European Journal of Mathematics (2016) 2:1013–1022
DOI 10.1007/s40879-016-0119-z

RESEARCH ARTICLE

A uniform estimate for rate functions in large deviations

Luchezar Stoyanov[1]

Received: 24 July 2016 / Revised: 26 September 2016 / Accepted: 10 October 2016 /
Published online: 24 October 2016

Abstract Given Hölder continuous functions f and ψ on a subshift of finite type Σ_A^+ such that ψ is not cohomologous to a constant, the classical large deviation principle holds with a rate function $I_\psi \geqslant 0$ such that $I_\psi(p) = 0$ iff $p = \int \psi \, d\mu$, where $\mu = \mu_f$ is the equilibrium state of f. In this paper we derive a uniform estimate from below for I_ψ for p outside an interval containing $\widetilde{\psi} = \int \psi \, d\mu$, which depends only on the subshift Σ_A^+, the function f, the norm $|\psi|_\infty$, the Hölder constant of ψ and the integral $\widetilde{\psi}$. Similar results can be derived in the same way, e.g. for Axiom A diffeomorphisms on basic sets.

Keywords Large deviations · Rate function · Subshift of finite type · Equilibrium state

Mathematics Subject Classification 37A05 · 37B10 · 37D20

1 Introduction

Let $T: X \to X$ be a transformation preserving an ergodic probability measure μ on a set X. Given an observable $\psi: X \to \mathbb{R}$, Birkhoff's ergodic theorem implies that

$$\frac{\psi_n(x)}{n} = \frac{\psi(x) + \psi(T(x)) + \cdots + \psi(T^{n-1}(x))}{n} \to \int_X \psi \, d\mu$$

✉ Luchezar Stoyanov
luchezar.stoyanov@uwa.edu.au

[1] School of Mathematics and Statistics, University of Western Australia, 35 Stirling Hwy, Crawley, Perth, WA 6009, Australia

for μ-almost all $x \in X$. It follows from general large deviation principles (see [3,6,12]) that if X is a mixing basic set for an Axiom A diffeomorphism T, and f and ψ are Hölder continuous functions on X with *equilibrium states* $\mu = \mu_f$ and μ_ψ, respectively, and ψ is not cohomologous to a constant (see the definition below), then there exists a real-analytic *rate function* $I = I_\psi : \mathrm{Int}(\mathcal{J}_\psi) \to [0, \infty)$, where $\mathcal{J}_\psi = \left\{\int \psi\, dm : m \in \mathcal{M}_T\right\}$, such that

$$\lim_{\delta \to 0} \lim_{n \to \infty} \frac{1}{n} \log \mu\left(\left\{x \in X : \frac{\psi_n(x)}{n} \in (p - \delta, p + \delta)\right\}\right) = -I_\psi(p) \qquad (1)$$

for all $p \in \mathrm{Int}(\mathcal{J}_\psi)$. Here $\mathcal{M}_T$ is the *set of all T-invariant Borel probability measures* on X. Moreover, $I(p) = 0$ if and only if $p = \int \psi\, d\mu$, and the (closed) interval $\mathcal{J}_\psi$ is non-trivial, since ψ is not cohomologous to a constant.

Similar large deviation principles apply for any subshift of finite type $\sigma : \Sigma_A^+ \to \Sigma_A^+$ on a one-sided shift space

$$\Sigma_A^+ = \left\{\xi = (\xi_0, \xi_1, \dots, \xi_m, \dots) : 1 \leqslant \xi_i \leqslant s_0,\ A(\xi_i, \xi_{i+1}) = 1 \text{ for all } i \geqslant 0\right\}.$$

Here A is an $s_0 \times s_0$-matrix of 0's and 1's ($s_0 \geqslant 2$). We assume that A is *aperiodic*, i.e. there exists an integer $M > 0$ such that $A^M(i, j) > 0$ for all i, j (see, e.g. [7, Chapter 1]). The *shift map* σ is defined by $\sigma(\xi) = \xi'$, where $\xi'_i = \xi_{i+1}$ for all $i \geqslant 0$. We consider Σ_A^+ with a metric d_θ defined for some constant $\theta \in (0, 1)$ by $d_\theta(\xi, \eta) = 0$ if $\xi = \eta$ and $d_\theta(\xi, \eta) = \theta^k$ if $\xi \neq \eta$ and $k \geqslant 0$ is the maximal integer with $\xi_i = \eta_i$ for $0 \leqslant i \leqslant k$.

