepsilon) = \left\{z \in X : \left|\int_0^T \varphi(f_t(z))\, dt\right| > \varepsilon T\right\}$$

and

$$D_Y(\varphi \circ \mathbf{i}, \varepsilon) = \left\{z \in Y : \left|\int_0^T (\varphi \circ \mathbf{i})(f_t(z))\, dt\right| > \varepsilon T\right\}$$

are related as follows:

$$D_X(\varphi, \varepsilon) = h(D_Y(\varphi \circ \mathbf{i}, \varepsilon)).$$

So if we can relate two flows as above and identify the class of functions ψ for which there exists $\varphi : X \to \mathbb{R}$ satisfying $\psi = \varphi \circ \mathbf{i}$ and a large deviation estimate for the system $(X, f_t, \mathcal{B}, \mu)$, then we can pass the same estimates for that class of functions on to the system $(Y, g_t, \mathcal{A}, \nu)$.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

Remark 4.2. However, if the given isomorphism does not respect other measures, then we may not be able to interpret the deviation rates for the system (Z, Y^t, m) as the variational bounds in Theorems B and C. See Proposition 1.2(3).

Note that the roof functions $r_n : X \to \mathbb{R}_+$ in Proposition 1.2 are Hõlder and bounded away from zero, so they are automatically log-Hölder as well: if $r_n(\omega) \geq b_n > 0$ for all $\omega \in X$, then for $N \in \mathbb{N}$ and $\omega, \omega' \in X$ with $\omega' \in [\omega]_N$,

$$\left|1 - \frac{r_n(\omega)}{r_n(\omega')}\right| = \frac{|r_n(\omega) - r_n(\omega')|}{|r_n(\omega')|} \leq \frac{C\alpha^N}{b_n}.$$

Moreover, if we fix $n \in \mathbb{N}$ and a connected component $\mathcal{H}$ of $\mathcal{M}_\kappa$, a function $\varphi : \mathcal{H} \to \mathbb{R}$ which is bounded and Hölder in the sense of Veech will induce a function $\theta : \mathcal{V}_0^{(1)}(\mathcal{R}) \to \mathbb{R}$ such that $\varphi \circ \pi_{\mathcal{R}} = \theta$, and then the function $\psi = \theta \circ \mathbf{i}_{\mathbf{n}} : X_{r_n} \to \mathbb{R}$ is such that $\psi_{r_n} : X \to \mathbb{R}$ is Hölder.

Finally, each roof function r_n has exponential tail (with respect to μ_κ), since $\tau_K \geq c$ for some positive constant c for the compact set $K \subset \mathcal{H}$ in (1.1).

Hence, we can use Theorem C with ψ as the observable to estimate the rate of decay of the deviation sets for φ.

Acknowledgements. We are deeply grateful to Boris M. Gurevich, Amir Dembo and Dmitry Dolgopyat for useful discussions. Part of this work was done while the first author was visiting Rice University, and another part while the second author was visiting the Instituto Nacional de Matemática Pura e Aplicada (IMPA), Brazil. We would like to thank these host institutions for their warm hospitality. The first author was partially supported by CNPq, FAPERJ and PRONEX (Brazil). The second author was partially supported by the National Science Foundation under grant DMS-0604386, by the Edgar Odell Lovett Fund at Rice University and by the Programme on Mathematical Control Theory of the Presidium of the Russian Academy of Sciences.

References

[1] V. Araújo. Large deviations bound for semiflows over a non-uniformly expanding base. *Bull. Braz. Math. Soc. (N.S.)* **38**(3) (2007), 335–376.

[2] J. S. Athreya. Quantitative recurrence and large deviations for Teichmüller geodesic flow. *Geom. Dedicata* **119** (2006), 121–140.

[3] A. Avila, S. Gouëzel and J.-C. Yoccoz. Decay of correlations for Teichmüller flows. *Publ. Math. Inst. Hautes Études Sci.* **104** (2006), 143–211.

