

# Large Deviations, Averaging and Periodic Orbits of Dynamical Systems

**Yuri Kifer**<sup>1,2</sup>

Institute of Mathematics, Hebrew University, Jerusalem, Israel

Received: 13 January 1993/in revised form: 5 August 1993

**Abstract:** The paper deals with large deviation bounds for the proportion of periodic orbits with irregular behavior for expansive dynamical systems with specification, in particular, we obtain estimates for large deviations from the equidistribution for closed geodesics on negatively curved manifolds. We derive also large deviation bounds in the averaging principle when the fast motion is the shift along periodic orbits.

## 1. Introduction

Let  $f^t: M \rightarrow M$  be a group of homeomorphisms of a compact metric space with either discrete time  $t \in \mathbb{Z}$  or continuous time  $t \in \mathbb{R}$ . A point  $x \in M$  is called periodic if  $f^t x = x$  for some  $t > 0$  and the orbit  $\{f^s x, s \in \mathbb{Z} \text{ or } s \in \mathbb{R}\}$  of such  $x$  is called a closed (or periodic) orbit containing  $x$ . Denote by  $CO$  the set of all closed orbits and by  $CO_\delta(t)$  those orbits from  $CO$  with some period in the interval  $[t - \delta, t + \delta]$ . Let  $\gamma \in CO$ ,  $x \in \gamma$ , and  $\tau(\gamma)$  denotes the least period of  $\gamma$ . Then the map  $t \rightarrow f^t x$  sends the Lebesgue measure on  $[0, \tau(\gamma)]$  to the measure  $\tau(\gamma)\zeta_\gamma$  on  $\gamma$  where  $\zeta_\gamma = (\tau(\gamma))^{-1} \int_0^{\tau(\gamma)} \delta_{f^s x} ds$  in the continuous time case and  $\zeta_\gamma = (\tau(\gamma))^{-1} \sum_{i=1}^{\tau(\gamma)} \delta_{f^i x}$  in the discrete time case, with  $\delta_y$  standing for the unit mass at  $y$ . Set  $\mu_{t,\delta} = N_{t,\delta}^{-1} \sum_{\gamma \in CO_\delta(t)} \zeta_\gamma$ , where  $N_{t,\delta} = \# \{CO_\delta(t)\}$  is the number of elements in  $CO_\delta(t)$  which is finite if  $f^t$  is an expansive dynamical system (see, for instance, [BW]). By [B1, B2, and B3] (see also [F] and [Pa])  $\mu_{t,\delta}$  weakly converges as  $t \rightarrow \infty$  to the measure  $\mu_{\max}$  with maximal entropy for  $f^t$  provided  $f^t$  is a hyperbolic dynamical system and, in fact, the more general conditions of expansiveness and specification will suffice. This was called by R. Bowen the equidistribution of closed orbits. For  $\Gamma \subset CO$  set  $\nu_{t,\delta}(\Gamma) = N_{t,\delta}^{-1} \# \{\Gamma \cap CO_\delta(t)\}$ . Then by [La],

$$\lim_{t \rightarrow \infty} \nu_{t,\delta} \{ \gamma \in CO : \zeta_\gamma \notin U_{\mu_{\max}} \} = 0 \tag{1.1}$$

<sup>1</sup> Partially supported by US-Israel BSF

<sup>2</sup> Partially sponsored by the Edmund Landau Center for research in Mathematical Analysis, supported by the Minerva Foundation, (Germany)

for each open neighborhood  $U_{\mu_{\max}}$  of  $\mu_{\max}$  with respect to the topology of weak convergence on the space  $\mathcal{P}(M)$  of probability measures on  $M$ . This means that except for a small (tending to zero) proportion long periodic orbits are nearly uniformly distributed according to  $\mu_{\max}$ . The same results hold true if one considers all periodic orbits  $\gamma$  with  $\tau(\gamma) \leq t$ , i.e., when one replaces  $CO_\delta(t)$ ,  $N_{t,\delta}$ , and  $v_{t,\delta}$  by  $CO_t = CO_{t/2}(t/2)$ ,  $N_t = N_{t/2,t/2}$ , and  $v_t = v_{t/2,t/2}$ , respectively. In order to estimate this small proportion of “irregular” periodic orbits one needs the large deviations approach. I will prove here that for any  $\delta > 0$  small enough,

$$\begin{aligned} & \limsup_{t \rightarrow \infty} t^{-1} \log v_{t,\delta} \{ \gamma \in CO : \zeta_\gamma \in K \} \\ &= \limsup_{t \rightarrow \infty} t^{-1} \log v_t \{ \gamma \in CO : \zeta_\gamma \in K \} \leq -\inf \{ I(\mu) : \mu \in K \} \end{aligned} \quad (1.2)$$

for any closed  $K \subset \mathcal{P}(M)$  and

$$\begin{aligned} & \liminf_{t \rightarrow \infty} t^{-1} \log v_{t,\delta} \{ \gamma \in CO : \zeta_\gamma \in G \} \\ &= \liminf_{t \rightarrow \infty} t^{-1} \log v_t \{ \gamma \in CO : \zeta_\gamma \in G \} \geq -\inf \{ I(\mu) : \mu \in G \} \end{aligned} \quad (1.3)$$

for any open  $G \subset \mathcal{P}(M)$ , where

$$I(\mu) = \begin{cases} h(f^1) - h_\mu(f^1) & \text{if } \mu \in \mathcal{P}(M) \text{ is } f^t\text{-invariant} \\ \infty, & \text{otherwise.} \end{cases} \quad (1.4)$$

Here  $h(f^1) = \sup \{ h_\mu(f^1) : \mu \in \mathcal{P}(M), \mu \text{ is } f^t\text{-invariant} \}$  is the topological entropy of  $f^1$  and  $h_\mu(f^1)$  is the Kolmogorov–Sinai entropy of  $f^1$  with respect to  $\mu$ . In particular, this yields bounds of large deviations from the equidistribution for closed geodesics on negatively curved manifolds since the geodesic flow on such manifolds is a hyperbolic (even Anosov) dynamical system. Remark that (1.2) implies (1.1) but the main interest in (1.2) and (1.3) is in the precise bounds for sets disjoint with some neighborhood of  $\mu_{\max}$ . The above results require only the expansivity and specification properties for  $f^t$  together with the density of functions having unique equilibrium state in the space  $C(M)$  of all continuous functions on  $M$ . So the results remain true for homeomorphisms of the Smale spaces (see [Ru]) and for expanding maps of Riemannian manifolds. The latter can be reduced to the homeomorphism case by taking the natural extension as in [Ru]. Of course, the inequalities (1.2) and (1.3) are very close to the corresponding large deviation results for dynamical systems from [Ki1] and they follow in view of two possible ways of obtaining the topological pressure: by means of  $(\varepsilon, t)$ -separated sets and by means of periodic orbits. Another problem which I treat in this paper is connected with the averaging. Let  $M$  and  $N$  be compact Riemannian manifolds and  $B$  be a bounded and Lipschitz continuous map from  $N \times M$  to the tangent bundle  $TN$ . Let

$$\frac{dX^\varepsilon(t)}{dt} = B(X^\varepsilon(t), f^{t/\varepsilon} y), \quad X^\varepsilon(0) = x, \quad (1.5)$$

where  $X^\varepsilon(t) = X_{x,y}^\varepsilon(t) \in N$ ,  $\varepsilon > 0$  is a small parameter, and  $f^t$  is a smooth flow on  $M$  generated by another  $C^1$  vector field on  $M$  by means of the ordinary differential equation