For any function $g : \Sigma_A^+ \to \mathbb{R}$ set

$$\mathrm{var}_k g = \sup\left\{|g(\xi) - g(\eta)| : \xi_i = \eta_i, 0 \leqslant i \leqslant k\right\}, \quad |g|_\theta = \sup\left\{\frac{\mathrm{var}_k g}{\theta^k} : k \geqslant 0\right\},$$
$$|g|_\infty = \sup\left\{|g(\xi)| : \xi \in \Sigma_A^+\right\}, \qquad \|g\|_\theta = |g|_\theta + |g|_\infty.$$

Denote by $\mathcal{F}_\theta(\Sigma_A^+)$ the *space of all functions* g on Σ_A^+ with $\|g\|_\theta < \infty$.

Two functions f, g on $\mathcal{F}_\theta(\Sigma_A^+)$ are called *cohomologous* if there exists a continuous function h on Σ_A^+ such that $f = g + h \circ \sigma - h$.

The *Ruelle transfer operator* $L_f : C(\Sigma_A^+) \to C(\Sigma_A^+)$ is defined by

$$L_f g(x) = \sum_{\sigma(y) = x} e^{f(y)} g(y).$$

Here $C(\Sigma_A^+)$ denotes the space of all continuous functions $g : \Sigma_A^+ \to \mathbb{R}$ with respect to the metric d_θ. Denote by $\Pr(\psi)$ the *topological pressure*

$$\Pr(\psi) = \sup_{m \in \mathcal{M}_\sigma} \left(h_\sigma(m) + \int \psi\, dm\right)$$

of ψ with respect to the map σ, where $\mathcal{M}_\sigma$ is the *set of all σ-invariant probability measures on* Σ_A^+ and $h_\sigma(m)$ is the *measure theoretic entropy* of m with respect to σ (see [7] or [10]). Given $\psi \in \mathcal{F}_\theta(\Sigma_A^+)$, there exists a unique σ-invariant probability measure μ_ψ on Σ_A^+ such that

$$\Pr(\psi) = h_\sigma(\mu_\psi) + \int \psi \, d\mu_\psi$$

(see, e.g. [7, Theorem 3.5]). The measure μ_ψ is called the *equilibrium state* of ψ.

For brevity throughout we write $\int h dm$ for $\int_{\Sigma_A^+} h dm$. In what follows we assume that $\theta \in (0, 1)$ is a fixed constant, $f : \Sigma_A^+ \to \mathbb{R}$ is a fixed function in $\mathcal{F}_\theta(\Sigma_A^+)$ and $\mu = \mu_f$.

As we mentioned earlier, it follows from the Large Deviation Theorem [3,6,12] that if ψ is not cohomologous to a constant, then there exists a real analytic *rate function* $I = I_\psi : \mathrm{Int}(\mathcal{I}_\psi) \to [0, \infty)$ with $I(p) = 0$ iff $p = \int \psi \, d\mu$ for which (1) holds. More precisely, we have

$$-I(p) = \inf\left\{\Pr(f + q\psi) - \Pr(f) - qp : q \in \mathbb{R}\right\}. \tag{2}$$

It is also known that

$$\left[\frac{d}{dq}\Pr(f + q\psi)\right]_{q=\eta} = \int \psi \, d\mu_{f+\eta\psi}, \tag{3}$$

and $\Pr(f + q\psi)$ is a strictly convex function of q (see [7,10] or [4]).

In his paper we derive an estimate from below for $I_\psi(p)$ for p outside an interval containing

$$\widetilde{\psi} = \int \psi \, d\mu.$$

The estimate depends only on $|\psi|_\infty$, $\widetilde{\psi}$, $|\psi|_\theta$ and some constants determined by the given function f. In what follows we use the notation $\min \psi = \min_{x \in \Sigma_A^+} \psi(x)$,

$$b = b_\psi = \max\{1, |\psi|_\theta\}, \qquad B_\psi = \widetilde{\psi} - \min\{0, \min \psi\}.$$

Since $\widetilde{\psi} > \min \psi$ (ψ is not cohomologous to a constant), we have $\widetilde{\psi} - \min \psi > 0$, so $B_\psi > 0$ always.

