Ergod. Th. & Dynam. Sys. (1997), **17**, 1339–1358
Printed in the United Kingdom © *1997 Cambridge University Press*

The averaging method for perturbations of mixing flows

H. SCOTT DUMAS† and FRANÇOIS GOLSE‡§

† *Department of Mathematics, University of Cincinnati, Cincinnati, OH 45221–0025, USA*

‡ *School of Mathematics, Institute for Advanced Study, Olden Lane, Princeton, NJ 08540, USA*

(*Received* 2 *August* 1995 *and accepted in revised form* 24 *July* 1996)

Abstract. For a class of multiphase averaging problems in which the unperturbed flow of the fast variables is a transitive Anosov flow or is 'sufficiently rapidly mixing', we obtain what we believe is an optimal power-law estimate, in the small parameter, of the rate at which the average maximum difference between solutions of the exact and averaged problems converges to zero. This verifies and extends an earlier conjectural remark by V. I. Arnold.

1. *Introduction*

The method of averaging is the most rudimentary perturbation technique for approximating the solutions of systems of differential equations with 'fast' and 'slow' variables. In such systems, it is tempting to approximate the leading-order evolution of the slow variables by first considering them fixed, then 'averaging' the contribution of the fast variables over the appropriate geometric object (or 'fiber'). Use of this idea goes back at least as far as Lagrange and Laplace; but in spite of the simplicity and intuitive appeal of averaging methods, the rigorous theory has been slow to develop, in part because it is often more effective to tailor the method to specific problems than to appeal to general theorems. Yet even in applications, theoretical results serve to indicate the optimal result one might expect in a given problem. In this paper we extend the theory for multiphase problems (i.e. problems with more than one fast variable) in which the unperturbed flow of the fast variables is a transitive Anosov flow or is 'sufficiently rapidly mixing'. We obtain what we believe is an optimal power-law estimate, in the small parameter, of the rate at which the average maximum difference between solutions of the exact and averaged problems converges to zero.

For a survey of results in multiphase averaging theory, as well as their applications and connections with other topics, the reader may consult [**Ar3**], [**Ar4**], or [**LM**]; these

§ Permanent address: Université Paris VII & Ecole Normale Supérieure, D.M.I., 45 rue d'Ulm, 75230 Paris Cedex 05, France.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

texts also cite a large body of more specialized research. Here we place our result in context by recalling that the most general multiphase averaging theorem is of the type originally due to D. V. Anosov, concerning a system of ODEs of the form

$$\frac{dX}{dt} = \varepsilon f(X, Y), \quad \frac{dY}{dt} = \omega(X, Y) + \varepsilon g(X, Y)$$

and its associated (slow) averaged system

$$\frac{dW}{dt} = \varepsilon \overline{f}(W) \equiv \varepsilon \int_{\mathcal{M}} f(W, Y)\, d\mu_Y.$$

Here, $\varepsilon \geq 0$ is a small parameter, the slow variables X and W evolve in $\mathbb{R}^m$, Y belongs to the smooth compact n-dimensional manifold $\mathcal{M}$, and $d\mu_Y$ denotes the invariant measure of the unperturbed flow (i.e. the flow of the system with $\varepsilon = 0$) on $\mathcal{M}$, normalized so that $\int_{\mathcal{M}} d\mu_Y = 1$.

Anosov's theorem states roughly that given $\rho > 0$, the relative measure of the set of ('bad') initial conditions for which

$$\max_{0 \leq t \leq 1/\varepsilon} |X(t) - W(t)| > \rho \quad \text{with } X(0) = W(0)$$

vanishes together with ε. Anosov's original formulation of this theorem [**An**] requires that $\mathcal{M}$ be immersed, that the functions f, g, ω be C^1 in their arguments, that the divergence of ω vanish, and that the unperturbed flow be $d\mu_Y$-ergodic for almost all X. A closely related theorem proved under less restrictive hypotheses may be found in [**DGL**]; the brief article [**Du**] discusses how the methods used to prove this stronger theorem may be used to simplify the proof of Anosov's original theorem.

Anosov's theorem makes no mention of the rate at which the measure of the set of 'bad' initial conditions converges to 0 with ε, nor of the size of the difference between X and W for 'good' initial conditions. Estimating these quantities requires some measure of how fast the unperturbed flow is uniformly distributed over $\mathcal{M}$. One approach, pioneered mainly by Arnold [**Ar1**] and Neishstadt [**Ne1,2**], restricts to the case where $\mathcal{M}$ is the n-torus $\mathbb{T}^n \equiv \mathbb{R}^n/\mathbb{Z}^n$, and where the unperturbed frequency ω is independent of the fast variables Y; i.e. $\omega = \omega(X)$. In this case one may use Fourier analysis techniques, and the 'bad' initial conditions are identified as those too close to resonance, i.e. as those X where 'divisors' of the form $k \cdot \omega(X)$, $k \in \mathbb{Z}^n \backslash \{0\}$ are too small. With an appropriate nondegeneracy assumption on ω, it is possible to accurately estimate the measure of those X giving rise to small divisors and, taking into account solutions' passage through resonance, Neishtadt in fact achieves optimal order estimates in [**Ne2**]. This approach is obviously motivated by the important class of examples where the unperturbed vector field is Hamiltonian and integrable in the sense of Liouville (as defined in §49 of [**Ar2**]); yet the results and the proofs of [**Ar1**] and [**Ne1,2**] do not depend upon any such assumption and apply to more general situations.

In this article we obtain order estimates for a class of problems which is not more general than, but rather disjoint from that considered by Arnold and Neishtadt. Specifically, we do not assume that $\mathcal{M}$ is a torus, but instead that ω is independent of the *slow* variables ($\omega = \omega(Y)$), and we strengthen the uniform distribution hypothesis from ergodicity to mixing. In fact, we consider a class of mixing flows whose autocorrelation

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

functions decay sufficiently fast, or, alternately, the class of transitive Anosov flows on $\mathcal{M}$ (see assumptions (iv) and (iv′) below).

Our resulting estimates of the decay of the error with ε verify and extend the following conjectural statements due to Arnold [**Ar3**, §18-D pp. 160–161]:

> 'Neishtadt's result can be interpreted as an indication of the statistical independence of the increments of the deviation of I from J on subsequent time intervals of length one†. [. . .]
>
> The idea of independence of the deviation of I from J can apparently be justified much more completely when the fast motion is not quasi-periodic but is an Anosov system. This is indicated, in particular, by the central limit theorem‡ for functions on the phase space [. . .]
>
> This theorem justifies the probabilistic approach when both the slow and fast motions are independent of the slow variable [. . .]'

In contrast with M. Ratner's central limit theorem (to which Arnold refers above), our method allows a dependence of the slow motion on the slow variable. Moreover, even in the case where the slow motion does not depend on the slow variable, a further restriction must be imposed on the vector field in order to apply Ratner's central limit theorem; we shall briefly discuss this later. It appears that one should not seek to apply Ratner's result directly as suggested by Arnold, but instead use some of the authors' arguments adapted to the classical proofs of averaging theorems. In any case it is worth noticing that the results stated in the present paper do not follow from a direct application of Ratner's central limit theorem. But we shall see that in the case where the slow motion does not depend on the slow variable, our error estimate is consistent with the central limit theorem, which seems to indicate that our method recovers the optimal power-law estimate in terms of ε.

Very recently, we became aware of some closely related, independent work by Yu. Kifer [**Ki**]. Kifer's results hold for Anosov flows: he proves that the difference between the solution of the perturbed system and that of the averaged system, conveniently scaled, converges weakly to a diffusion process. Our results apply to the case of flows having the 'locally uniform bounded variance property' and not to Anosov flows only. Even in the case of Anosov flows, it does not seem that our results can be deduced from [**Ki**] in any straightforward manner.

2. *Set-up and statement of results*

We begin by describing the setting for our averaging theorems, which is closely related—but not identical—to the setting for Anosov's theorem.

† In this article, I and J are denoted respectively by X and W.

‡ Arnold refers to two announcements of the central limit theorem in special cases: one by Ya. Sinai (for geodesic flows on manifolds with constant negative curvature) and another by M. Ratner (for Anosov flows on three-dimensional manifolds). We have chosen not to cite these references, partly because they do not contain complete proofs, but mostly since both are superseded by Ratner's paper [**Ra2**] which takes advantage of the general constructions of Markov partitions for flows due to Bowen [**Bo**] and Ratner [**Ra1**]. Both prior announcements are listed in the reference section of [**Ra2**].

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

Let $\mathcal{M}$ be a smooth compact manifold of dimension n equipped with a smooth density $|dY|$ normalized to satisfy the relation $\int_{\mathcal{M}} |dY| = 1$. For small $\varepsilon \geq 0$, we consider the system of ODEs on $\mathbb{R}^m \times \mathcal{M}$

$$\frac{dX}{dt} = \varepsilon f(X, Y) \tag{1a}$$

$$\frac{dY}{dt} = \omega(Y) \tag{1b}$$

associated to the perturbation $(\varepsilon f, \omega)$ of the vector field $(0, \omega)$. Here X and Y denote elements of $\mathbb{R}^m$ and $\mathcal{M}$, respectively. We group the assumptions on (1) into four hypotheses, the first three of which are:

(i) $f \in C^1(\mathbb{R} \times \mathcal{M})$ and $\omega \in C^1(\mathcal{M})$;

(ii) the density $|dY|$ is invariant under the flow generated by ω;

(iii) $f(X, Y) = 0$ for all $X \notin B_R(0)$, the open ball of radius $R > 0$ centered on 0 in $\mathbb{R}^m$.