$$\frac{df^t y}{dt} = b(f^t y), \quad f^0 y = y. \tag{1.6}$$

If  $y$  belongs to a periodic orbit  $\gamma$  of  $f^t$  then by the standard averaging principle (see, for instance, [SV]) the solution  $X_{x,y}^\varepsilon(t)$  of (1.5) is close for small  $\varepsilon$  and  $t \in [0, T]$  to the solution  $\bar{X}_\gamma(t) = \bar{X}_{x,\gamma}(t)$  of the equation

$$\frac{d\bar{X}_\gamma(t)}{dt} = \bar{B}_\gamma(\bar{X}_\gamma(t)), \quad \bar{X}_\gamma(0) = x, \tag{1.7}$$

where  $\bar{B}_\gamma(x) = (\tau(\gamma))^{-1} \int_0^{\tau(\gamma)} B(x, f^t y) dt$  and, recall,  $\tau(\gamma)$  is the least period of  $\gamma$ . Still, if one wants to describe  $X_{x,y}^\varepsilon(t)$  for small but fixed  $\varepsilon$  and all  $y$  lying on periodic orbits then the standard averaging principle does not help since  $X_{x,y}^\varepsilon(t)$  is  $\delta$ -close to  $\bar{X}_{x,\gamma}(t)$  only for  $\varepsilon \leq \varepsilon(\gamma, \delta)$ , where  $\varepsilon(\gamma, \delta)$  depends not only on  $\tau(\gamma)$  but on  $\gamma \in CO$  itself. Thus in order to obtain a global picture on the averaging along periodic orbits one has to employ the large deviation approach to averaging similarly to [Ki2]. In a certain sense, this will give a more detailed description of long periodic orbits than (1.2) and (1.3). Namely, let  $C_{0T}(N)$  be the space of continuous curves  $\varphi = \{\varphi_t \in N, t \in [0, T]\}$  on  $N$  with the metric

$$\rho_{0T}(\varphi, \psi) = \sup_{0 \leq t \leq T} d_N(\varphi_t, \psi_t), \quad \varphi, \psi \in C_{0T}(N), \tag{1.8}$$

where  $d_M$  and  $d_N$  denote distances on  $M$  and  $N$ , respectively. Let  $CO^*$ ,  $CO_\delta^*(t)$ , and  $CO_t^*$  denote the sets of points lying on orbits from  $CO$ ,  $CO_\delta(t)$ , and  $CO_t$ , respectively. Introduce the probability measures  $\nu_{t,\delta}^*$  and  $\nu_t^*$  on  $CO_\delta^*(t)$  and  $CO_t^*$  by

$$\nu_{t,\delta}^*(\Gamma) = \int \zeta_\gamma(\Gamma \cap \gamma) d\nu_{t,\delta}(\gamma) = N_{t,\delta}^{-1} \sum_{\gamma \in CO_\delta(t)} \zeta_\gamma(\Gamma \cap \gamma) \tag{1.9}$$

for  $\Gamma \subset CO^*$  and  $\nu_t^*(\Gamma)$  equal to the same expression with  $\nu_{t,\delta}$ ,  $N_{t,\delta}$ , and  $CO_\delta(t)$  replaced by  $\nu_t$ ,  $N_t$ , and  $CO_t$ , respectively.

I will show that for  $\varphi \in C_{0T}(N)$ ,  $\varphi_0 = x$ ,

$$\nu_{T/\varepsilon,\delta}^* \{y \in CO^* : \rho_{0T}(X_{x,y}^\varepsilon, \varphi) < \delta\} \asymp_{\varepsilon \rightarrow 0} \exp\left(-\frac{1}{\varepsilon} S_{0T}(\varphi)\right), \tag{1.10}$$

where  $\asymp$  means a logarithmic asymptotic in the sense specified in Sect. 3 and  $S_{0T}(\varphi) \geq 0$  is a functional which differs from  $\infty$  if and only if the curve  $\varphi$  is a combination of averaged motions  $\bar{X}_\mu(t)$  satisfying

$$\frac{d\bar{X}_\mu(t)}{dt} = \bar{B}_\mu(\bar{X}_\mu(t)), \quad \bar{B}_\mu(x) = \int_M B(x, y) d\mu(y) \tag{1.11}$$

with  $\mu$  being an invariant measure of the flow  $f^t$ . Furthermore,  $S_{0T}(\varphi) = 0$  if and only if  $\varphi_t = \bar{X}_{\mu_{\max}}(t)$  for all  $t \in [0, T]$ . In this sense, the most “probable” behavior of  $X_{x,y}^\varepsilon$  has to be close to the motion averaged with respect to the measure with maximal entropy. The relation (1.10) remains true if  $\nu_{t,\delta}^*$  is replaced by  $\nu_{T/\varepsilon,\delta}^*$ . Maybe for the first time in connection with the averaging problems I apply here the

contraction principle from large deviations which, in particular, simplifies some of the arguments from [Ki2].

Assuming certain properties of  $\bar{B}_{\mu_{\max}}$  one can obtain further estimates corresponding to unbounded time intervals. In the last section I consider averaging for parabolic second order partial differential equations and stochastic differential equations employing the contraction principle.

## 2. Large Deviations

In the discrete time case a dynamical system  $f^t$  on  $M$  is called expansive if there exists  $\delta_0 > 0$  such that for any pair of points  $x \neq y$  in  $M$  there is  $n \in \mathbb{Z}$  so that  $d_M(f^n x, f^n y) > \delta_0$ . In the continuous time case  $f^t, t \in \mathbb{R}$  is called expansive if for every  $\eta > 0$  there exists  $\delta_0(\eta) > 0$  so that  $d_M(f^t x, f^{s(t)} y) < \delta_0(\eta)$  for all  $t \in \mathbb{R}$ , for a pair of points  $x, y \in M$ , and for a continuous function  $s: \mathbb{R} \rightarrow \mathbb{R}$  with  $s(0) = 0$  implies  $y = f^u x$  for some  $u \in [-\eta, \eta]$ . Next,  $f^t, t \in \mathbb{Z}$  satisfies specification if for each  $\eta$  there is an integer  $L(\eta)$  so that if  $t_0, \dots, t_{n+1} \in \mathbb{Z}$  satisfy  $t_{i+1} \geq t_i + L(\eta)$ , then for any  $x_0, \dots, x_n \in M$  there exists  $x \in M$  with  $f^{t_{n+1}-t_0+L(\eta)} x = x$  and  $d_M(f^l x, f^{l-t_i} x_i) < \eta$  for all  $l \in [t_i, t_{i+1} - L(\eta)]$ ,  $i = 0, 1, \dots, n$ . In the continuous time case  $f^t, t \in \mathbb{R}$  is said to satisfy specification if for each  $\eta$  there is  $L(\eta)$  so that if  $t_0, \dots, t_{n+1} \in \mathbb{R}$  satisfy  $t_{i+1} \geq t_i + L(\eta)$ , then for any  $x_0, \dots, x_n \in M$  there exists  $x \in CO_{\eta}^*(t_{n+1} - t_0)$  and a step function  $s(t)$  constant on each  $[t_i, t_{i+1}]$  with  $s[t_0, t_1] = 0$  and  $|s[t_{i+1}, t_{i+2}] - s[t_i, t_{i+1}]| < \eta$  such that  $d_M(f^{t+s(t)} x, f^{t-t_i} x_i) < \eta$  for all  $t \in [t_i, t_{i+1} - L(\eta)]$ .

For  $g \in C(M)$  set  $S_t g(x) = \sum_{i=0}^{t-1} g(f^i x)$  for the discrete time case, and  $S_t g(x) = \int_0^t g(f^s x) ds$  for the continuous time case. Let  $V(f)$  be the set of  $g \in C(M)$  for which there exist  $\delta, K > 0$  such that  $\sup_{0 \leq s < t} d_M(f^s x, f^s y) \leq \delta$  implies  $|S_t g(x) - S_t g(y)| \leq K$ .