[4] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics, 470)*. Springer, Berlin, 1975.

[5] A. Bufetov and B. Gurevich. Existence and uniqueness of the measure of maximal entropy for the Teichmüller flow on the moduli space of abelian differentials. *Preprint*, 2008, arXiv.org.

[6] A. I. Bufetov. Decay of correlations for the Rauzy–Veech–Zorich induction map on the space of interval exchange transformations and the central limit theorem for the Teichmüller flow on the moduli space of abelian differentials. *J. Amer. Math. Soc.* **19**(3) (2006), 579–623.

[7] J. Buzzi and O. Sarig. Uniqueness of equilibrium measures for countable Markov shifts and multidimensional piecewise expanding maps. *Ergod. Th. & Dynam. Sys.* **23**(5) (2003), 1383–1400.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press

[8] I. P. Cornfeld, S. V. Fomin and Y. G. Sinaĭ. *Ergodic Theory (Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 245)*. Springer, New York, 1982; translated from the Russian by A. B. Sosinskiĭ.

[9] G. Forni. Deviation of ergodic averages for area-preserving flows on surfaces of higher genus. *Ann. of Math.* (2) **155**(1) (2002), 1–103.

[10] B. M. Gurevich and S. V. Savchenko. Thermodynamic formalism for symbolic Markov chains with a countable number of states. *Uspekhi Mat. Nauk* **53**(2(320)) (1998), 3–106.

[11] J. Hubbard and H. Masur. Quadratic differentials and foliations. *Acta Math.* **142**(3–4) (1979), 221–274.

[12] A. Katok. Lyapunov exponents, entropy and periodic orbits for diffeomorphisms. *Publ. Math. Inst. Hautes Études Sci.* **51** (1980), 137–173.

[13] M. Kontsevich. Lyapunov Exponents and Hodge Theory. *The Mathematical Beauty of Physics: A Memorial Volume for Claude Itzykson (Saclay, France, 5–7 June 1996) (Advanced Series in Mathematical Physics, 24)*. Eds. J. M. Drouffe and J. B. Zuber. World Scientific, Singapore, 1997, pp. 318–332.

[14] M. Kontsevich and A. Zorich. Connected components of the moduli spaces of abelian differentials with prescribed singularities. *Invent. Math.* **153**(3) (2003), 631–678.

[15] H. Masur. Interval exchange transformations and measured foliations. *Ann. of Math.* (2) **115**(1) (1982), 169–200.

[16] V. A. Rokhlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Transl.* **10** (1962), 1–52; translated from *Mat. Sb.* **25** (1949), 107–150.

[17] V. A. Rokhlin. Lectures on the entropy theory of transformations with invariant measure. *Russian Math. Surveys* **22**(5) (1967), 1–52; translated from *Uspekhi Mat. Nauk* **22**(5) (1967), 3–56.

[18] O. M. Sarig. Thermodynamic formalism for countable Markov shifts. *Ergod. Th. & Dynam. Sys.* **19**(6) (1999), 1565–1593.

[19] W. A. Veech. Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.* (2) **115**(1) (1982), 201–242.

[20] W. A. Veech. The Teichmüller geodesic flow. *Ann. of Math.* (2) **124**(3) (1986), 441–530.

[21] S. Waddington. Large deviations asymptotics for Anosov flows. *Ann. Inst. H. Poincaré Sect. C* **13**(4) (1996), 445–484.

[22] P. Walters. *An Introduction to Ergodic Theory*. Springer, Berlin, 1982.

[23] L. S. Young. Some large deviation results for dynamical systems. *Trans. Amer. Math. Soc.* **318**(2) (1990), 525–543.

[24] A. Zorich. Finite Gauss measure on the space of interval exchange transformations. Lyapunov exponents. *Ann. Inst. Fourier (Grenoble)* **46**(2) (1996), 325–370.

https://doi.org/10.1017/S0143385710000349 Published online by Cambridge University Press