Theorem 1.1 *Let $f, \psi \in \mathcal{F}_\theta(\Sigma_A^+)$ be real-valued functions. Assume that ψ is not cohomologous to a constant, and let $0 < \delta_0 < B_\psi$. Then for all $p \notin [\widetilde{\psi} - \delta_0, \widetilde{\psi} + \delta_0]$ we have*

$$I_\psi(p) \geqslant \frac{\delta_0 q_0}{2},$$

where $q_0 = \min\{C, 1/b\}$ *for some constant* $C > 0$ *depending only on* $|f - \Pr(f)|_\infty$, $|f|_\theta$, $|\psi|_\infty$, $\widetilde{\psi}$ *and* δ_0.

The motivation to try to obtain estimates of the kind presented in Theorem 1.1 comes from attempts to get some kind of an 'approximate large deviation principle' for characteristic functions χ_K of arbitrary compact sets K of positive measure. In the special case when the boundary ∂K of K is 'relatively regular' (e.g. $\mu(\partial K) = 0$) large deviation results were established by Leplaideur and Saussol in [5], and also by Kachurovskii and Podvigin [2]. The next example presents a first step in the case of an arbitrary compact set K of positive measure.

Example 1.2 Let K be a compact subset of Σ_A^+ with $0 < \mu(K) < 1$, let $0 < \delta_0 \leqslant \mu(K)$, and let ψ be a Hölder continuous function that approximates χ_K from above, i.e. $0 \leqslant \psi \leqslant 1$, $\psi = 1$ on K and $\psi = 0$ outside a small neighbourhood V of K. Then $b = |\psi|_\theta \gg 1$ if V is sufficiently small, so q_0 in Theorem 1.1 has the form $q_0 = 1/b$. It then follows from Theorem 1.1 (in fact, from Lemma 2.3) that $I_\psi(p) \geqslant \delta_0/(2|\psi|_\theta)$ for $p \notin [\widetilde{\psi} - \delta_0, \widetilde{\psi} + \delta_0]$.

A result similar to Theorem 1.1 can be stated, e.g. for Axiom A diffeomorphisms on basic sets. Recall that if $F\colon M \to M$ is a C^1 Axiom A diffeomorphism on a Riemannian manifold M, a non-empty subset Λ of M is called a *basic set* for F if Λ is a locally maximal compact F-invariant subset of M which is not a single orbit, F is hyperbolic and transitive on Λ, and the periodic points of F in Λ are dense in Λ (see, e.g. [1] or [7, Appendix III]). It follows from the existence of Markov partitions that there exists a two-sided subshift of finite type $\sigma\colon \Sigma_A \to \Sigma_A$ and a continuous surjective map $\pi\colon \Sigma_A \to \Lambda$ such that: (i) $F \circ \pi = \pi \circ \sigma$, and (ii) for every Hölder continuous function g on Λ, $f = g \circ \pi \in \mathcal{F}_\theta$ for some $\theta \in (0, 1)$ and π is one-to-one almost everywhere with respect to the equilibrium state of f. Given a Hölder continuous function g on Λ, the rate function I_g is naturally related to the rate function I_f of $f = g \circ \pi$. On the other hand, f is cohomologous to a function $f' \in \mathcal{F}_{\sqrt{\theta}}(\Sigma_A)$ which depends on forward coordinates only, so $f' \in \mathcal{F}_{\sqrt{\theta}}(\Sigma_A^+)$. Applying Theorem 1.1 to f' provides a similar result for f and therefore for g.

For some hyperbolic systems, large deviation principles similar to (1), however with shrinking intervals, have been established recently in [8,9].

2 Proof of Theorem 1.1

2.1 The Ruelle–Perron–Frobenius Theorem

For convenience of the reader we state here a part of the estimates in [11] that will be used in this section.