**2.1. Theorem.** *Suppose that  $f^t$  is an expansive dynamical system on  $M$  satisfying specification and such that  $V(f)$  is separable and dense in  $C(M)$ . Then (1.2)–(1.4) hold true for all  $\delta > 0$  small enough.*

*Proof.* For  $\gamma \in CO_\delta(t)$  let  $\tau_{t,\delta}(\gamma)$  denotes a period of  $\gamma$  in the interval  $[t - \delta, t + \delta]$ . Then

$$|\tau_{t,\delta}(\gamma)^{-1}(t - \tau_{t,\delta}(\gamma)) S_{\tau_{t,\delta}(\gamma)} g(x)| \leq \delta \|g\|, \quad (2.1)$$

where  $\gamma \in CO_\delta(t)$  and  $\|g\| = \sup\{|g(x)|: x \in M\}$ . In the same way as in Lemma 4 of [B3] and Lemma 2.8 of [F] it follows that for any  $g \in C(M)$  and  $\delta$  small enough,

$$\lim_{t \rightarrow \infty} t^{-1} \log \sum_{\gamma \in CO_\delta(t)} \exp(S_{\tau_{t,\delta}(\gamma)} g(x)) = P(g), \quad (2.2)$$

where  $P(g)$  is the topological pressure of  $g$  which satisfies the variational principle (see [Wa]),

$$P(g) = \sup \left\{ \int g d\mu + h_\mu(f^1): \mu \in \mathcal{P}(M) \text{ is } f^1\text{-invariant} \right\}. \quad (2.3)$$

In particular, taking  $g \equiv 0$ ,

$$\lim_{t \rightarrow \infty} t^{-1} \log N_{t,\delta} = h(f^1). \quad (2.4)$$

Thus by (2.1)–(2.4),

$$\begin{aligned}
& \lim_{t \rightarrow \infty} t^{-1} \log \int_{CO} \exp(t \int g d\zeta_\gamma) dv_{t,\delta}(\gamma) \\
&= \lim_{t \rightarrow \infty} t^{-1} \log \left( N_{t,\delta}^{-1} \sum_{\gamma \in CO_\delta(t)} \exp(t \int g d\zeta_\gamma) \right) \\
&= -h(f^1) + \lim_{t \rightarrow \infty} t^{-1} \log \sum_{\gamma \in CO_\delta(t)} \exp[S_{\tau_{t,\delta}(\gamma)} g(x)] \exp[\tau_{t,\delta}(\gamma)^{-1}(t - \tau_{t,\delta}(\gamma)) S_{\tau_{t,\delta}(\gamma)} g(x)] \\
&= P(g) - h(f^1) = \sup_{\mu \in \mathcal{P}(M)} (\int g d\mu - I(\mu)) \tag{2.5}
\end{aligned}$$

with  $I(\mu)$  defined by (1.4) which is a convex lower semi-continuous functional since  $f^t$  is expansive, and so the entropy  $h_\mu(f^1)$  is upper semicontinuous, and affine in  $\mu$  (see, for instance, [DGS]). By [B4 and F] expansivity and specification yield that for any  $g \in V(f)$  there exists a unique measure  $\mu_g \in \mathcal{P}(M)$ , called the equilibrium state for  $g$ , such that

$$P(g) = \int g d\mu_g - I(\mu_g) + h(f^1). \tag{2.6}$$

Since  $V(f)$  is dense in  $C(M)$  then by Theorem 2.1 from [Kil] these imply the large deviation bounds (1.2) and (1.3) for the measures  $v_{t,\delta}$ . In order to obtain (1.2) and (1.3) for  $v_t$  I remark, first, that

$$\lim_{t \rightarrow \infty} t^{-1} \log N_t = h(f^1). \tag{2.7}$$

Indeed,

$$N_{t-\delta,\delta} \leq N_t \leq N_{t,\delta} + N_{t-2\delta,\delta} + \dots + N_{t-2k\delta,\delta} + N_{0,\delta}, \tag{2.8}$$

where  $k$  is the integral part of  $t/2\delta$ . Thus (2.7) follows from (2.4) and (2.8). Next, similarly to (2.8) for any  $\Gamma \subset CO$ ,

$$\begin{aligned}
N_{t-\delta,\delta} v_{t-\delta,\delta}(\Gamma) &\leq N_t v_t(\Gamma) \leq N_{t,\delta} v_{t,\delta}(\Gamma) + N_{t-2\delta,\delta}(\Gamma) \\
&\quad + N_{t-2k\delta,\delta} v_{t-2k\delta,\delta}(\Gamma) + N_{0,\delta} v_{0,\delta}(\Gamma). \tag{2.9}
\end{aligned}$$

Now (1.2) and (1.3) for  $v_{t,\delta}$  together with (2.4), (2.7), and (2.9) yield (1.2) and (1.3) for  $v_t$  completing the proof.  $\square$

**2.2. Remark.** The same large deviation bounds hold true for the measures  $\omega_{t,\delta}$  and  $\omega_t$  defined for  $\Gamma \subset CO$  by  $\omega_{t,\delta}(\Gamma) = (\sum_{\gamma \in CO_\delta(t)} \tau(\gamma))^{-1} (\sum_{\gamma \in \Gamma \cap CO_\delta(t)} \tau(\gamma))$  and  $\omega_t = \omega_{t/2,t/2}$ . Indeed, for  $\Gamma \subset CO$ ,

$$(t + \delta)^{-1} N_{t,\delta}^{-1} (N_{t,\delta} v_{t,\delta}(\Gamma) - N_1) \leq \omega_{t,\delta}(\Gamma) \leq (t + \delta) N_{t,\delta}(\Gamma) v_{t,\delta}(\Gamma) (N_{t,\delta} - N_1)^{-1} \tag{2.10}$$

and

$$t^{-1} N_t^{-1} (N_t v_t(\Gamma) - N_1) \leq \omega_t(\Gamma) \leq t N_t v_t(\Gamma) (N_t - N_1)^{-1}, \tag{2.11}$$

and so (1.2) and (1.3) for  $\omega_{t,\delta}$  and  $\omega_t$  follow from (1.2) and (1.3) for  $v_{t,\delta}$  and  $v_t$ .

**2.3. Remark.** By [B1 and B4] the conditions of Theorem 2.1 are satisfied if  $f^t$  is a smooth dynamical system restricted to a basic hyperbolic set, in particular, in case of an Anosov dynamical system. Then  $V(f)$  contains all Hölder continuous functions, and so it is dense in  $C(M)$ . Since the geodesic flow on a compact

manifold of negative curvature is an Anosov flow and closed geodesics are in one-to-one correspondence to periodic orbits of the corresponding geodesic flow, then (1.2) and (1.3) give large deviation bounds for closed geodesics on negatively curved manifolds where  $CO$ ,  $CO_\delta(t)$ , and  $CO_t$  should stand for sets of closed geodesics, for those with length in the interval  $[t - \delta, t + \delta]$ , and for those with length not exceeding  $t$ , respectively.

*2.4. Remark.* Theorem 2.1 remains valid for subshifts of finite type after making it aperiodic by the usual trick (see [Si] and Example 3 in [B3]). By [We] and [Si] irreducible sofic systems satisfy specification and since the functions constant on cylinder sets are both in  $V(f)$  and dense in  $C(M)$  then these dynamical systems also satisfy the conditions of Theorem 2.1 (see also Sect. 6 in [KT]). I thank B. Weiss for clarifying this point for me. In view of the equivalence of subshifts of finite type and finite oriented graphs we obtain as a corollary large deviations results for numbers of closed paths in such graphs with prescribed average properties.