As discussed in the introduction, the goal of averaging theory is to approximate the X-components of solutions of (1) by solutions of the associated (slow) averaged system in $\mathbb{R}^m$:

$$\frac{dW}{dt} = \varepsilon \overline{f}(W) \equiv \varepsilon \int_{\mathcal{M}} f(W, Y)|dY|. \tag{2}$$

Standard ODE theory shows that, under hypotheses (i)–(iii), the systems (1) and (2) generate global flows which leave the respective domains $D_R \equiv \overline{B_R}(0) \times \mathcal{M}$ and $\overline{B_R}(0)$ invariant (here $\overline{B_R}(0)$ is the closed ball of radius R centered on the origin in $\mathbb{R}^m$). We shall denote these flows by $(X_t^\varepsilon, Y_t) \equiv (X_t^\varepsilon, Y_t)(x, y)$ and $W_t^\varepsilon \equiv W_t^\varepsilon(x)$ respectively. This notation means that $t \mapsto (X_t^\varepsilon, Y_t) \equiv (X_t^\varepsilon, Y_t)(x, y)$ is the integral curve of the vector field $(\varepsilon f, \omega)$ passing through the initial condition (x, y) at time $t = 0$. It is perhaps worth pointing out that the unperturbed flow is consistently denoted (X_t^0, Y_t).

Definition. For every $\varphi \in L^2(\mathcal{M}; \mathbb{C})$, the autocorrelation function C_φ of φ with respect to the flow Y_t is defined as

$$C_\varphi(t) = \int_{\mathcal{M}} \varphi(Y_t(y))\varphi(y)|dy| - \left(\int_{\mathcal{M}} \varphi(y)|dy| \right)^2.$$

With this notation, our final hypothesis (a mixing assumption on the system (1)) takes one of two alternate forms. The first form is:

(iv) for each bounded set B of $C^1(\mathcal{M})$, there exists $\alpha > 1$ such that

$$\mathbf{c}_\alpha(B) \equiv \sup_{t>0} \sup_{\varphi \in B} t^\alpha |C_\varphi(t)| < \infty.$$

The theorem below also holds if hypothesis (iv) is replaced by its second form:

(iv′) the flow Y_t is a C^2 transitive Anosov flow on $\mathcal{M}$ with a smooth invariant measure.

These alternate hypotheses are discussed below in Remark (1), and in the course of the proof.

Finally, we introduce the pointwise error E_t^ε between solutions of the perturbed system (1) and its associated averaged system (2), defined as

$$E_t^\varepsilon(x, y) = X_t^\varepsilon(x, y) - W_t^\varepsilon(x).$$

We may now state our results as Theorem A and Theorem B; Theorem A applies to general systems of the form (1a) and (1b).

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

THEOREM A. *Assume hypotheses (i)–(iv), or (i)–(iv′) pertaining to systems (1) and (2), and fix $T > 0$. Then the average maximum error*

$$\int_{D_R} \sup_{0\le t\le T/\varepsilon} |E_t^\varepsilon(x, y)|\, dx|dy| = O(\varepsilon^{1/3}) \quad \text{as } \varepsilon \to 0.$$

Theorem B states the result obtained by following the proof of Theorem A in the case where the slow motion (i.e. equation (1a)) does not depend on the slow variable X. In this case

$$X_t^\varepsilon(x, y) = x + \varepsilon \int_0^t f(Y_s(y))\, ds, \quad W_t^\varepsilon(x) = x + \varepsilon t \overline{f}$$

and

$$E_t^\varepsilon(y) = \varepsilon \int_0^t f(Y_s(y))\, ds - \varepsilon t \overline{f}.$$

THEOREM B. *Assume hypotheses (i), (ii) and (iv), or (i), (ii) and (iv′) pertaining to systems (1) and (2), and fix $T > 0$. If, moreover, the function f does not depend on the X variable, the maximum average error*

$$\sup_{0\le t\le T/\varepsilon} \int_{\mathcal{M}} |E_t^\varepsilon(y)||dy| = O(\varepsilon^{1/2}) \quad \text{as } \varepsilon \to 0.$$

Before stating some remarks concerning the two theorems above, let us briefly compare the result in Theorem B with Ratner's central limit theorem [**Ra2**] in the case where assumption (iv′) holds.

Assume that $f - \overline{f} = \widetilde{f}$ satisfies the 'homological condition':

$$f - \overline{f} = \widetilde{f} \notin \mathrm{Im}(\omega(Y) \cdot \nabla_Y), \tag{HC}$$

where $\omega(Y) \cdot \nabla_Y$ is viewed as an unbounded skew-adjoint operator on $L^2(\mathcal{M}, |dy|)$. (One often refers to the equation $\omega(Y) \cdot \nabla_Y S = \widetilde{f}$ as the 'homological equation' and to $\mathrm{Im}(\omega(Y) \cdot \nabla_Y)$ as the space of 'coboundaries'.) Under the assumption of condition (HC), Ratner's central limit theorem applied to our situation asserts the existence of a positive constant σ_f such that, for all fixed $t \in]0, T]$ and $\alpha > 0$,

$$\mathrm{meas}\{y \in \mathcal{M} \mid |E_{t/\varepsilon}^\varepsilon(y)| > (\alpha\sigma_f\sqrt{t})\varepsilon^{1/2}\} \to \mathrm{meas}(\mathcal{M})\sqrt{\frac{2}{\pi}} \int_\alpha^{+\infty} e^{-u^2/2}\, du \quad \text{as } \varepsilon \to 0^+ \tag{CL}$$

(see formula (1) of [**Ra2**]). Define

$$\rho_\varepsilon(z) = \mathrm{meas}\{y \in \mathcal{M} \mid \varepsilon^{-1/2}|E_{t/\varepsilon}^\varepsilon(y)| > z\}.$$

Then for all $t \in]0, T]$, as $\varepsilon \to 0$ one has

$$\frac{1}{\sqrt{\varepsilon}} \int_{\mathcal{M}} |E_{t/\varepsilon}^\varepsilon(y)||dy| = \int_0^{+\infty} \rho_\varepsilon(z)\, dz \to \mathrm{meas}(\mathcal{M}) \int_0^{+\infty} \sqrt{\frac{2}{\pi}} \int_{z/\sigma_f\sqrt{t}}^{+\infty} e^{-u^2/2}\, du\, dz$$

which is consistent with the result of Theorem B.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

This brief discussion shows that one should not attempt to use Ratner's result directly to establish either Theorem A or B. In fact, doing so would impose the restriction (HC) on $\widetilde{f}$ in the case where f does not depend on X and, even worse, impose it on $\widetilde{f}(X,\cdot)$ for all X, in the general case. This would be highly unnatural: one expects the final result to be easier to prove when $\widetilde{f}$ does not satisfy (HC), since it is precisely in this case that the homological equation can be solved.

In the case where assumption (iv′) is satisfied, the method we use to prove Theorems A and B borrows crucial estimates from [**Ra2**]. We have stated Theorem B and compared the error estimate it provides with (CL) under assumption (iv′) to show that our method is sharp (in the sense that it recovers the order of ε in the error computed by a direct application of Ratner's central limit theorem).

Remarks. (1) The alternate hypotheses (iv) and (iv′) are simply different ways of ensuring that the unperturbed flow has a certain property crucial to the proof below. This property, which we call 'bounded variance', is one ingredient in the central limit theorem, and we have chosen not to use it as the fourth hypothesis above because we believe Theorem A is more interesting as stated. On the other hand, we are unable to consolidate (iv) and (iv′) into a single assumption because their relationship appears to be unknown. We do know, however, that (iv) is satisfied by at least one important class of flows: according to Ratner [**Ra3**], the geodesic and horocycle flows on the unit tangent bundle of a compact surface with constant negative curvature have the property of exponential decay for the autocorrelation coefficients of smooth functions. It is apparently unknown whether geodesic flows in the case of varying sectional curvature satisfy (iv). We derive the bounded variance property from (iv) in §5 below, where it appears as a variant of the Green–Kubo formula of statistical physics. For flows satisfying (iv′) (i.e. for C^2 transitive Anosov flows on compact manifolds), the central limit theorem, and in particular the finite variance property, were proved by Ratner [**Ra1**], and carry over to the present case provided that certain uniformity considerations are taken into account (cf. §5 and the appendix below).

(2) As mentioned in the introduction, Theorem A above can be viewed as an amplification of Anosov's averaging theorem in the case of mixing flows in much the same way that Neishtadt's results [**Ne1,2**] are an amplification for the case of perturbations of quasi-periodic flows on flat tori. In the latter case, the 'rate of ergodization' is measured in terms of small divisors, and the error estimate depends on the measure of the set of nearly resonant frequencies. This is not so in the present case: the value 1 of the parameter α is 'critical' in assumption (iv), and the error estimate is not affected by the value of α when $\alpha > 1$.