Theorem 2.1 (Ruelle–Perron–Frobenius) *Let the* $s_0 \times s_0$*-matrix* A *and* $M > 0$ *be as in Sect. 1, let* $f \in \mathcal{F}_\theta(\Sigma_A^+)$ *be real-valued, and let* $b_f = \max\{1, |f|_\theta\}$. *Then:*

(i) *There exist a unique* $\lambda = \lambda_f > 0$, *a probability measure* $\nu = \nu_f$ *on* Σ_A^+ *and a positive function* $h = h_f \in \mathcal{F}_\theta(\Sigma_A^+)$ *such that* $L_f h = \lambda h$ *and* $\int h\, d\nu = 1$. *The* *spectral radius of L_f as an operator on $\mathcal{F}_\theta(\Sigma_A^+)$ is λ, and its essential spectral radius is $\theta\lambda$. The eigenfunction h satisfies*

$$\|h\|_\theta \leqslant \frac{6s_0^M b_f}{\theta^2(1-\theta)}\, e^{4b_f/(1-\theta)}\, e^{2M|f|_\infty}, \qquad \min h \geqslant \frac{1}{e^{2b_f/(1-\theta)}\, s_0^M\, e^{2M|f|_\infty}}.$$

Moreover,

$$\frac{\min h}{|h|_\infty}\,\lambda^n \leqslant L_f^n 1 \leqslant \frac{|h|_\infty}{\min h}\,\lambda^n,$$

for any integer $n \geqslant 0$.

(ii) *The probability measure $\widehat{\nu} = h\nu$ (this is the so-called* equilibrium state *of f) is σ-invariant and $\widehat{\nu} = \nu_{\widehat{f}}$, where $\widehat{f} = f - \log(h\circ\sigma) + \log h - \log\lambda$. Moreover $L_{\widehat{f}} 1 = 1$.*

(iii) *For every $g \in \mathcal{F}_\theta(\Sigma_A^+)$ and every integer $n \geqslant 0$ we have*

$$\left\| \frac{1}{\lambda^n}\, L_f^n g - h\int g\, d\nu \right\|_\theta \leqslant D\rho^n\, \|g\|_\theta,$$

where we can take

$$\rho = \left(1 - \frac{1-\theta}{4s_0^{2M}\, e^{8\theta b_f/(1-\theta)}\, e^{4M|f|_\infty}}\right)^{1/2M} \in (0,1)$$

and

$$D = 10^8\, \frac{b_f^7}{\theta^{10}(1-\theta)^8}\, s_0^{17M}\, e^{40 b_f/(1-\theta)}\, e^{33M|f|_\infty}.$$

Remark 2.2 The constants that appear in the above estimates are not optimal. The proof of [11, Theorem 2] follows that in [1, Section 1.B] with a more careful analysis of the estimates involved. The main point here is that, apart from their obvious dependence on parameters related to the subshift of finite type $\sigma : \Sigma_A^+ \to \Sigma_A^+$, these constants can be taken to depend only on $|f|_\theta$ and $|f|_\infty$.

2.2 Reductions

Let $f \in \mathcal{F}_\theta(\Sigma_A^+)$ be the fixed function from Sect. 1 and let $\mu = \mu_f$ be as before. It follows from the properties of pressure (see, e.g. [10] or [7]) that $\Pr(g+c) = \Pr(g)+c$ for every continuous function g and every constant $c \in \mathbb{R}$. Thus, replacing f by $f - \Pr(f)$, we may assume that $\Pr(f) = 0$. Moreover, if g and h are cohomologous continuous functions on Σ_A^+, then $\Pr(g) = \Pr(h)$ and the equilibrium states μ_g of g and μ_h of h on Σ_A^+ coincide. Since f is cohomologous to a function $\phi \in \mathcal{F}_\theta(\Sigma_A^+)$ with $L_\phi 1 = 1$ (see, e.g. [7]), it is enough to prove the main result with f replaced by such a function ϕ. Moreover, $|\phi|_\infty$ and $|\phi|_\theta$ can be estimated by means of $|f - \Pr(f)|_\infty$ and $|f|_\theta$ [see e.g. Theorem 2.1 (ii)].

So, from now on we will assume that $L_\phi 1 = 1$. It then follows that $\Pr(\phi) = 0$. Let $\mu = \mu_\phi$ be the *equilibrium state* of ϕ on Σ_A^+.