*2.5. Remark.* By the usual contraction type argument (see Corollary 2.1 in [Ki1]) (1.2) and (1.3) yield large deviation bounds for integrals of functions. Namely, let  $g_1, \dots, g_n \in C(M)$  and define  $\varphi_n: \mathcal{P}(M) \rightarrow \mathbb{R}^n$  by  $\varphi_n \mu = (\int g_1 d\mu, \int g_2 d\mu, \dots, \int g_n d\mu)$ . Set  $J_n(\alpha) = \inf \{I(\mu): \varphi_n \mu = \alpha\}$  if  $\alpha \in \varphi_n \mathcal{P}(M)$  and  $J_n(\alpha) = \infty$  if  $\alpha \notin \varphi_n \mathcal{P}(M)$ . Then for any closed  $K \subset \mathbb{R}^n$ ,

$$\begin{aligned} & \limsup_{t \rightarrow \infty} t^{-1} \log v_{t, \delta} \{ \gamma \in CO: \varphi_n \zeta_\gamma \in K \} \\ &= \limsup_{t \rightarrow \infty} t^{-1} \log v_t \{ \gamma \in CO: \varphi_n \zeta_\gamma \in K \} \leq - \inf_{\alpha \in K} J_n(\alpha), \end{aligned} \quad (2.12)$$

and for any open  $G \subset \mathbb{R}^n$ ,

$$\begin{aligned} & \liminf_{t \rightarrow \infty} t^{-1} \log v_{t, \delta} \{ \gamma \in CO: \varphi_n \zeta_\gamma \in G \} \\ &= \liminf_{t \rightarrow \infty} t^{-1} \log v_t \{ \gamma \in CO: \varphi_n \zeta_\gamma \in G \} \geq - \inf_{\alpha \in G} J_n(\alpha). \end{aligned}$$

Again by the contraction principle it is easy to obtain also large deviation bounds for measures uniformly distributed on periodic orbits with additional restrictions  $|\varphi_n \zeta_\gamma - \alpha| \leq \delta t^{-1}$  as suggested in [La]. Then one ends up with functionals  $I(\mu)$  with  $\varphi_n \mu = \alpha$ .

### 3. Averaging: Bounded Time

I start with the following result similar to Theorem 2.1 from [Ki2].

**3.1. Theorem.** *Let  $f^t$  be an expansive dynamical system on  $M$  satisfying specification. Then for any continuous function  $q(t, y) = q_t(y)$  on  $[0, T] \times M$ ,  $T > 0$  and  $\delta > 0$  small enough the limit*

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \log \int_{CO^*} \exp \left( \varepsilon^{-1} \int_0^T q_t(f^{t/\varepsilon} y) dt \right) dv_{T/\varepsilon, \delta}^*(y) = \int_0^T P(q_t) dt - Th(f^1) \quad (3.1)$$

exists. The limit remains the same if  $v_{T/\varepsilon, \delta}^*$  is replaced by  $v_{T/\varepsilon}^*$ ,  $\omega_{T/\varepsilon, \delta}^* = \zeta_\gamma d\omega_{T/\varepsilon, \delta}(\gamma)$ , or  $\omega_{T/\varepsilon}^* = \zeta_\gamma d\omega_{T/\varepsilon}(\gamma)$ .

*Proof.* By the uniform continuity of  $q(t, y)$  one concludes in the same way as in Sect. 3 from [Ki2] that the assertion of Theorem 3.1 is equivalent to the statement that for any continuous functions  $q_1, \dots, q_k \in C(M)$  and for any numbers  $0 = t_0 < t_1 < \dots < t_{k-1} < t_k = T$  one has

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \varepsilon \log \int_{CO^*} \exp \left( \varepsilon^{-1} \sum_{i=1}^k \int_{t_{i-1}}^{t_i} q_i(f^{s/\varepsilon} y) ds \right) d\nu_{T/\varepsilon, \delta}^*(y) \\ & = \sum_{i=0}^k (t_i - t_{i-1}) P(q_i) - Th(f^1). \end{aligned} \tag{3.2}$$

Set  $a_i = t_i T^{-1}$ ,  $i = 1, \dots, k$  and

$$D(t, y) = \exp \left( \sum_{i=1}^k \int_{a_{i-1}t}^{a_i t} q_i(f^s y) ds \right). \tag{3.3}$$

Taking  $t = T\varepsilon^{-1}$  we see that (3.2) is equivalent to

$$\lim_{t \rightarrow \infty} t^{-1} \log \int_{CO^*} D(t, y) d\nu_{t, \delta}^*(y) = \sum_{i=0}^k (a_i - a_{i-1}) P(q_i) - h(f^1), \tag{3.4}$$

which, in turn, is equivalent to

$$\lim_{t \rightarrow \infty} t^{-1} \log \sum_{\gamma \in CO_\delta(t)} \int_0^{\tau(\gamma)} D(t, f^s y_\gamma) ds = \sum_{i=0}^k (a_i - a_{i-1}) P(q_i), \tag{3.5}$$

where  $y_\gamma$  is any point on  $\gamma \in CO_\delta(t)$ . For fixed  $\kappa > 0$  set  $y_\gamma^j = f^{j\kappa} y_\gamma$ , for  $j = 0, 1, \dots, n_\gamma$  with  $n_\gamma = \text{integral part of } \kappa^{-1} \tau(\gamma)$ , where  $y_\gamma \in CO$ . Then (3.5) is equivalent to

$$\lim_{t \rightarrow \infty} t^{-1} \log \sum_{\gamma \in CO_\delta(t)} \sum_{j=0}^{n_\gamma} D(t, y_\gamma^j) = \sum_{i=0}^k (a_i - a_{i-1}) P(q_i). \tag{3.6}$$

Recall that a set  $E \subset M$  is called  $(\beta, t)$ -separated if  $z_1, z_2 \in E$ ,  $z_1 \neq z_2$  implies  $\sup_{0 \leq s \leq t} d_M(f^s z_1, f^s z_2) > \beta$ . It follows from Lemmas 1–3 in [B3] and Lemmas 2.1–2.6 in [F] that if  $\beta > 0$  is small enough then for any family  $E_{\beta, t}$  of maximal  $(\beta, t)$ -separated sets

$$\lim_{t \rightarrow \infty} t^{-1} \log \sum_{z \in E_{\beta, t}} D(t, z) = \sum_{i=0}^k (a_i - a_{i-1}) P(q_i). \tag{3.7}$$

The equality of the left-hand sides of (3.6) and (3.7) for small  $\delta$  and  $\beta$  follows in the same way as in the proof of Lemma 4 in [B3] and Lemma 2.8 in [F], completing the proof of (3.1). In the same way as in Sect. 2 one derives that the limit remains the same if  $\nu_{T/\varepsilon, \delta}^*$  is replaced by  $\nu_{T/\varepsilon}^*$ , by  $\omega_{T/\varepsilon, \delta}^* = \zeta_\gamma d\omega_{T/\varepsilon, \delta}(\gamma)$ , or by  $\omega_{T/\varepsilon}^* = \zeta_\gamma d\omega_{T/\varepsilon}(\gamma)$ .  $\square$

Next, I will derive large deviation estimates for solutions  $X_{x, y}^\varepsilon(t)$  of (1.5) from Theorem 2.1 in [Ki1] and Theorem 3.1 above by the ‘‘contraction principle’’ argument. Set  $\pi_y^\varepsilon = T^{-1} \int_0^T \delta_{(s, f^{s/\varepsilon} y)} ds$ ,  $y \in M$  which belongs to the space

$\mathcal{P}([0, T] \times M)$  of probability measures on  $[0, T] \times M$ . If  $q(t, y) = q_t(y) \in C([0, T] \times M)$ , then I can rewrite (3.1) as

$$\begin{aligned} & \lim_{\varepsilon \rightarrow 0} \varepsilon \log \int_{CO^*} \exp(\varepsilon^{-1} \int_0^T q d\pi_y^\varepsilon) dv_{T/\varepsilon, \delta}^*(y) \\ &= \int_0^T (P(T^{-1} q_t) - h(f^1)) dt \\ &= \sup_{\eta \in \mathcal{P}([0, T] \times M)} (\int q d\eta - I_{0T}(\eta)), \end{aligned} \quad (3.8)$$

where

$$I_{0T}(\eta) = \begin{cases} \int_0^T I(\eta_t) dt & \text{if } d\eta = \frac{1}{T} d\eta_t dt \text{ and} \\ & \eta_t \in \mathcal{P}(M) \text{ is } f^s\text{-invariant;} \\ \infty, & \text{otherwise} \end{cases} \quad (3.9)$$

and  $I(\eta_t)$  is defined by (1.4).