(3) Theorem A above is not a consequence of the result of Freidlin and Wentzell on 'normal deviations from an averaged system' (cf. [**FW**, Chapter 7, §3]). In fact, Freidlin and Wentzell consider a case where the analogue of the flow Y_t is a random process with an adapted filtration of σ-algebras satisfying a strong mixing assumption which is reminiscent of the very weak Bernoulli property (cf. [**OW**]). By contrast, since we are not concerned with approximations by Brownian paths as in [**FW**], we require only the classical mixing property of ergodic theory (concerning two-point correlation for individual functions), stated in such a way as to ensure that the unperturbed flow has

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

the bounded variance property.

(4) In certain applications it would be desirable to allow ω to depend on X as well as Y. Our methods appear to permit this dependence, but at the expense of a badly deteriorated error estimate. Further study and perhaps other methods are needed to achieve a good result in the case $\omega = \omega(X, Y)$.

3. *Proof of Theorem A*

In order to examine the relation between solutions of (1a) and (2), we follow Neishtadt [**Ne2**] by introducing the smooth auxiliary function $S^\varepsilon \in C_0^1(\mathbb{R}^m \times \mathcal{M}; \mathbb{R}^m)$ (to be determined later) and the near-identity transformation $P = X + \varepsilon S^\varepsilon(X, Y)$.

Now integrating the difference $(d/dt)(X - W) = (d/dt)(P - W - \varepsilon S^\varepsilon)$ from 0 to t leads to

$$\begin{aligned} &E_t^\varepsilon(x, y) \\ &= X_t^\varepsilon(x, y) - W_t^\varepsilon(x) = \varepsilon \int_0^t (\overline{f}(X_u^\varepsilon(x, y)) - \overline{f}(W_u^\varepsilon(x)))\, du \\ &\quad -\varepsilon(S^\varepsilon(X_t^\varepsilon(x, y)), Y_t(y)) - S^\varepsilon(x, y) + \varepsilon \int_0^t R^\varepsilon(X_u^\varepsilon(x, y), Y_u(y))\, du, \end{aligned} \tag{3a}$$

where

$$R^\varepsilon = \varepsilon f \cdot \nabla_X S^\varepsilon + \omega \cdot \nabla_Y S^\varepsilon + \widetilde{f}, \quad \text{with } \widetilde{f} = f - \overline{f}. \tag{3b}$$

In estimating the error, the usual strategy is to demand that the $O(1)$ part of R^ε (i.e. $\omega \cdot \nabla_Y S^\varepsilon + \widetilde{f}$) vanishes, or is sufficiently small with ε, giving rise to a homological equation. Because of the presence of 0 in the spectrum of $\omega \cdot \nabla_Y$, solving the homological equation is problematic. Various approaches are possible; here we proceed to pose a regularized version of the homological equation by setting $R^\varepsilon = \lambda_\varepsilon S^\varepsilon$, where λ_ε is an order parameter ($0 \le \varepsilon \ll \lambda_\varepsilon$, where $\varepsilon \ll \lambda_\varepsilon$ means that $\varepsilon = o(\lambda_\varepsilon)$ as $\varepsilon \to 0$) to be chosen optimally later.

If the ansatz $R^\varepsilon = \lambda_\varepsilon S^\varepsilon$ is admitted provisionally, (3a) becomes

$$\begin{aligned} &E_t^\varepsilon(x, y) \\ &= \varepsilon \int_0^t (\overline{f}(X_u^\varepsilon(x, y)) - \overline{f}(W_u^\varepsilon(x)))\, du \\ &\quad -\varepsilon(S^\varepsilon(X_t^\varepsilon(x, y)), Y_t(y)) - S^\varepsilon(x, y) + \varepsilon\lambda_\varepsilon \int_0^t S^\varepsilon(X_u^\varepsilon(x, y), Y_u(y))\, du. \end{aligned} \tag{4}$$

In view of assumption (ii), it is easy to verify that the Jacobian of the reverse flow $(X_{-t}^\varepsilon, Y_{-t})$ is $O(1)$ over $O(\varepsilon^{-1})$-time intervals. We use this fact to ensure that the dilation or compression of the image measure $(X_t^\varepsilon, Y_t)_*(dx|dy|)$ is $O(1)$, bounded away from 0, on $O(\varepsilon^{-1})$-time intervals. More precisely, we deduce the existence of a constant $C > 1$ such that, given $T > 0$, for all $0 < \varepsilon \le 1$, the inequality

$$C^{-1}\, dx|dy| \le (X_t^\varepsilon, Y_t)_*(dx|dy|) \le C\, dx|dy| \tag{5}$$

holds for all $0 \le t \le T/\varepsilon$.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

Again following Neishtadt, we introduce $Z(t) = \int_{D_R} |E_t^\varepsilon(x, y)|dx|dy|$ (recall $D_R = \overline{B_R}(0) \times \mathcal{M}$). Then integrating (3a) over all initial conditions $(x, y) \in D_R$ and using (5), we obtain

$$Z(t) \leq L\varepsilon \int_0^t Z(u)\, du + C\varepsilon(2 + \lambda_\varepsilon t)\|S^\varepsilon\|_{L^1(D_R)} \tag{6}$$

(L is the X-Lipschitz constant for f, $\overline{f}$ and $\widetilde{f}$). We then use the following version of the Gronwall inequality:

$$0 \leq Z(t) \leq A + B\int_0^t Z(s)\, ds \Longrightarrow A + B\int_0^t Z(s)\, ds \leq Ae^{Bt}$$

to show that for $0 \leq t \leq T/\varepsilon$ the right-hand side of (6) is bounded by

$$C(2\varepsilon + \lambda_\varepsilon T)e^{LT}\|S^\varepsilon\|_{L^1(D_R)}. \tag{7}$$

Now returning to the estimate of the error, we take the supremum of the norm of (4) over $[0, T/\varepsilon]$, which gives

$$\begin{aligned}\sup_{0\leq t\leq T/\varepsilon} |E_t^\varepsilon(x, y)| \quad \leq \quad & \varepsilon \int_0^{T/\varepsilon} |\overline{f}(X_u^\varepsilon(x, y)) - \overline{f}(W_u^\varepsilon(x))|\, du \\ & + 2\varepsilon\|S^\varepsilon\|_{L^\infty(D_R)} + \varepsilon\lambda_\varepsilon \int_0^{T/\varepsilon} |S^\varepsilon(X_u^\varepsilon(x, y), Y_u(x, y))|\, du. \end{aligned} \tag{8}$$

Then integrating (8) over initial conditions D_R, extracting the Lipschitz constant L, using Fubini's theorem, and combining with (6) and (7), we arrive at

$$\int_{D_R} \sup_{0\leq t\leq T/\varepsilon} |E_t^\varepsilon(x, y)|\, dx|dy| \leq C(2\varepsilon + \lambda_\varepsilon T)e^{LT}\|S^\varepsilon\|_{L^1(D_R)} + 2\varepsilon(\text{meas } D_R)\|S^\varepsilon\|_{L^\infty(D_R)}. \tag{9}$$

The remainder of the proof consists of estimating $\|S^\varepsilon\|_{L^1(D_R)}$ and $\|S^\varepsilon\|_{L^\infty(D_R)}$; these norms are the subject of §5 below. We shall prove in §5 (estimates (11) and (17)) that

$$\|S^\varepsilon\|_{L^\infty(D_R)} \leq {\lambda_\varepsilon}^{-1}\|\widetilde{f}\|_{L^\infty(D_R)}$$

and that

$$\|S^\varepsilon\|_{L^1(D_R)} = O(\varepsilon {\lambda_\varepsilon}^{-2}) + O({\lambda_\varepsilon}^{-1/2}).$$

Inserting these estimates into (9), we find that the left-hand side of (9) is not greater than

$$\begin{aligned} & C(2\varepsilon + \lambda_\varepsilon T)[O(\varepsilon {\lambda_\varepsilon}^{-2}) + O({\lambda_\varepsilon}^{-1/2})]e^{LT} + 2\varepsilon {\lambda_\varepsilon}^{-1}\,\text{meas}(D_R)\|\widetilde{f}\|_{L^\infty(D_R)} \\ & = O(\varepsilon^2 {\lambda_\varepsilon}^{-2}) + O(\varepsilon {\lambda_\varepsilon}^{-1/2}) + O({\lambda_\varepsilon}^{1/2}) + O(\varepsilon {\lambda_\varepsilon}^{-1}). \end{aligned}$$

With $0 < \varepsilon \ll \lambda_\varepsilon$, taking the distinguished limit as $\varepsilon \to 0$ gives $\lambda_\varepsilon = O(\varepsilon^{2/3})$, which makes the last expression $O(\varepsilon^{1/3})$, as claimed. □

4. *Proof of Theorem B*

The proof of Theorem B parallels that of Theorem A, but is considerably simpler. Because f is independent of x, the average $\overline{f}$ of f over $\mathcal{M}$ is constant, so the first