For the proof of Theorem 1.1 we may assume that $\psi \geqslant 0$. Indeed, assuming the statement of the theorem is true in this case, suppose ψ takes negative values. Set $\psi_1 = \psi + c$, where $c = -\min \psi$. Then $\psi_1 \geqslant 0$. Moreover, $\widetilde{\psi}_1 = \int \psi_1 d\mu = \widetilde{\psi} + c$, $B_{\psi_1} = B_\psi$, and for $p_1 = p + c$ we have

$$\begin{aligned}\Gamma_1(q) &= p_1 q - \Pr(\phi + q\psi_1)\\ &= (p+c)q - \Pr(\phi + q\psi + qc) = pq - \Pr(\phi + q\psi) = \Gamma(q)\end{aligned}$$

for all $q \in \mathbb{R}$. Therefore (2) implies

$$\begin{aligned}I_\psi(p) &= \sup\{pq - \Pr(\phi + q\psi) : q \in \mathbb{R}\}\\ &= \sup\{p_1 q - \Pr(\phi + q\psi_1) : q \in \mathbb{R}\} = I_{\psi_1}(p_1).\end{aligned}$$

Moreover, if $0 < \delta_0 < B_\psi = B_{\psi_1}$, then $p \notin [\widetilde{\psi} - \delta_0, \widetilde{\psi} + \delta_0]$ is equivalent to $p_1 = p + c \notin [\widetilde{\psi}_1 - \delta_0, \widetilde{\psi}_1 + \delta_0]$. Since $|\psi_1|_\theta = |\psi|_\theta$ and $|\psi_1|_\infty \leqslant 2|\psi|_\infty$, using Theorem 1.1 for $I_{\psi_1}(p_1)$ and changing appropriately the value of the constant q_0, we get a similar estimate for $I_\psi(p)$.

2.3 Proof of Theorem 1.1 for $\psi \geqslant 0$

From now on we will assume that $\phi, \psi \in \mathcal{F}_\theta(\Sigma_A^+)$ are fixed real-valued functions such that $\psi \geqslant 0$, ψ is not cohomologous to a constant, and

$$L_\phi 1 = 1. \tag{4}$$

Given any $q \in \mathbb{R}$, set

$$f_q = \phi + q\psi, \qquad L_q = L_{f_q}.$$

In what follows we will assume

$$|q| \leqslant q_0 \leqslant \frac{1}{b} \tag{5}$$

for some constant $q_0 > 0$ which will be chosen below. Then $|f_q|_\theta \leqslant |\phi|_\theta + 1$ for all q with (5), and also $|f_q|_\infty \leqslant |\phi|_\infty + |\psi|_\infty$. Thus, setting

$$C_0 = \|\phi\|_\theta + 2\max\{|\psi|_\infty, 1\} \geqslant 1,$$

we have

$$\|f_q\|_\theta \leqslant C_0, \qquad |q| \in [0, q_0]. \tag{6}$$

Let ν_q be the *probability measure* on Σ_A^+ with

$$L_q^* \nu_q = \lambda_q \nu_q, \tag{7}$$

where λ_q is the *maximal eigenvalue* of $L_q = L_{f_q}$, and let $h_q > 0$ be a *corresponding normalised eigenfunction*, i.e. $h_q \in \mathcal{F}_\theta(\Sigma_A^+)$, $L_q h_q = \lambda_q h_q$ and $\int h_q d\nu_q = 1$. Then $\mu_q = h_q \nu_q$ is the *equilibrium state* of f_q, i.e. $\mu_q = \mu_{\phi+q\psi}$. Clearly $h_0 = 1$ and $\mu_0 = \mu$.

Using the uniform estimates in Theorem 2.1, it follows from (6) that there exist constants $D \geqslant 1$ and $\rho \in (0, 1)$, depending on C_0 but not on q_0, such that

$$\left\| \frac{1}{\lambda_q^n} L_q^n g - h_q \int g \, d\nu_q \right\|_\theta \leqslant D\rho^n \|g\|_\theta \tag{8}$$

for all integers $n \geqslant 0$, all functions $g \in \mathcal{F}_\theta(\Sigma_A^+)$ and all q with $|q| \in [0, q_0]$.