Assume that  $V(f)$  is dense in  $C(M)$ , then since  $V(f)$  is closed under linear operations it is easy to see that the set

$$V_{0T}(f) = \{q(t, y) = q_t(y) \in C([0, T] \times M): q_t \in V(f) \text{ for all } t \in [0, T]\}$$

is dense in  $C([0, T] \times M)$ . In view of [B3 and F] for each  $q \in V_{0T}(f)$  there exists a unique  $\eta_q \in \mathcal{P}([0, T] \times M)$  on which the supremum in the right-hand side of (3.8) is attained and, in fact,  $d\eta_q = d\mu_{T^{-1}q_t} dt$ , where  $\mu_g$  for  $g \in V(f)$  is the unique measure satisfying (2.6). Thus Theorem 2.1 from [Kil] together with Theorem 3.1 above yield

$$\begin{aligned} & \limsup_{\varepsilon \rightarrow 0} \varepsilon \log v_{T/\varepsilon, \delta}^* \{y: \pi_y^\varepsilon \in K\} \leq -\inf \{I_{0T}(\mu): \mu \in K\} \\ & \text{for any closed } K \subset \mathcal{P}([0, T] \times M) \text{ and} \end{aligned} \quad (3.10)$$

$$\begin{aligned} & \liminf_{\varepsilon \rightarrow 0} \varepsilon \log v_{T/\varepsilon, \delta}^* \{y: \pi_y^\varepsilon \in G\} \geq -\inf \{I_{0T}(\mu): \mu \in G\} \\ & \text{for any open } G \subset \mathcal{P}([0, T] \times M). \end{aligned} \quad (3.11)$$

On the subspace  $\mathcal{M} \subset \mathcal{P}([0, T] \times M)$  of measures  $\mu$  such that  $d\mu = \frac{1}{T} d\mu_t dt$ ,  $\mu_t \in \mathcal{P}(M)$ ,  $t \in [0, T]$  define the map  $\Psi_x: \mathcal{M} \rightarrow C_{0T}(N)$ ,  $x \in N$  by  $\varphi = \Psi_x \mu$  with

$$\varphi_t = x + \int_0^t \int_M B(\varphi_s, y) d\mu_s(y) ds, \quad (3.12)$$

where  $B$  is the same as in (1.5), which is well defined since  $B$  is Lipschitz continuous. Then  $\Psi$  is a continuous map if one takes the topology of weak convergence on  $\mathcal{P}([0, T] \times M)$  and the metric (1.8) on  $C_{0T}(N)$ . Clearly,  $X_{x,y}^\varepsilon = \Psi_x \pi_y^{\varepsilon, T}$ . Then (3.10) and (3.11) imply that for  $S_{0T}(\varphi) = \inf_\mu \{I_{0T}(\mu): \Psi_x \mu = \varphi\}$ ,

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon \log v_{T/\varepsilon, \delta}^* \{y: X_{x,y}^\varepsilon \in \tilde{K}\} \leq -\inf \{S_{0T}(\varphi): \mu \in \tilde{K}\} \quad (3.13)$$

for any closed  $\tilde{K} \subset C_{0T}(N)$  and

$$\liminf_{\varepsilon \rightarrow 0} \varepsilon \log v_{T/\varepsilon, \delta}^* \{y: X_{x,y}^\varepsilon \in \tilde{G}\} \geq -\inf \{S_{0T}(\varphi): \varphi \in \tilde{G}\} \quad (3.14)$$

for any open  $\tilde{G} \subset C_{0T}(N)$ , where  $S_{0T}(\varphi)$  is defined to be  $\infty$  if  $\varphi \notin \Psi_x \mathcal{M}$ . The transition from (3.10) and (3.11) to (3.13) and (3.14) is one of the versions of the, so-called, ‘‘contraction principle’’ and I appreciate a useful remark of R. Liptser concerning it. In view of (3.9) and (3.12),

$$S_{0T}(\varphi) = \int_0^T \inf \{I(\mu): \dot{\varphi}_t = \bar{B}_\mu(\varphi_t), \mu \in \mathcal{P}(M) \text{ is } f^s\text{-invariant}\} dt \quad (3.15)$$

with  $\bar{B}_\mu$  defined in (1.11), provided for Lebesgue almost all  $t \in [0, T]$  there exists an  $f^s$ -invariant  $\mu_t$  for which  $\dot{\varphi}_t = \bar{B}_{\mu_t}(\varphi)$ , and  $S_{0T}(\varphi) = \infty$ , otherwise. It follows from Sect. 9.1 of [IT] that  $S_{0T}(\varphi)$  is a lower semicontinuous functional on  $C_{0T}(N)$ ,  $S_{0T} \geq 0$ , and  $S_{0T}(\varphi) = 0$  if and only if  $\dot{\varphi}_t = \bar{B}_{\mu_{\max}}(\varphi_t)$  for all  $t \in [0, T]$ , where  $\mu_{\max}$  is the unique  $f^t$ -invariant probability measure such that  $\mu_{\mu_{\max}}(f^1) = h(f^1)$ .

Set  $\Phi_{0T}^a(x) = \{\varphi \in C_{0T}(N): \varphi_0 = x, S_{0T}(\varphi) \leq a\}$ . Since  $S_{0T}(\varphi)$  is lower semicontinuous then  $\Phi_{0T}^a(x)$  is a closed set. Moreover, it is easy to see that  $\Phi_{0T}^a(x)$  is compact for any  $a < \infty$ . Then one can rewrite (3.13) and (3.14) in the following form (cf. Sect. 3.3 in [FW]).

**3.2. Theorem.** *Suppose that the conditions of Theorem 2.1 are satisfied and  $B$  in (1.5) is Lipschitz continuous. Then for any  $\alpha, \beta, \lambda > 0$ , each  $\delta > 0$  small enough, and every  $\varphi \in C_{0T}(N)$ ,  $\varphi_0 = x$  there exists  $\varepsilon_0 > 0$  such that for all  $\varepsilon \in (0, \varepsilon_0)$ ,*

$$v_{T/\varepsilon, \delta}^* \{y \in CO^*: \rho_{0T}(X_{x,y}^\varepsilon, \varphi) < \beta\} \geq \exp \left\{ -\frac{1}{\varepsilon} S_{0T}(\varphi) + \lambda \right\} \quad (3.16)$$

and

$$v_{T/\varepsilon, \delta}^* \{y \in CO^*: \rho_{0T}(X_{x,y}^\varepsilon, \Phi_{0T}^a(x)) \geq \beta\} \leq \exp \left\{ -\frac{1}{\varepsilon} (a - \lambda) \right\} \quad (3.17)$$

with  $\rho_{0T}$  defined by (1.8). The functional  $S_{0T}(\varphi)$  is finite if and only if  $\dot{\varphi}_t = \bar{B}_{\nu_t}(\varphi_t)$  for some  $f^s$ -invariant  $\nu_t \in \mathcal{P}(M)$  and Lebesgue almost all  $t \in [0, T]$  and  $S_{0T}(\varphi) = 0$  if and only if  $\dot{\varphi}_t = \bar{B}_{\mu_{\max}}(\varphi_t)$  for all  $t \in [0, T]$ . The estimates (3.16) and (3.17) remain true if  $v_{T/\varepsilon, \delta}^*$  is replaced by  $v_{T/\varepsilon}^*$ ,  $\omega_{T/\varepsilon, \delta}^*$ , or  $\omega_{T/\varepsilon}^*$ .