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

integral on the right-hand side of (3a) vanishes identically. Therefore, taking $L^1(\mathcal{M})$-norms of both sides of (4), then using (5) leads to

$$\int_{\mathcal{M}} |E_t^\varepsilon(y)||dy| \le C\varepsilon(2+\lambda_\varepsilon t)\|S^\varepsilon\|_{L^1(\mathcal{M})},$$

and so

$$\sup_{0\le t\le T/\varepsilon} \int_{\mathcal{M}} |E_t^\varepsilon(y)||dy| \le C(2\varepsilon+\lambda_\varepsilon T)\|S^\varepsilon\|_{L^1(\mathcal{M})}. \tag{9′}$$

We then replace (17) of §5 by the appropriate estimate when f is independent of X. This estimate, which holds under the assumptions of Theorem B, is

$$\|S^\varepsilon\|_{L^1(\mathcal{M})} = O({\lambda_\varepsilon}^{-1/2})$$

(listed as (17′) in §5 below). Inserting this into (9′) above and taking the distinguished limit $\lambda_\varepsilon = \varepsilon$ gives precisely the result stated in Theorem B. □

5. *The homological equation*

5.1. *Reduction to the bounded variance property.*

According to our ansatz, we have

$$\lambda_\varepsilon S^\varepsilon(x,y) = R^\varepsilon(x,y) = \tilde{f}(x,y) + \frac{d}{dt} S^\varepsilon((X_t^\varepsilon, Y_t)(x,y))|_{t=0}$$

which we solve as

$$S^\varepsilon(x,y) = \int_0^\infty e^{-\lambda_\varepsilon t}\, \tilde{f}(X_t^\varepsilon, Y_t)(x,y)\, dt. \tag{10}$$

This leads immediately to the first required estimate:

$$\|S^\varepsilon\|_{L^\infty(D_R)} \le \|\tilde{f}\|_{L^\infty(D_R)} \int_0^\infty e^{-\lambda_\varepsilon t}\, dt = {\lambda_\varepsilon}^{-1}\|\tilde{f}\|_{L^\infty(D_R)}. \tag{11}$$

The estimate of $\|S^\varepsilon\|_{L^1(D_R)}$ is more crucial, as it depends on the bounded variance of the flow Y_t (and so ultimately on the statistical properties of Y_t as embodied in assumption (iv) or (iv′)).

We begin by rewriting (10) as

$$\begin{aligned} S^\varepsilon(x,y) &= \int_0^\infty e^{-\lambda_\varepsilon t}(\tilde{f}(X_t^\varepsilon, Y_t)(x,y) - \tilde{f}(x, Y_t(y)))\, dt + \int_0^\infty e^{-\lambda_\varepsilon t}\tilde{f}(x, Y_t(y))\, dt \\ &= I_1(x,y) + I_2(x,y). \end{aligned} \tag{12}$$

The first of these terms is easily estimated, since

$$\begin{aligned} |I_1(x,y)| &\le L\int_0^\infty e^{-\lambda_\varepsilon t}|X_t^\varepsilon(x,y) - x|\, dt \\ &\le L\int_0^\infty e^{-\lambda_\varepsilon t}\varepsilon t\|f\|_{L^\infty(D_R)}\, dt = \varepsilon{\lambda_\varepsilon}^{-2}L\|f\|_{L^\infty(D_R)} \end{aligned}$$

implies that

$$\|I_1\|_{L^1(D_R)} \le \varepsilon{\lambda_\varepsilon}^{-2}L(\text{meas } D_R)\|f\|_{L^\infty(D_R)}. \tag{13}$$

The term I_2 is controlled in terms of the 'variance' of $\tilde{f}$, which we define below.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

Definition. Let φ be a continuous function defined on $\mathcal{M}$. The variance of φ over the time interval $[0, T]$ along the flow Y_t is

$$V_\varphi(T) = \int_{\mathcal{M}} \left(\frac{1}{\sqrt{T}} \int_0^T \varphi(Y_t(y))\, dt - \sqrt{T}\overline{\varphi} \right)^2 |dy|,$$

where $\overline{\varphi}$ is the mean value of φ:

$$\overline{\varphi} = \int_{\mathcal{M}} \varphi(y)|dy|$$

(recall that $|dy|$ is normalized so that $\int_{\mathcal{M}} |dy| = 1$).

The way information on the variance of $\widetilde{f}$ is translated into information on I_2 is via an integral analogue of Abel's limit theorem [**Ah**, p. 41]. We write

$$\lambda_\varepsilon \int_{\mathcal{M}} |I_2(x, y)|^2 |dy| = \int_{\mathcal{M}} \left({\lambda_\varepsilon}^{1/2} \int_0^\infty e^{-\lambda_\varepsilon t} \widetilde{f}(x, Y_t(y))\, dt \right)^2 |dy|. \tag{14}$$

Now integrating by parts inside the parentheses, noting that the boundary term vanishes, and making the change of variable $\vartheta = \lambda_\varepsilon t$, this becomes

$$\lambda_\varepsilon \int_{\mathcal{M}} |I_2(x, y)|^2 |dy| = \int_{\mathcal{M}} \left(\int_0^\infty \left(\frac{\lambda_\varepsilon}{\vartheta} \right)^{1/2} \int_0^{\vartheta/\lambda_\varepsilon} \widetilde{f}(x, Y_t(y))\, ds\, \vartheta^{1/2} e^{-\vartheta}\, d\vartheta \right)^2 |dy|. \tag{15}$$

Then using the Cauchy–Schwarz inequality (with respect to the measure $\vartheta^{1/2} e^{-\vartheta}\, d\vartheta$) and the fact that for all $x \in \overline{B}_R(0)$, $\widetilde{f}(x, \cdot)$ has zero mean by construction, the last expression is bounded by

$$\begin{aligned} &\int_{\mathcal{M}} \left(\int_0^\infty \vartheta^{1/2} e^{-\vartheta}\, d\vartheta \right) \left(\int_0^\infty \left[\left(\frac{\lambda_\varepsilon}{\vartheta} \right)^{1/2} \int_0^{\vartheta/\lambda_\varepsilon} \widetilde{f}(x, Y_t(y))\, ds \right]^2 \vartheta^{1/2} e^{-\vartheta}\, d\vartheta \right) |dy| \\ &\quad = \Gamma(3/2) \int_0^\infty \left(\int_{\mathcal{M}} \left[\left(\frac{\lambda_\varepsilon}{\vartheta} \right)^{1/2} \int_0^{\vartheta/\lambda_\varepsilon} \widetilde{f}(x, Y_t(y))\, dt \right]^2 |dy| \right) \vartheta^{1/2} e^{-\vartheta}\, d\vartheta \\ &\quad = \Gamma(3/2) \int_0^\infty V_{\widetilde{f}(x,\cdot)} \left(\frac{\vartheta}{\lambda_\varepsilon} \right) \vartheta^{1/2} e^{-\vartheta}\, d\vartheta. \end{aligned} \tag{16}$$

This calculation shows that eventually everything depends upon the following.

Bounded variance property. We say that the flow Y_t defined on $\mathcal{M}$ has:

(i) the bounded variance property (abbreviated BVP) if and only if, for all $\varphi \in C^1(\mathcal{M})$, there exists $K(\varphi) > 0$ such that for all $T > 0$, $V_\varphi(T) \le K(\varphi)$;

(ii) the locally uniform bounded variance property (abbreviated LUBVP) if and only if, for any bounded subset $\mathcal{S}$ of $C^1(\mathcal{M})$, there exists $K(\mathcal{S}) > 0$ such that for all $T > 0$ and all $\varphi \in \mathcal{S}$, $V_\varphi(T) \le K(\mathcal{S})$.

The inequalities above lead directly to the following.

LEMMA 1. *Assume hypotheses (i) and (ii) from §2. Then:*

(1) *if f does not depend on X and the flow Y_t on $\mathcal{M}$ has the BVP, then the function S^ε given by (10) satisfies*

$$\|S^\varepsilon\|_{L^1(\mathcal{M})} \le {\lambda_\varepsilon}^{-1/2} \sqrt{\frac{\pi}{4} K(\widetilde{f})\, \mathrm{meas}(\mathcal{M})}; \tag{17′}$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

(2) *if f satisfies (iii) and the flow Y_t on $\mathcal{M}$ has the LUBVP, then the function S^ε given by (10) satisfies*

$$\|S^\varepsilon\|_{L^1(D_R)} \le {\lambda_\varepsilon}^{-1/2} \operatorname{meas}(B_R(0)) \sqrt{\frac{\pi}{4} K_f \operatorname{meas}(\mathcal{M})} + \varepsilon {\lambda_\varepsilon}^{-2} L(\operatorname{meas} D_R) \|f\|_{L^\infty(D_R)}, \tag{17}$$

where $K_f = K(\{\widetilde{f}(x,\cdot),\ x \in \overline{B}_R(0)\})$.