Set $L = L_\phi$. Given $x \in \Sigma_A^+$ and $m \geqslant 0$, set $g_m(x) = g(x) + g(\sigma x) + \cdots + g(\sigma^{m-1} x)$.
It follows from (7) with $g = 1$ that $\lambda_q = \int L_q 1 d\nu_q$. Now

$$(L_q 1)(x) = \sum_{\sigma y = x} e^{f_q(y)} = \sum_{\sigma y = x} e^{\phi(y) + q\psi(y)} \leqslant e^{q_0 |\psi|_\infty} (L1)(x) = e^{q_0 |\psi|_\infty}$$

for all $x \in \Sigma_A^+$ implies $\lambda_q \leqslant e^{q_0|\psi|_\infty}$. Similarly, $\lambda_q \geqslant e^{-q_0|\psi|_\infty}$. Thus,

$$e^{-q_0 C_0} \leqslant \lambda_q \leqslant e^{q_0 C_0}, \qquad |q| \leqslant q_0. \tag{9}$$

To estimate h_q for q with (5), first use (8) with $g = 1$ to get

$$\left\| \frac{1}{\lambda_q^n} L_q^n 1 - h_q \right\|_\theta \leqslant D\rho^n.$$

Using (4), this gives

$$\begin{aligned} h_q(x) &\leqslant \frac{(L_q^n 1)(x)}{\lambda_q^n} + D\rho^n = \frac{1}{\lambda_q^n} \sum_{\sigma^n y = x} e^{(\phi + q\psi)_n(y)} + D\rho^n \\ &\leqslant \frac{e^{q_0 C_0 n}}{\lambda_q^n} (L1)(x) + D\rho^n \leqslant e^{2q_0 C_0 n} + D\rho^n \end{aligned}$$

for all $x \in \Sigma_A^+$ and $n \geqslant 0$. Similarly,

$$h_q \geqslant \frac{e^{-q_0 C_0 n}}{\lambda_q^n} (L1) - D\rho^n \geqslant e^{-2q_0 C_0 n} - D\rho^n$$

for all $n \geqslant 0$. Thus,

$$\max\left\{0, e^{-2q_0C_0n} - D\rho^n\right\} \leqslant h_q \leqslant e^{2q_0C_0n} + D\rho^n, \qquad n \geqslant 0, \quad |q| \leqslant q_0. \tag{10}$$

From now on we will assume that $p \notin [\widetilde{\psi} - \delta_0, \widetilde{\psi} + \delta_0]$ *is fixed*. Consider the function

$$\Gamma(q) = pq - \Pr(\phi + q\psi), \qquad q \in \mathbb{R}.$$

Then $I(p) = \sup_{q\in\mathbb{R}} \Gamma(q)$. Clearly, $\Gamma(0) = 0$ and moreover by (3),

$$\Gamma'(q) = p - \int \psi \, d\mu_{\phi+q\psi}. \tag{11}$$

In particular, $\Gamma'(0) = p - \widetilde{\psi} \notin [-\delta_0, \delta_0]$.

We will now estimate the integral in the right-hand side of (11). Let $\alpha > 0$ be the constant so that $\rho_1 = \max\{\rho, \theta\} = e^{-\alpha}$.

Lemma 2.3 *Assume that* $\psi \geqslant 0$ *on* Σ_A^+ *and* $0 < \delta_0 < B_\psi = \widetilde{\psi}$. *Set*

$$q_0 = \min\left\{\frac{\delta_0}{100C_0^2 n_0}, \frac{1}{b}\right\}, \tag{12}$$

where n_0 *is the integer with*

$$n_0 - 1 \leqslant \frac{1}{\alpha}\left|\log \frac{\delta_0}{16C_0D}\right| < n_0. \tag{13}$$

Then $\Gamma(q_0) \geqslant \delta_0 q_0/2$ *and* $\Gamma(-q_0) \geqslant \delta_0 q_0/2$.

Proof For any $q \in [0, q_0]$ and any integer $n \geqslant 0$, (7), (9) and (10) yield

$$\begin{aligned}\int \psi \, d\mu_q &= \int \psi h_q \, d\nu_q = \frac{1}{\lambda_q^n}\int L_q^n(\psi h_q)\, d\nu_q \\ &= \frac{1}{\lambda_q^n}\int \sum_{\sigma^n y = x} e^{(\phi+q\psi)_n(y)}\psi(y) h_q(y)\, d\nu_q(x) \\ &\leqslant e^{2q_0C_0n}\left(e^{2qC_0n} + D\rho^n\right)\int L^n\psi \, d\nu_q.\end{aligned}$$