This theorem says that ‘‘most likely’’  $X_{x,y}^\varepsilon(t)$  will stay close to the averaged motion  $\dot{\varphi}_t = \bar{B}_{\mu_{\max}}(\varphi_t)$ ,  $\varphi_0 = x$ . Still, for a set of  $y$ 's with an exponentially small in  $\varepsilon^{-1}$  measure  $X_{x,y}^\varepsilon(t)$  may stay near any curve  $\varphi \in C_{0T}(N)$  satisfying  $\dot{\varphi}_t = \bar{B}_{\nu_t}(\varphi_t)$  with  $f^s$ -invariant  $\nu_t \in \mathcal{P}(M)$ . Other motions for  $X_{x,y}^\varepsilon(t)$  may occur only for  $y$ 's from sets whose measure decrease as  $\varepsilon \rightarrow 0$  faster than exponentially in  $\varepsilon^{-1}$ .

**3.3. Remark.** One can derive Theorem 3.2 via Theorem 2.1 from [Fr] in the same way as in [Ki2] but [Fr] deals with limits of the type (3.1) for integrals with respect to a measure independent of  $\varepsilon$  though here  $v_{T/\varepsilon, \delta}^*$  depends on  $\varepsilon$ , and so in order to use [Fr] one has to generalize a bit Theorem 2.1 from there which is also possible.

**3.4. Remark.** The dynamical systems from Remark 2.3 satisfy the conditions of Theorem 3.2 since these conditions are the same as in Theorem 2.1.

3.5. *Remark.* Define the map  $\tilde{\Psi}_x : \mathcal{P}(M) \rightarrow C_{0T}(N)$ ,  $x \in N$  by  $\varphi = \tilde{\Psi}_x \mu$  with  $\varphi_t = x + \int_0^t \int_M B(\varphi_s, y) d\mu(y) ds$ , then  $\tilde{\Psi}_x$  is a continuous map. For  $\varphi \in \tilde{\Psi}_x \mathcal{P}(M)$  set

$$S(\varphi) = \inf \{ I(\mu) : \mu \in \mathcal{P}(M) \text{ and } \tilde{\Psi}_x \mu = \varphi \} \quad (3.18)$$

and  $S(\varphi) = \infty$  if  $\varphi \notin \tilde{\Psi}_x \mathcal{P}(M)$ . Since  $\bar{X}_{x,y} = \tilde{\Psi}_x \zeta_y$  with  $\bar{X}_{x,y}$  defined by (1.7), then Theorem 2.1 together with the ‘‘contraction principle’’ type argument yield that for any  $T, a, \beta, \lambda > 0$ , each  $\delta > 0$  small enough, and every  $\varphi \in C_{0T}(N)$ ,  $\varphi_0 = x$  there exists  $t_0 > 0$  such that for all  $t \geq t_0$ ,

$$v_{t,\delta} \{ \gamma \in CO : \rho_{0T}(\bar{X}_{x,\gamma}, \varphi) < \beta \} \geq \exp \left\{ -\frac{1}{\varepsilon} (S(\varphi) + \lambda) \right\} \quad (3.19)$$

and

$$v_{t,\delta} \{ \gamma \in CO : \rho_{0T}(\bar{X}_{x,\gamma}, \tilde{\Phi}_{0T}^a(x)) \geq \beta \} \leq \exp \left\{ -\frac{1}{\varepsilon} (a - \lambda) \right\}, \quad (3.20)$$

where  $\tilde{\Phi}_{0T}^a(x) = \{ \varphi \in C_{0T}(N) : \varphi_0 = x, S(\varphi) \leq a \}$ . The same inequalities will hold true if  $v_{t,\delta}$  is replaced by  $v_t, \omega_{t,\delta}$ , or  $\omega_t$ . One can rewrite (3.18) in the form

$$S(\varphi) = \inf \{ I(\mu) : \mu \in \mathcal{P}(M) \text{ is } f^s\text{-invariant and } \varphi_t = \bar{B}_\mu(\varphi_t) \text{ for all } t \in [0, T] \} \quad (3.21)$$

with  $S(\varphi) = \infty$  if it is not defined by (3.21). Thus  $S(\varphi) < \infty$  if and only if  $\varphi = \bar{X}_\mu$  for some  $f^s$ -invariant  $\mu$  and  $S(\varphi) = 0$  if and only if  $\varphi = \bar{X}_{\mu_{\max}}$ .

3.6. *Remark.* For a domain  $V \subset N$  set

$$\tau_{x,y}^\varepsilon(V) = \inf \{ t \geq 0 : X_{x,y}^\varepsilon(t) \notin V \} \quad (3.22)$$

and  $\tau_{x,y}^\varepsilon(V) = \infty$  if  $X_{x,y}^\varepsilon(t) \in V$  for all  $t \geq 0$ . Then it follows easily from Theorem 3.2 (cf. [Fr]) that

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \varepsilon \log v_{T/\varepsilon, \delta}^* \{ y \in CO^* : \tau_{x,y}^\varepsilon(V) < T \} \\ = \inf \{ S_{0t}(\varphi) : 0 \leq t \leq T, \varphi \in C_{0T}(N), \varphi_0 = x, \varphi_t \notin V \}, \end{aligned} \quad (3.23)$$

provided  $\delta > 0$  is small enough and  $V$  is a connected open domain with a compact closure  $\bar{V}$ . The limit remains the same if  $v_{T/\varepsilon, \delta}^*$  is replaced by  $v_{T/\varepsilon}^*, \omega_{T/\varepsilon, \delta}^*$ , and  $\omega_{T/\varepsilon}^*$ .

3.7 *Remark.* One can consider other types of large deviations in this situation. Consider, for instance, the equation

$$\frac{d\tilde{X}_{x,y}^\varepsilon(t)}{dt} = B^\varepsilon(\tilde{X}_{x,y}^\varepsilon(t), f^t y) \text{ with } B^\varepsilon(x, y) = \int_0^1 B(x, f^{s/\varepsilon} y) ds. \quad (3.24)$$

Let  $\hat{\Psi}_x$  be the map from  $\mathcal{P}(M)$  to  $C_{0T}(\varphi)$  defined by  $\hat{\Psi}_x \mu = \varphi$  with

$$\varphi_t = x + \int_0^t ds \int_M B(\varphi_s, f^s y) d\mu(y). \quad (3.25)$$

Then  $\tilde{X}_{x,y}^\varepsilon = \hat{\Psi}_x \zeta_y^\varepsilon$ , where  $\zeta_y^\varepsilon = \varepsilon \int_0^{1/\varepsilon} \delta_{f^s y} ds$ . Note that if  $\mu$  is  $f^t$ -invariant then (3.25) becomes  $\varphi_t = x + \int_0^t ds \int_M B(\varphi_s, y) d\mu(y)$ . Thus if one has the large deviation estimates for  $\zeta_y^\varepsilon$  with the rate function  $I(\mu)$ , then for  $\tilde{X}_{x,y}^\varepsilon$  one has estimates of the form (3.16) and (3.17) with  $S_{0T}(\varphi) = \inf \{ I(\mu) : \hat{\Psi}_x \mu = \varphi \} = \inf \{ I(\mu) : \mu \in \mathcal{P}(M) \text{ is } f^s\text{-invariant and } \varphi_t = \bar{B}_\mu(\varphi_t) \text{ for all } t \in [0, T] \}$ , i.e. with the same functional as in Remark 3.5.