Proof. (1) According to the BVP, the right-hand side of (16) is bounded by

$$K(\widetilde{f})\Gamma(3/2) \int_0^\infty \vartheta^{1/2} e^{-\vartheta}\, d\vartheta = K(\widetilde{f})\Gamma(3/2)^2 = K(\widetilde{f})\pi/4.$$

Hence

$$\lambda_\varepsilon \|I_2\|^2_{L^2(\mathcal{M})} \le \frac{\pi}{4} K(\widetilde{f})$$

from which (17′) follows after applying the Cauchy–Schwarz inequality, since the term I_1 in the decomposition (12) is zero because f is independent of X.

(2) Since $f \in C^1(\mathcal{M})$, the set $\{\widetilde{f}(x,\cdot),\ x \in \overline{B}_R(0)\}$ is bounded in $C^1(\mathcal{M})$, so that $K_f = K(\{\widetilde{f}(x,\cdot),\ x \in \overline{B}_R(0)\})$ is well-defined assuming that Y_t has the LUBVP. As in the proof of (1), after integrating both sides of (16) over $B_R(0)$, the LUBVP shows that

$$\lambda_\varepsilon \|I_2\|^2_{L^2(D_R)} \le \frac{\pi}{4} \operatorname{meas}(B_R(0)) K_f$$

and hence

$$\|I_2\|^2_{L^1(D_R)} \le {\lambda_\varepsilon}^{-1/2} \sqrt{\frac{\pi}{4} \operatorname{meas}(B_R(0)) \operatorname{meas}(D_R) K(\widetilde{f})}.$$

Combining this with (12) and (13) gives the announced result. □

5.2. *Proof of the bounded variance property.* Now it follows from the previous sections that the proof of Theorem A depends upon verifying that the flow Y_t has the LUBVP, whereas Theorem B depends upon verifying only that it has the BVP.

PROPOSITION 2. *Assume that ω satisfies assumption (i), and that (ii) and (iv) hold. Then the flow Y_t has the LUBVP (hence also the BVP).*

Proof. Let $\mathcal{S}$ be a bounded subset of $C^1(\mathcal{M})$. Without loss of generality, we can assume that each function $\varphi \in \mathcal{S}$ is of zero mean value on $\mathcal{M}$ and that the parameter α in assumption (iv) satisfies $1 < \alpha < 2$. We write

$$\begin{aligned} V_\varphi(T) &= \int_{\mathcal{M}} \left(\frac{1}{\sqrt{T}} \int_0^T \varphi(Y_t(y))\, dt \right)^2 |dy| \\ &= \int_{\mathcal{M}} \frac{2}{T} \int_0^T \int_0^t \varphi(Y_t(y)) \varphi(Y_s(y))\, ds\, dt |dy| \\ &= \frac{2}{T} \int_0^T \int_0^t \int_{\mathcal{M}} \varphi(Y_t(y)) \varphi(Y_s(y)) |dy|\, ds\, dt \\ &= \frac{2}{T} \int_0^T \int_0^t \int_{\mathcal{M}} \varphi(Y_{t-s}(y)) \varphi(y) |dy|\, ds\, dt. \end{aligned}$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

Since $\int_{\mathcal{M}} \varphi(y)|dy| = 0$ for all $\varphi \in \mathcal{S}$, the last expression becomes

$$\begin{aligned}\frac{2}{T}\int_0^T \int_0^t C_\varphi(t-s)\,ds\,dt &= \frac{2}{T}\int_0^T \int_0^t C_\varphi(\tau)\,d\tau\,dt \\ &= \frac{2}{T}\int_0^T C_\varphi(\tau)\,d\tau \int_\tau^T dt \\ &= 2\int_0^T C_\varphi(s)\,ds - \frac{2}{T}\int_0^T sC_\varphi(s)\,ds.\end{aligned}$$

Assumption (iv) therefore shows that

$$|V_\varphi(T)| \le 4\int_0^T |C_\varphi(s)|\,ds \le 4\,\mathrm{meas}(\mathcal{M})\|\phi\|^2_{L^\infty(\mathcal{M})} + \frac{4\mathbf{c}_\alpha}{\alpha-1},$$

which gives the LUBVP with

$$K(\mathcal{S}) = 4\,\mathrm{meas}(\mathcal{M})\|\phi\|^2_{L^\infty(\mathcal{M})} + \frac{4\mathbf{c}_\alpha}{\alpha-1}. \qquad \square$$

When assumption (iv′) holds, Theorem 3.1 of [**Ra2**] implies that the flow Y_t has the BVP. In fact, more is true, and we have the following.

PROPOSITION 3. *Assume that ω satisfies (i) and that (ii) and (iv′) hold. Then the flow Y_t has the LUBVP.*

None of the theorems stated in [**Ra2**] implies Proposition 3; however, Proposition 3 can be proved by using Ratner's estimates and checking carefully that the bounding constants involved are locally uniform in the $C^1(\mathcal{M})$ topology. In fact, since the bounded variance property is a weaker statement than the central limit theorem, this local uniformity does not need to be checked in each estimate of Ratner's article; rather, the relevant estimates have to be extracted from [**Ra2**] or established along the same lines. While we assume that a specialist of ergodic theory would encounter no difficulty in verifying Proposition 3, we believe that a complete proof of Theorems A and B should be accessible to readers not necessarily familiar with the symbolic coding of Anosov flows. For this reason, and since the question of uniformity is crucial for our results, we give a complete proof of Proposition 3 in the appendix below.

Acknowledgements. H. S. Dumas was supported by the National Science Foundation under Grants DMS–9106812 and DMS–9505017. F. Golse was supported by the Alfred P. Sloan Foundation during his stay at the Institute for Advanced Study.

Appendix. Proof of Proposition 3

The proof of Proposition 3 follows some (not all) of the steps in Ratner's proof of the central limit theorem [**Ra2**]. For the most part we shall adopt the notation in [**Ra2**]. We begin by recalling the basic results of symbolic dynamics for flows in the form needed here; we refer to Bowen [**Bo**] for proofs of these results.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

A.1. *Symbolic dynamics.* We first recall some material on the coding of Anosov flows and introduce some notation used in what follows.

By assumption (iv′), the flow Y_t defined on $\mathcal{M}$ is a C^2 transitive Anosov flow. By assumption (ii), Y_t has a smooth invariant measure, the Riemannian volume element of $\mathcal{M}$, denoted either by $|dy|$ or (as a measure) by m. Such flows exist (e.g. the geodesic flow on the unit tangent bundle of a compact surface of negative curvature).

Finite Markov partitions with parallelepipeds of arbitrarily small diameters have been constructed for Anosov flows by Bowen [**Bo**] and Ratner [**Ra1**]. These partitions provide the following elements for a representation of the flow Y_t.

(i) A matrix $\pi \in M_r(\{0,1\})$ (where r is the cardinality of the Markov partition) and a set

$$X_\pi = \{x \in \{1,\dots,r\}^{\mathbb{Z}} \mid \pi_{x_i x_{i+1}} = 1, \forall i \in \mathbb{Z}\},$$

equipped with the metric

$$\delta(x,x') = \sum_{i\in\mathbb{Z}} 2^{-|i|}\mathbf{e}(x_i,x_i'),$$

where $\mathbf{e}(z,z') = 0$ if $z = z'$ and $\mathbf{e}(z,z') = 1$ if $z \neq z'$. That the flow Y_t is topologically transitive is reflected by the existence of a positive integer s such that π^s has strictly positive entries.

(ii) The shift operator ϕ defined on X_π by $(\phi x)_i = x_{i-1}$ for all $i \in \mathbb{Z}$. Moreover, there exists a positive Hölder continuous function l defined on X. This allows us to construct the 'special flow' S_t as follows. We first introduce the suspension of X_π by l, defined as the quotient space

$$W = \{(x,y) \in X_\pi \times \mathbb{R}^+ \mid 0 \le y \le l(x)\}/\sim,$$

where the equivalence $\sim$ is given by $(x, l(x)) \sim (\phi x, 0)$. The space W is equipped with the quotient topology defined by the metric

$$d([(x,y)],[(x',y')]) = \inf_{\substack{(x_1,y_1)\sim(x,y);\\(x_1',y_1')\sim(x',y')}} (\delta(x_1,x_1') + |y_1 - y_1'|)$$

(where the notation $[(x,y)]$ designates the equivalence class of (x,y) modulo $\sim$). The special flow S_t is then defined for $t < \inf_{x\in X_\pi} l(x)$ by

$$\begin{aligned} S_t(x,y) &= (x, y+t), \quad \text{if } y + t < l(x),\\ S_t(x,y) &= (\phi x, y + t - l(x)), \quad \text{if } y + t \ge l(x). \end{aligned}$$

(iii) A continuous surjective mapping $\psi : W \to \mathcal{M}$ and a Borel measure ν on W (invariant under the special flow S) such that $\psi \circ S_t = Y_t \circ \psi$ for all $t \in \mathbb{R}$, and such that $\psi_*\nu = m$. Moreover, the set on which ψ fails to be one-to-one has ν-measure 0, and the dynamical systems (W, ν, S_t) and $(\mathcal{M}, m, Y_t)$ are metrically isomorphic.