It follows from (8) with $q = 0$ and $g = \psi$ and the choice of C_0 that

$$\left|L^n\psi - \int \psi \, d\mu\right| = \left|L^n\psi - \int \psi \, d\nu\right| \leqslant D\rho^n C_0, \tag{14}$$

therefore $L^n\psi \leqslant \widetilde{\psi} + C_0D\rho^n$. Combining this with the above gives

$$\int \psi \, d\mu_q \leqslant e^{2q_0C_0n}\left(e^{2q_0C_0n} + D\rho^n\right)\left(\widetilde{\psi} + C_0D\rho^n\right). \tag{15}$$

Let $n_0 = n_0(f, \theta, \delta_0) \geqslant 1$ be the *integer* such that

$$e^{-n_0\alpha} < \frac{\delta_0}{16C_0D} \leqslant e^{-(n_0-1)\alpha}. \tag{16}$$

Then $-n_0\alpha < \log \delta_0/(16C_0D) \leqslant -(n_0-1)\alpha$, so n_0 satisfies (13). With this choice of n_0 define q_0 by (12). Then for $q \in [0, q_0]$ we have $12qC_0^2n_0 \leqslant \delta_0/8$ and so $12qC_0n_0 \leqslant 1$. It now follows from (15) with $q \in [0, q_0]$ and $n = n_0$, $0 < \delta_0 \leqslant B_\psi = \widetilde{\psi} \leqslant C_0$, (16) and the fact that $e^x \leqslant 1 + 3x$ for $x \in [0, 1]$ that

$$\begin{aligned}
\int \psi\, d\mu_q &\leqslant \left(e^{4q_0C_0n_0} + De^{2q_0C_0n_0}e^{-\alpha n_0}\right)\left(\widetilde{\psi} + C_0De^{-\alpha n_0}\right)\\
&\leqslant \left(1 + 12q_0C_0n_0 + (1 + 6q_0C_0n_0)\,\frac{\delta_0}{16C_0}\right)\left(\widetilde{\psi} + \frac{\delta_0}{16}\right)\\
&\leqslant \widetilde{\psi} + 12q_0C_0^2n_0 + (1 + 6q_0C_0n_0)\,\frac{\delta_0}{16} + \left(2 + 2\,\frac{\delta_0}{16C_0}\right)\frac{\delta_0}{16}\\
&\leqslant \widetilde{\psi} + \frac{\delta_0}{8} + \frac{\delta_0}{8} + \frac{3\delta_0}{16} < \widetilde{\psi} + \frac{\delta_0}{2}.
\end{aligned}$$

Thus, in the case $p \geqslant \widetilde{\psi} + \delta_0$, it follows from (11) that $\Gamma'(q) \geqslant \delta_0/2$ for all $q \in [0, q_0]$, and therefore $\Gamma(q_0) \geqslant \delta_0 q_0/2$.

Next, assume that $p \leqslant \widetilde{\psi} - \delta_0$. We will now estimate $\int \psi\, d\mu_q$ from below for $q \in [-q_0, 0]$. As in the previous estimate, using (9) and (10), for such q we get

$$\begin{aligned}
\int \psi\, d\mu_q &= \int \psi h_q\, d\nu_q = \frac{1}{\lambda_q^{n_0}} \int L_q^{n_0}(\psi h_q)\, d\nu_q\\
&= \frac{1}{\lambda_q^{n_0}} \int \sum_{\sigma^{n_0} y = x} e^{(\phi + q\psi)_{n_0}(y)}\, \psi(y) h_q(y)\, d\nu_q(x)\\
&\geqslant e^{-2q_0C_0n_0}\left(e^{-2q_0C_0n_0} - D\rho^{n_0}\right) \int L^{n_0}\psi\, d\nu_q.
\end{aligned}$$

Notice that by the choice of q_0 and n_0 we have $e^{-2q_0C_0n_0} - D\rho^{n_0} > 0$. In fact, it follows from $e^{-x} > 1 - x$ for $x > 0$ that $e^{-2q_0C_0n_0} > 1 - 2q_0C_0n_0$, while (16) implies $D\rho^{n_0} < \delta_0/(16C_0)$. Thus,

$$e^{-2q_0C_0n_0} - D\rho^{n_0} > 1 - 2q_0C_0n_0 - \frac{\delta_0}{16C_0} > 1 - \frac{\delta_0}{8C_0}.$$