#### 4. Averaging: Unbounded Time

Similarly to [Ki2] one can consider solutions of (1.5) on unbounded time intervals. Denote by  $\bar{X}(t) = \bar{X}^{(x)}(t)$  the solution of (1.11) with  $\mu = \mu_{\max}$  and  $\bar{X}(0) = x$ . Define the flow  $F^t$  on  $N$  by  $F^t x = \bar{X}^{(x)}(t)$ . Let  $V$  be a connected open set in  $N$  with a compact closure  $\bar{V}$ . Put

$$R(x, z) = \inf \{ S_{0T}(\varphi) : T \geq 0, \varphi \in C_{0T}(N), \varphi_0 = x, \varphi_T = z \} \quad (4.1)$$

and  $R_\delta(x) = \inf \{ R(x, z) : z \in \partial V \}$ , where  $\partial V$  is the boundary of  $V$ .

**4.1. Theorem.** *Let  $M$  be a basic hyperbolic set of the flow  $f^t$  (see, for instance, [B1]). Assume also that  $\mathcal{O}$  is an attracting fixed point of the dynamical system  $F^t$  whose domain of attraction contains a compact set  $\bar{V}$  which is the closure of a domain  $V$  with the smooth boundary  $\partial V$  and for each  $z \in \partial V$ ,  $(\bar{B}_{\mu_{\max}}(z), n(z)) < 0$  but there exists a  $f^t$ -invariant  $\mu$  such that  $(\bar{B}_\mu(z), n(z))_z > 0$ , where  $n(z)$  is the exterior normal to  $\partial V$  at  $z$  and  $(\cdot)_z$  denotes the Riemannian inner product in  $T_z N$ . Then for any  $x \in V$  and each  $\lambda > 0$ ,*

$$\lim_{\varepsilon \rightarrow 0} v_{T(\varepsilon)}^* \{ y \in CO^* : \exp(\varepsilon^{-1}(R(\mathcal{O}) - \lambda)) \leq \tau_{x,y}^\varepsilon(V) \leq \exp(\varepsilon^{-1}(R_\delta(\mathcal{O}) + \lambda)) \} = 1, \quad (4.2)$$

provided  $T(\varepsilon) \geq \exp(\varepsilon^{-1} R(0))$ .

*Proof.* I only sketch the proof here since ideologically it is close to the proof of Theorem 2.3 from [Ki2]. The main idea consists in the following. Suppose that  $R_\delta(\mathcal{O}) < \infty$ . Then for any  $\alpha > 0$  small enough and each  $\beta > 0$  there exist  $T_1 = T_{1;\alpha,\beta}$  and  $\varphi^\theta = \varphi^{\theta,\alpha,\beta} \in C_{0T}(N)$  such that  $\varphi_0^\theta = \mathcal{O}$ ,  $d_N(\varphi_{T_1}^\theta, V) \leq \alpha$ , and  $S_{0T_1}(\varphi^\theta) \leq R_\delta(\mathcal{O}) + \beta$ . One can pick up also  $T_2 = T_{2,\lambda}$  such that  $F^{T_2} \bar{V} \subset \{ y : d_M(y, \mathcal{O}) \leq \lambda \}$  with  $\lambda$  small enough. Next, one applies a Markov property type argument saying that either  $X_{x,y}^\varepsilon(t)$ ,  $x$  close to  $\mathcal{O}$ , exits from  $V$  staying close to  $\varphi^\theta$  for the time  $T_1$  or  $X_{x,y}(T_1) \in V$  and  $X_{x,y}(T_1 + T_2)$  is again close to  $\mathcal{O}$  and we repeat this process. The  $v_{\varepsilon^{-1}T_1, \delta}^*$ -measure of those  $y \in CO^*$  for which  $X_{x,y}^\varepsilon(t)$  performs the first type of motion can be estimated by Theorem 3.2. So first one has to approximate certain sequences of periodic orbits of periods  $\varepsilon^{-1}T_1$  and  $\varepsilon^{-1}T_2$  by a periodic orbit  $\gamma \in CO_\delta(T(\varepsilon))$  and then to estimate  $v_{T(\varepsilon)}^*$ -measure of points on these long orbits via products of measures  $v_{\varepsilon^{-1}T_1}^*$  and  $v_{\varepsilon^{-1}T_2}^*$  of points  $y$  on orbits from  $CO_{\varepsilon^{-1}T_1}$  and  $CO_{\varepsilon^{-1}T_2}$  for which  $X_{x,y}^\varepsilon(t)$  performs the motions described above. To do this one needs, in particular, the so-called shadowing property and not just the specification, since the latter provides only a periodic orbit of a period  $T(\varepsilon) \pm \text{const } \varepsilon T(\varepsilon)$  approximating the above shorter periodic orbits which is not good enough for (4.2). I refer the reader to [Ki2] in order to recover the details of the proof of Theorem 4.1 along the same lines.

#### 5. Averaging of PDE and SDE

In this section I consider large deviations via the contraction principle for parabolic second order partial differential equations and stochastic differential equations with coefficients incorporating a fast motion. Let  $M$  and  $N$  be compact

Riemannian manifolds and

$$L_y = \sum_{1 \leq i, j \leq n} a_{ij}(x, y) \frac{\partial^2}{\partial x^i \partial x^j} + \sum_{1 \leq i \leq n} b_i(x, y) \frac{\partial}{\partial x^i}, \quad x \in N, y \in M$$

be a family of uniformly elliptic operators on  $N$  with coefficients  $C^2$  in  $x$  and continuous in  $y$ . For  $\mu \in \mathcal{P}(M)$  I denote also by  $L_\mu$  the operator with the coefficients  $a_{ij}^\mu(x) = \int_M a_{ij}(x, y) d\mu(y)$  and  $b_i^\mu(x) = \int_M b_i(x, y) d\mu(y)$ ,  $1 \leq i, j \leq n$ .

Set, again,  $\mathcal{M} = \{\mu \in \mathcal{P}([0, T] \times M) : d\mu = \frac{1}{T} d\mu_t dt, \mu_t \in \mathcal{P}(M)\}$ . Consider the map  $\Psi_g : \mathcal{M} \rightarrow C([0, T] \times N)$  given by  $\Psi_g \mu = u$  with  $g \in C(N)$  and

$$u(t, x) = g(x) - \int_0^t L_{\mu_s} u(s, x) ds, \quad 0 \leq t \leq T. \quad (5.1)$$

Employing, for instance, the probabilistic representation of solutions of second order parabolic equations (see [KS]) one can show that  $\Psi_g$  is well-defined and that it is a continuous map provided  $\mathcal{M}$  is taken with the topology of weak convergence and  $C([0, T] \times N)$  with the supremum norm. Now if  $u_{y,g}^e(t, x)$  is the solution of

$$\frac{\partial u_{y,g}^e(t, x)}{\partial t} + L_{f^{t/e} u_{y,g}^e}(t, x) = 0, \quad u_{y,g}^e(0, x) = g \quad (5.2)$$

and  $\pi_y^e = T^{-1} \int_0^T \delta_{(s, f^{s/e} y)} ds$ , then  $u_{y,g}^e = \Psi_g \pi_y^e$ . Again knowing that  $\pi_y^e$  has large deviation estimates with the rate functional  $I_{0T}$  given by (3.9), with  $I$  defined either by (1.4) or by (3.14) from [Kil] depending on the measure for which the large deviation bounds are needed, one obtains the large deviation bounds in the averaging principle of the form

$$\mu_{T,\varepsilon} \{y : \rho_{0T}(u_{y,g}^e, u) < \beta\} \geq \exp\{-\varepsilon^{-1}(S_{0T}(u) + \lambda)\}, \quad (5.3)$$