(iv) A ϕ-invariant probability measure μ on X_π such that $d\nu = (d\mu(x) \times dy)/\bar{l}$, where $\bar{l} = \int_X l(x)\,d\mu(x)$. The dynamical system (X_π, μ, ϕ) is a K-automorphism with strong mixing of type $\Upsilon_{\gamma,\alpha}$ for some constants $0 < \gamma < 1$ and $\alpha > 0$. Strong mixing of this type is defined as follows. Introduce the partition $\mathcal{P} = \{U_1, U_2, \dots, U_r\}$ defined by

$$U_k = \{x \in X_\pi \mid x_0 = k\}, \quad \text{for } 1 \le k \le r.$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

For any pair of integers a and b such that $a \leq b$, we denote $\phi^{-a}\mathcal{P} \vee \cdots \vee \phi^{-b}\mathcal{P}$ by $\mathcal{P}_a^b$. That (X_π, μ, ϕ) is a K-automorphism with strong mixing of type $\Upsilon_{\gamma,\alpha}$ means that there exists a $C_0 > 0$ such that, for all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$, for any family of pairwise disjoint sets $B_i \in \mathcal{P}_{k+n}^{\infty}$ and any $A \in \mathcal{P}_{-\infty}^{k}$, we have

$$\sum_i |\mu(B_i \cap A) - \mu(A)\mu(B_i)| \leq C_0 \mu(A) \gamma^{n^\alpha}.$$

Before going further, we introduce the following classes of functions.

Definition and notation. Let $A > 0$, $\rho \in]0, 1[$ and $\kappa > 0$. We say that

— the function $F : X_\pi \to \mathbb{R}$ belongs to the class $\mathcal{F}_{\rho,\kappa,A}(X)$ iff $|F(x) - F(x')| \leq A\rho^{n^\kappa}$ for all $n \geq 0$ and all pairs of sequences (x, x') such that $x_i = x'_i$ for $|i| \leq n$;

— the function $f : W \to \mathbb{R}$ belongs to the class $\mathcal{F}_{\rho,\kappa,A}(W)$ iff the function defined on X_π by $x \mapsto \int_0^{l(x)} f(x, y)\, dy$ belongs to $\mathcal{F}_{\rho,\kappa,A}(X)$.

(v) There exist constants $A > 0$, $\rho \in]0, 1[$ and $\kappa > 0$ such that the function l defined in (ii) belongs to $\mathcal{F}_{\rho,\kappa,A}(X)$. (This follows from the Hölder continuity of l stated in (ii); one can even take $\kappa = 1$.)

Most of this is actually taken from the introduction of [**Ra2**, pp. 182–183]. However, the following statement is not to be found explicitly in [**Ra2**]†, but can be extracted from [**Bo**] (see Lemma 2.2(i), p. 435 of [**Bo**] and the definition of the map ρ (our ψ) p. 436).

(vi) The map ψ defined in (iii) is Hölder continuous (with respect to the metric d defined on W in (ii) and the geodesic metric on $\mathcal{M}$).

A.2. *Reduction to the case of a discrete time variable.*

Notation. For any function $f : W \to \mathbb{R}$, we shall denote $F(x) = \int_0^{l(x)} f(x, y)\, dy$.

Since X_π is compact and l is continuous and positive on X_π, there exist $0 < l_1 < l_2$ such that for all $x \in X_\pi$, $l_1 \leq l(x) \leq l_2$.

Let $f \in \mathcal{F}_{\rho,\kappa,A}(W) \cap L^\infty(W, \nu)$ be such that $\int_W f\, d\nu = 0$. Define

$$V_f(T) = \int_W \left(\frac{1}{\sqrt{T}} \int_0^T f(S_t(x, y))\, dt \right)^2 d\nu(x, y). \tag{A.1}$$

We shall see shortly that the LUBVP is simply a particular case of Lemma 3.2 of [**Ra2**], but with explicit dependence of the bounding constants on the parameters ρ, κ and A. Showing this, though technical, is necessary in order to obtain the local uniformity needed here.

LEMMA A.1. *Let $F \in \mathcal{F}_{\rho,\kappa,C_1}(X_\pi) \cap L^\infty(X_\pi, \mu)$ be such that $\int_{X_\pi} F\, d\mu = 0$. Then*

$$\int_{X_\pi} \left| \sum_{i=1}^{n} F \circ \phi^{-i} \right|^4 d\mu \;\leq\; C_2(F) n^2.$$

In the inequality above, $C_2(F) = (C_3^4 + 4C_1C_3^3\Sigma_1 + C_0C_3^4\Sigma_2)$, where

$$\Sigma_1 = \sum_{i \geq 0} 2^{4i+7} \rho^{[2^{i\alpha/2}]^\kappa} \quad \text{and} \quad \Sigma_2 = \sum_{i \geq 0} 2^{4i+7} r^{6[2^{i\alpha/2}]+3} \gamma^{(2^{i\alpha})},$$

and where C_0, γ and α are the constants in (iv) of §A.1 above, and $C_3 = \|F\|_{L^\infty(X_\pi)}$.

† Although it is used implicitly in Remark (ii) p. 192 of [**Ra2**].

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

With Lemma A.1 above, we can state the following lemma, which contains the key to Proposition 3.

LEMMA A.2. *There exists a continuous positive function $\mathcal{K}$ on $\mathbb{R}_+^* \times \mathbb{R}^+ \times \mathbb{R}^+$ such that, for all $f \in \mathcal{F}_{\rho,\kappa,A}(W) \cap L^\infty(W, \nu)$ such that $\int_W f\, d\nu = 0$, one has*

$$|V_f(T)| \leq \mathcal{K}(T, \|f\|_{L^\infty(W,\nu)}, C_2(F)),$$

where $C_2(F)$ is the expression defined in Lemma A.1.

We defer the rather technical proof of Lemma A.1 to §A.3 below. Let us briefly indicate the following proof.

Proof of Proposition 3. Let $\mathcal{S}$ be a bounded set in $C^1(\mathcal{M})$; it follows from (vi) above that there exists $A(\mathcal{S}) > 0$, $\rho \in]0, 1[$ and $\kappa > 0$ such that

$$\mathcal{S} \circ \psi = \{g \circ \psi \mid g \in \mathcal{S}\} \subset \mathcal{F}_{\rho(\mathcal{S}),\kappa(\mathcal{S}),A(\mathcal{S})}. \tag{A.2}$$

Without loss of generality, we can assume that all functions in $\mathcal{S}$ have zero mean on $\mathcal{M}$. Applying Lemma A.2 shows that, for all $T > 0$ and for all $g \in \mathcal{S}$,

$$|V_g(T)| \leq \sup_{f \in \mathcal{S}\circ\psi} \mathcal{K}(T, \|f\|_{L^\infty(W,\nu)}, C_2(F)),$$

which gives the LUBVP as announced. (The fact that the right-hand side of the inequality above might not be bounded as $T \to 0^+$ causes no harm, as the variance can easily be controlled by the $L^\infty(\mathcal{M})$ norm for $T \in [0, 1]$.) □

We now give the following proof.

Proof of Lemma A.2. First observe that

$$\begin{aligned}\left|V_f(T) - \int_W \left(\frac{1}{\sqrt{T}}\int_0^T f(S_t(x,0))\,dt\right)^2 d\nu(x,y)\right| \\ \leq \frac{4\|f\|_\infty^2 l_2^2}{T} + 2\sqrt{\frac{4\|f\|_\infty^2 l_2^2}{T}}\left(\int_W \left(\frac{1}{\sqrt{T}}\int_0^T f(S_t(x,0))\,dt\right)^2 d\nu(x,y)\right)^{1/2}.\end{aligned} \tag{A.3}$$

We introduce

$$n(t,x) = \sup\{k \in \mathbb{N} \mid l(x) + l(\phi^{-1}x) + \cdots + l(\phi^{-k}x) \leq t\}, \tag{A.4}$$

and

$$D(t,x) = \frac{1}{\sqrt{T}} \sum_{i=0}^{n(t,x)} F(\phi^{-i}x). \tag{A.5}$$

One easily sees that

$$\begin{aligned}\left|\int_W \left(\frac{1}{\sqrt{T}}\int_0^T f(S_t(x,0))\,dt\right)^2 d\nu(x,y) - \int_W D(t,x)^2\,d\nu\right| \\ \leq l_2^2\|f\|_\infty^2 + 2l_2\|f\|_\infty\left(\int_W D(t,x)^2\,d\nu(x,y)\right)^{1/2}.\end{aligned} \tag{A.6}$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

To prove that f has the bounded variance property, it therefore suffices to bound

$$\int_W D(t,x)^2\,d\nu(x,y).$$

To do this, we follow Ratner [**Ra2**, §3] by introducing

$$\xi(t,x)=\frac{1}{\sqrt{t}}\left(n(t,x)-\frac{t}{\bar{l}}\right) \tag{A.7}$$

and the sets $A_{kt}=\{x \mid k\le \xi(t,x)<k+1\}$ for all $t>0$ and $k\in\mathbb{Z}$. It is easy to check that $X=\bigcup_{k\in\mathbb{Z}} A_{kt}$ for all $t>0$. Define $s(t,x)=s_k(t,x)$ for $x\in A_{kt}$, where

$$s_k(t,x)=\frac{1}{\sqrt{t}}\sum_{i=0}^{[t/\bar{l}+k\sqrt{t}]} F(\phi^{-i}x). \tag{A.8}$$