On the other hand, (14) yields $\int L^{n_0}\psi\, d\nu_q \geqslant \widetilde{\psi} - DC_0\rho^{n_0} > \widetilde{\psi} - \delta_0/16$. Hence for $q \in [-q_0, 0]$ we get

$$\int \psi \, d\mu_q \geqslant (1 - 2q_0 C_0 n_0)\left(1 - \frac{\delta_0}{8C_0}\right)\left(\widetilde{\psi} - \frac{\delta_0}{16}\right)$$
$$\geqslant \left(1 - 2q_0 C_0 n_0 - \frac{\delta_0}{8C_0}\right)\left(\widetilde{\psi} - \frac{\delta_0}{16}\right)$$
$$\geqslant \widetilde{\psi} - \widetilde{\psi}\left(2q_0 C_0 n_0 + \frac{\delta_0}{8C_0}\right) - \frac{\delta_0}{16} \geqslant \widetilde{\psi} - \frac{\delta_0}{50} - \frac{\delta_0}{8} - \frac{\delta_0}{16} > \widetilde{\psi} - \frac{\delta_0}{2}.$$

Thus, for $q \in [-q_0, 0]$ we have

$$\Gamma'(q) = p - \int \psi \, d\mu_q \leqslant \widetilde{\psi} - \delta_0 - (\widetilde{\psi} - \delta_0/2) \leqslant -\frac{\delta_0}{2},$$

and therefore $\Gamma(-q_0) \geqslant \delta_0 q_0/2$. □

Proof of Theorem 1.1 Assume again that $\psi \geqslant 0$. Let $p \geqslant \widetilde{\psi} + \delta_0$. Then $I(p) = \sup_{q \in \mathbb{R}} \Gamma(q)$, so by Lemma 2.3, $I(p) \geqslant \Gamma(q_0) \geqslant \delta_0 q_0/2$. Similarly, for $p \leqslant \widetilde{\psi} - \delta_0$ we get $I(p) \geqslant \delta_0 q_0/2$.

As explained in Sect. 2.2, the case of an arbitrary real-valued $\psi \in \mathcal{F}_\theta(\Sigma_A^+)$ follows from the case $\psi \geqslant 0$. □

Acknowledgments Thanks are due to the referees for their valuable comments and suggestions.

References

1. Bowen, R.: Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms. Lecture Notes in Mathematics, vol. 470. Springer, Berlin (1975)
2. Kachurovskii, A.G., Podvigin, I.V.: Large deviations and rates of convergence in the Birkhoff ergodic theorem: from Hölder continuity to continuity. Dokl. Math. **93**(1), 6–8 (2016)
3. Kifer, Yu.: Large deviations in dynamical systems and stochastic processes. Trans. Amer. Math. Soc. **321**(2), 505–524 (1990)
4. Lalley, S.P.: Distribution of periodic orbits of symbolic and Axiom A flows. Adv. in Appl. Math. **8**(2), 154–193 (1987)
5. Leplaideur, R., Saussol, B.: Large deviations for return times in non-rectangle sets for Axiom A diffeomorphisms. Discrete Contin. Dyn. Syst. **22**(1–2), 327–344 (2008)
6. Orey, S., Pelikan, S.: Deviations of trajectory averages and the defect in Pesin's formula for Anosov diffeomorphisms. Trans. Amer. Math. Soc. **315**(2), 741–753 (1989)
7. Parry, W., Pollicott, M.: Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics. Astérisque, vol. 187–188. Société Mathmatique de France, Paris (1990)
8. Petkov, V., Stoyanov, L.: Sharp large deviations for some hyperbolic systems. Ergodic Theory Dynam. Systems **35**(1), 249–273 (2015)
9. Pollicott, M., Sharp, R.: Large deviations, fluctuations and shrinking intervals. Comm. Math. Phys. **290**(1), 321–334 (2009)
10. Ruelle, D.: Thermodynamic Formalism. Encyclopedia of Mathematics and its Applications, vol. 5. Addison-Wesley, Reading (1978)
11. Stoyanov, L.: On the Ruelle–Perron–Frobenius theorem. Asymptot. Anal. **43**(1–2), 131–150 (2005)
12. Young, L.-S.: Large deviations in dynamical systems. Trans. Amer. Math. Soc. **318**(2), 525–543 (1990)