$u(0, x) = g(x)$ , and

$$\mu_{T,\varepsilon} \{y : \rho_{0T}(u_{y,g}^e, \Phi_{0T}^a(g)) \geq \beta\} \leq \exp\{-\varepsilon^{-1}(a - \lambda)\}, \quad (5.4)$$

where  $\rho_{0T}(u, \tilde{u}) = \sup_{x \in M, t \in [0, T]} |u(t, x) - \tilde{u}(t, x)|$ ,  $\Phi_{0T}^a(g) = \{u \in C([0, T] \times N) : u(0, x) = g \text{ and } S_{0T}(u) \leq a\}$ ,

$$S_{0T}(u) = \inf_{\mu} \{I_{0T}(\mu) : \Psi_g \mu = u\} \\ = \int_0^T \inf \left\{ I(\mu) : \frac{\partial u}{\partial t} + L_\mu u = 0, \mu \in \mathcal{P}(M) \text{ is } f^s\text{-invariant} \right\} dt, \quad (5.5)$$

provided for Lebesgue almost all  $t \in [0, T]$  there exists a  $f^s$ -invariant  $\mu_t$  for which  $\frac{\partial u}{\partial t} + L_{\mu_t} u = 0$  and  $S_{0T}(u) = \infty$ , otherwise. Here  $\mu_{T,\varepsilon}$  is one of the measures connected with periodic orbits as in Sect. 3 or the measure  $m$  from Theorem 3.4 of [Kil].

Next, consider the following stochastic differential equation on  $N$ ,

$$dX_{x,y}^e(t) = \sigma(X_{x,y}^e(t), f^{t/e} y) d\omega_t + b(X_{x,y}^e(t), f^{t/e} y) dt, \quad X_{x,y}^e(0) = x, \quad (5.6)$$

where the matrix  $\sigma(x, y)$  satisfies  $\sigma(x, y) \sigma^*(x, y) = (a_{ij}(x, y))$  and  $\omega_t$  denotes the Wiener process. Then  $u_{y,g}^e(t, x) = Eg(X_{x,y}^e(t))$ , where  $E$  denotes the expectation (see, for instance, [KS]). Again by the contraction principle type argument one can

obtain the large deviation bounds for the process  $X_{x,y}^\varepsilon$ , itself. Define the map  $\tilde{\Psi}_{x,\omega}: \mathcal{M} \rightarrow C_{0T}(N)$  by  $\tilde{\Psi}_{x,\omega}\mu = \varphi(t, \omega)$ , where

$$\varphi(t) = x + \int_0^t \sigma_{\mu_s}(\varphi(s)) d\omega_s + \int_0^t b_{\mu_s}(\varphi(s)) ds, \quad t \in [0, T], \quad (5.7)$$

which is well-defined for almost all  $\omega$ . Here  $\sigma_\mu$  and  $b_\mu$  denote the integrals of  $\sigma$  and  $b$  in  $y$  against  $\mu \in \mathcal{P}(M)$ . In view of continuous dependence of solutions of stochastic differential equations on coefficients the map  $\tilde{\Psi}_{x,\omega}$  is continuous for almost all  $\omega$ . Since  $X_{x,y}^\varepsilon(\cdot, \omega) = \tilde{\Psi}_{x,\omega}\pi_y^\varepsilon$  one obtains the large deviation bounds for almost all  $\omega$  by the same routine as above with the functional  $S_{0T}^\omega(\varphi) = \inf\{I_{0T}(\mu): \tilde{\Psi}_{x,\omega}\mu = \varphi\}$ .

One can consider also a fast process  $\zeta_{t/\varepsilon}$  in place of  $f^{t/\varepsilon}y$  in (5.6), where  $\zeta_t$  is, for instance, a diffusion on  $M$  independent of  $w_t$  in (5.6). The large deviations for the resulting process follow in the same way, as above.

The need in averaging methods and, in particular, in large deviation estimates there emerges when one considers small perturbations of Hamiltonian systems (see, [AFr], and Remark 2.5 in [Ki2]). Deterministic perturbations by a small vector field lead to an averaging of the type considered in Sect. 3 and in [Ki2]. Random perturbations by a small diffusion lead to an averaging of the type considered in this section.

## References

- [A] Arnold, V.I.: Geometric Methods in the Theory of Ordinary Differential Equations. Berlin, Heidelberg, New York: Springer, 1983
- [B1] Bowen, R.: Periodic orbits for hyperbolic flows. Am. J. Math. **94**, 1–30 (1972)
- [B2] Bowen, R.: The equidistribution of closed geodesics. Am. J. Math. **94**, 413–423 (1972)
- [B3] Bowen, R.: Some systems with unique equilibrium states. Math. Syst. Theory **8**, 193–202 (1975)
- [B4] Bowen, R.: Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Math. **470**, Berlin, Heidelberg, New York: Springer, 1975
- [BW] Bowen, R., Walters, P.: Expansive one-parameter flows J. Diff. Eq. **12**, 180–193 (1972)
- [DGS] Denker, M., Grillenberger, G., Sigmund, K.: Ergodic Theory on Compact Spaces Lecture Notes in Math. **257**, Berlin, Heidelberg, New York: Springer, 1976
- [F] Franko, E.: Flows with unique equilibrium state. Am. J. Math. **99**, 486–514 (1977)
- [Fr] Freidlin, M.I.: The averaging principle and theorems on large deviations. Russ. Math. Sur. **33**, no 5, 107–160 (1978)
- [FW] Freidlin, M.I., Wentzell, A.D.: Random Perturbations of Dynamical Systems. Berlin, Heidelberg, New York: Springer, 1984
- [IT] Ioffe, A.D., Tikhomirov, V.M.: Theory of Extremal Problems. Amsterdam: North-Holland, 1979
- [Ki1] Kifer, Y.: Large deviations in dynamical systems and stochastic processes Trans. Am. Math Soc **321**, 505–524 (1990)
- [Ki2] Kifer, Y.: Averaging in dynamical systems and large deviations. Invent Math **110**, 337–370 (1992)
- [KS] Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Berlin, Heidelberg, New York: Springer, 1988
- [KT] Kitchens, B., Tuncel, S.: Finitary Measures for Subshifts of Finite Type and Sofic Systems. Memoirs of A.M.S., **338** (1985)
- [La] Lalley, S.: Distribution of periodic orbits of symbolic and Axiom A flows. Adv. in Appl. Math. **8**, 154–193 (1987)

- [Pa] Parry, W.: Bowen's equidistribution theory and the Dirichlet density theorem *Ergodic Th. and Dyn. Sys.* **4**, 117–134 (1984)
- [Ru] Ruelle, D.: *Thermodynamic Formalism*, Reading, Mass. Addison–Wesley, 1978
- [Si] Sigmund, K.: On dynamical systems with the specification property. *Trans. Am. Math Soc.* **190**, 285–299 (1974)
- [SV] Sanders, J.A., Verhurst, F.: *Averaging Methods in Nonlinear Dynamical Systems*. Berlin, Heidelberg, New York: Springer, 1985
- [Wa] Walters, P.: *An Introduction to Ergodic Theory*. Berlin, Heidelberg, New York: Springer, 1982
- [We] Weiss, B.: Subshifts of finite type and sofic systems. *Monatsh. fur. Math.* **77**, 462–474 (1973)

Communicated by J.-P. Eckmann