As in equation (20) of [**Ra2**], we have

$$\left|\int_W D(T,x)^2\,d\nu-\int_W s(T,x)^2 d\nu\right|\le\int_W |D(T,x)^2-s(T,x)^2|\,d\mu_l(x), \tag{A.9}$$

where $d\mu_l(x)=l(x)\,d\mu(x)/\bar{l}$, and it is easy to see that

$$\int_W |D(T,x)^2-s(T,x)^2|\,d\mu_l(x)\le 4\|F\|_\infty^2+4\|F\|_\infty\left(\int_X s(T,x)^2\,d\mu_l(x)\right)^2. \tag{A.10}$$

We have therefore reduced the problem of showing that f has bounded variance to that of producing a bound on

$$\int_X s(T,x)^2\,d\mu_l(x).$$

We then mimic the argument in equation (22) of [**Ra2**]:

$$\begin{aligned}\int_X s(T,x)^2\,d\mu_l(x)&\\ &\le \sum_{k\in\mathbb{Z}}\int_{k\le\xi<k+1} s_k(T,x)^2\,d\mu_l(x)\\ &\le \sum_{k\in\mathbb{Z}}\left(\int_{k\le\xi<k+1} d\mu_l(x)\right)^{1/2}\left(\int_{k\le\xi<k+1} s_k(T,x)^4\,d\mu_l(x)\right)^{1/2}.\end{aligned} \tag{A.11}$$

For any positive integer r, there exists $I_r>0$ such that

$$\int_{k\le\xi<k+1} d\mu_l(x)\le\frac{I_r}{k^{2r}}, \tag{A.12}$$

and

$$\int_{k\le\xi<k+1} s_k(T,x)^4\,d\mu_l(x)\le\frac{l_2}{\bar{l}}E_\mu[s_k(T,\cdot)^4], \tag{A.13}$$

where E_μ denotes the expectation with respect to the probability measure μ.

Applying Lemma A.1 leads to

$$E_\mu[s_k(T,\cdot)^4]\le\frac{1}{T^2}C_2(F)\left(\frac{T}{\bar{l}}+k\sqrt{T}\right)^2=\frac{l_2}{\bar{l}}C_2(F)\left(\frac{1}{\bar{l}}+\frac{k}{\sqrt{T}}\right)^2. \tag{A.14}$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

Since the integration is performed on the set defined by $k \le \xi(T,x) < k+1$, it follows that

$$\left(\frac{1}{\bar{l}} + \frac{k}{\sqrt{T}}\right) \le \frac{n(T,x)}{T} \le \frac{1}{l_1}, \tag{A.15}$$

and finally

$$\int_X s(T,x)^2 \, d\mu_l(x) \le \left(\frac{l_2}{\bar{l}} C_2(F) \frac{1}{l_1^2}\right)^{1/2} \sum_{k\in\mathbb{Z}} \frac{I_r}{|k|^r}. \tag{A.16}$$

Gathering together (A.3), (A.6), (A.9), (A.10) and (A.16) establishes the bound announced in Lemma A.2. □

A.3. *Proof of Lemma A.1.* Let $F \in \mathcal{F}_{\rho,\kappa,C_1}(X_\pi) \cap L^\infty(X_\pi,\mu)$ ($0<\rho<1$, $\kappa>0$, $C_1>0$) be such that $\int_{X_\pi} F(x)\,d\mu(x) = 0$. Throughout this section, we shall denote by E the expectation with respect to the probability measure μ on X_π. We set

$$S_F^n(x) = \sum_{i=1}^{n} F(\varphi^{-i}x)$$

and we wish to estimate $E(|S_F^n|^4)$. Thus

$$\begin{aligned} E(|S_F^n|^4) &= \sum_{1\le k_1,k_2,k_3,k_4\le n} E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x)) \\ &= \sum_{i\ge 0} \sum_{2^i \le e(k,A) < 2^{i+1}} E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x)) \\ &\quad + \sum_{k\in A} E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x)), \end{aligned} \tag{A.17}$$

where we have used Ratner's notation:

$$\begin{aligned} A &= \{k=(k_1,k_2,k_3,k_4) \mid \forall i \in \{1,\dots,4\}\ \exists j \in \{1,\dots,4\},\ j\ne i,\ \text{s.t. } k_i = k_j\} \\ &= \{(k_1,k_2,k_3,k_4) \mid 1\le k_1\le n\} \cup \{(k_1,k_2,k_3,k_4) \mid 1\le k_1,k_3\le 3,\ k_1\ne k_3\}, \\ e(k,k') &= \sup_{1\le i\le 4} |k_i - k_i'|, \quad e(k,A) = \inf_{k'\in A} e(k,k'), \\ \sharp A &= n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} \le n^2. \end{aligned} \tag{A.18}$$

Now let k be such that $e(k,A) \in [2^i, 2^{i+1})$. We claim that there exists $j \in \{1,\dots,4\}$ such that given $l \ne j$, $l \in \{1,\dots,4\}$, we have $|k_j - k_l| \ge 2^i$. To see this, assume the contrary; assume that

$$\forall j \in \{1,\dots,4\},\ \exists l \in \{1,\dots,4\},\ l\ne j,\ \text{s.t. } |k_i - k_i'| < 2^i.$$

Without loss of generality, we reorder the components of k so that $k_1 \le k_2 \le k_3 \le k_4$. Then necessarily $k_2 - k_1 < 2^i$ and $k_4 - k_3 < 2^i$, so that $e(k,(k_1,k_1,k_3,k_3)) < 2^i$, which is absurd, since $(k_1,k_1,k_3,k_3) \in A$ and $e(k,A) \ge 2^i$.

An elementary counting procedure yields the following crude estimate:

$$\sharp\{k \mid 2^i \le e(k,A) < 2^{i+1}\} \le n^2 2^{4i+7}, \tag{A.19}$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

so that using (A.17) results in

$$\begin{aligned} & E(|S_F^n|) \\ \leq \quad & n^2 \sup_{k\in A} |E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x))| \\ & + n^2 \sum_{i\geq 0} 2^{4i+7} \sup_{2^i \leq e(k,A)<2^i+1} |E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x))|. \end{aligned} \tag{A.20}$$

It remains only to estimate

$$\sup_{2^i \leq e(k,A)<2^{i+1}} |E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x))|.$$

As in [**Ra2**, §1], we associate to each $F \in \mathcal{F}_{\rho,\kappa,C_1}$ the function F_m defined by

$$F_m \equiv F_m(x_{-m}, \dots, x_m)$$
$$\int_X (F - F_m)(x)\phi(x_{-m}, \dots, x_m)\, d\mu(x) = 0 \quad \forall \phi \equiv \phi(x_{-m}, \dots, x_m).$$

In other words, $F_m = E(F|\sigma(\mathcal{P}_{-m}^m))$, where $\sigma(\mathcal{P})$ is the σ-algebra generated by $\mathcal{P}$. Now

$$F \in \mathcal{F}_{\rho,\kappa,C_1} \Longrightarrow \|F - F_m\|_{L^2} \leq C_1 \rho^{(m^\kappa)} \tag{A.21}$$

(see [**Ra2**, equation (3)]) and

$$\|F_m\|_{L^\infty} \leq |F\|_{L^\infty}. \tag{A.22}$$

Therefore

$$\begin{aligned} & |E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x)) \\ & \qquad - E(F_m(\varphi^{-k_1}x)F_m(\varphi^{-k_2}x)F_m(\varphi^{-k_3}x)F_m(\varphi^{-k_4}x))| \\ \leq \quad & |E((F - F_m)(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x))| \\ & \qquad + |E(F(\varphi^{-k_1}x)(F - F_m)(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x))| \\ & \qquad + |E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)(F - F_m)(\varphi^{-k_3}x)F(\varphi^{-k_4}x))| \\ & \qquad + |E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)(F - F_m)(\varphi^{-k_4}x))| \\ \leq \quad & 4\|F\|_{L^\infty}^3 C_1 \rho^{(m^\kappa)}, \end{aligned} \tag{A.23}$$

where the last estimate follows from the Cauchy–Schwarz inequality.

We next estimate $|E(F_m(\varphi^{-k_1}x)F_m(\varphi^{-k_2}x)F_m(\varphi^{-k_3}x)F_m(\varphi^{-k_4}x))|$ under the hypotheses $k_1 \leq k_2 \leq k_3 \leq k_4$ and $k_2 - k_1 \geq 2^i$, which we may assume without loss of generality. Now

$$F_m = \sum_{A\in\mathcal{P}_{-m}^m} F_{m|A}\mathbf{1}_A$$

by definition of $\mathcal{P}_{-m}^m$, and since F_m depends only on $x_{-m}, \dots, x_m$, F_m is constant on the elements of $\mathcal{P}_{-m}^m$, so that for each $A \in \mathcal{P}_{-m}^m$, the function $F_{m|A}$ takes a single value in $\mathbb{R}$.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

Thus

$$\begin{aligned}|E(F_m(\varphi^{-k_1}x)F_m(\varphi^{-k_2}x)F_m(\varphi^{-k_3}x)F_m(\varphi^{-k_4}x))| \\ \leq \left|\sum_{A_1,A_2,A_3,A_4\in\mathcal{P}_{-m}^m} F_{m|A_1}F_{m|A_2}F_{m|A_3}F_{m|A_4}\mu(\varphi^{k_1}A_1)\mu(\varphi^{k_2}A_2\cap\varphi^{k_3}A_3\cap\varphi^{k_4}A_4)\right| \\ +\left|\sum_{A_1,A_2,A_3,A_4\in\mathcal{P}_{-m}^m} F_{m|A_1}F_{m|A_2}F_{m|A_3}F_{m|A_4}\cdot(\mu(\varphi^{k_1}A_1\cap\varphi^{k_2}A_2\cap\varphi^{k_3}A_3\cap\varphi^{k_4}A_4)\right. \\ \left.-\mu(\varphi^{k_1}A_1)\mu(\varphi^{k_2}A_2\cap\varphi^{k_3}A_3\cap\varphi^{k_4}A_4))\right|, \qquad (A.24)\end{aligned}$$

where we have used the preceding expression for F_m and the fact that $\mathbf{1}_A\mathbf{1}_B=\mathbf{1}_{A\cap B}$.

But the first term on the right-hand side of (A.24) vanishes, since it reduces to

$$\sum_{A_2,A_3,A_4\in\mathcal{P}_{-m}^m} F_{m|A_2}F_{m|A_3}F_{m|A_4}\mu(\varphi^{k_2}A_2\cap\varphi^{k_3}A_3\cap\varphi^{k_4}A_4)\cdot\sum_{A_1\in\mathcal{P}_{-m}^m} F_{m|A_1}\mu(\varphi^{k_1}A_1). \qquad (A.25)$$

Now $\mu(\varphi^{k_1}A_1)=\mu(A_1)$, so the second factor in the expression (A.25) above may be rewritten as

$$\sum_{A_1\in\mathcal{P}_{-m}^m} F_{m|A_1}\mu(\varphi^{k_1}A_1)=\int_X F_m\,d\mu(x)=\int_X F\,d\mu(x)=0.$$

Since $\varphi^{k_1}A_1\in\mathcal{P}_{-m+k_1}^{m+k_1}$ and $m+k_2-(m+k_1)\geq 2^i$, the second term on the right-hand side of (A.24) is bounded as follows:

$$\begin{aligned}&\|F\|_{L^\infty}^4\cdot(r^{2m+1})^3\cdot\sum_{\substack{A_1\in\mathcal{P}_{-m}^m\\ B\in\mathcal{P}_{m+k_2}^{+\infty}}}|\mu(\varphi^{k_1}A_1\cap B)-\mu(\varphi^{k_1}A_1)\mu(B)| \\ &\quad\leq\ \|F\|_{L^\infty}^4\cdot(r^{2m+1})^3\cdot C_0\gamma^{(2^{i\alpha})}\sum_{A_1\in\mathcal{P}_{-m}^m}\mu(\varphi^{k_1}A_1) \\ &\quad\leq\ |F\|_{L^\infty}^4\cdot(r^{2m+1})^3\cdot C_0\gamma^{(2^{i\alpha})}, \qquad (A.26)\end{aligned}$$

where $(r^{2m+1})^3$ is an upper bound on the number of elements of $\mathcal{P}_{m+k_2}^{+\infty}$ of the form $\varphi^{k_2}A_2\cap\varphi^{k_3}A_3\cap\varphi^{k_4}A_4$ (as is easily seen, since $\mathcal{P}_{-m}^m$ has r^{2m+1} elements).

Now assuming $\|F\|_{L^\infty}=C_3$ and concatenating the inequalities (A.23), (A.24) and (A.26), we find

$$\begin{aligned}\sup_{2^i\leq e(k,A)<2^i+1}&|E(F(\varphi^{-k_1}x)F(\varphi^{-k_2}x)F(\varphi^{-k_3}x)F(\varphi^{-k_4}x))| \\ &\leq 4C_1C_3^3\rho^{(m^\kappa)}+C_3^4r^{6m+3}C_0\gamma^{(2^{i\alpha})}. \qquad (A.27)\end{aligned}$$

Then (A.20) and (A.27) above give

$$E(|S_F^n|^4)\leq n^2C_3^4+n^2\sum_{i\geq 0}2^{4i+7}(4C_1C_3^3\rho^{(m_i^\kappa)}+C_3^4r^{6m_i+3}C_0\gamma^{(2^{i\alpha})}), \qquad (A.28)$$

where the m_i are chosen arbitrarily for each i. We then choose $m_i=[2^{i\alpha/2}]$, in which case $\alpha>0$, $0<\rho<1$ and $0<\gamma<1$ imply that the series

$$\sum_{i\geq 0}2^{4i+7}(4C_1C_2^3\rho^{[2^{i\alpha/2}]^\kappa}+C_2^4r^{6[2^{i\alpha/2}]+3}C_0\gamma^{(2^{i\alpha})})$$

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press

converges. Hence, (A.28) can be rewritten as

$$E(|S_F^n|^4) \leq n^2(C_3^4 + 4C_1C_3^3\Sigma_1 + C_0C_3^4\Sigma_2)$$

with

$$\Sigma_1 = \sum_{i\geq 0} 2^{4i+7}\rho^{[2^{i\alpha/2}]^\kappa} \quad \text{and} \quad \Sigma_2 = \sum_{i\geq 0} 2^{4i+7}r^{6[2^{i\alpha/2}]+3}\gamma^{(2^{i\alpha})}$$

which completes the proof of Lemma A.1. □

REFERENCES

[Ah] L. V. Ahlfors. *Complex Analysis*, 3rd edn. McGraw-Hill, New York, 1979.

[An] D. V. Anosov. Averaging in systems of ODEs with rapidly oscillating solutions. *Izv. Akad. Nauk. SSSR* **24** (1960), 721–742 (untranslated).

[Ar1] V. I. Arnold. Conditions for the applicability, and estimate of the error, of an averaging method for systems which pass through states of resonance in the course of their evolution. *Dokl. Akad. Nauk.* **161**(1) (1965), 331–334. English transl. *Soviet Math. Doklady* **6**(2) (1965), 581–585.

[Ar2] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, New York, 1980 (translated from Russian).

[Ar3] V. I. Arnold. *Geometrical Methods in the Theory of Ordinary Differential Equations*. Springer, New York, 1983 (translated from Russian).

[Ar4] V. I. Arnold (Ed.). *Dynamical Systems III. (Encyclopedia of Mathematical Sciences, Vol. III)*. Springer, Berlin, 1988 (translated from Russian).

[Bo] R. Bowen. Markov partitions for hyperbolic flows. *Amer. J. Math.* **95** (1973), 429–460.

[Du] H. S. Dumas. A new proof of Anosov's averaging theorem. *Hamiltonian Dynamical Systems (IMA Volumes in Mathematics and its Applications, Vol. 63)*. Eds. H. S. Dumas, K. R. Meyer and D. S. Schmidt. Springer, New York, 1995, pp. 127–136.

[DGL] H. S. Dumas, F. Golse and P. Lochak. Multiphase averaging for generalized flows on manifolds. *Ergod. Th. & Dynam. Sys.* **14** (1994), 53–67.

[FW] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, New York, 1984 (translated from Russian).

[Ki] Yu. Kifer. Limit theorems in averaging for dynamical systems. *Ergod. Th. & Dynam. Sys.* **15** (1995), 1143–1172.

[LM] P. Lochak and C. Meunier. *Multiphase Averaging for Classical Systems (Applied Mathematical Sciences Vol. 72)*. Springer, New York, 1988 (translated from French).

[Ne1] A. I. Neishtadt. Averaging in multi-frequency systems. *Dokl. Akad. Nauk. SSSR Mechanics* **223**(2) (1975), 314–317. English transl. *Soviet Phys. Dokl.* **20**(7) (1975), 492–494.

[Ne2] A. I. Neishtadt. Averaging in multi-frequency systems II. *Dokl. Akad. Nauk. SSSR Mechanics* **226**(6) (1976), 1295–1298. English transl. *Soviet Phys. Dokl.* **21**(2) (1976), 80–82.

[OW] D. Ornstein and B. Weiss. Geodesic flows are Bernoullian. *Israel J. Math.* **14** (1973), 184–198.

[Ra1] M. Ratner. Markov partitions for Anosov flows on n-dimensional manifolds. *Israel J. Math.* **15** (1973), 92–114.

[Ra2] M. Ratner. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. *Israel J. Math.* **16** (1973), 181–197.

[Ra3] M. Ratner. Anosov flows with Gibbs measures are also Bernoullian. *Israel J. Math.* **17** (1974), 380–391.

[Ra4] M. Ratner. The rate of mixing for geodesic and horocycle flows. *Ergod. Th. & Dynam. Sys.* **7** (1987), 267–288.

https://doi.org/10.1017/S0143385797097654 Published online by Cambridge University Press