ISSN 0001-4346, Mathematical Notes, 2013, Vol. 94, No. 4, pp. 524–531. © Pleiades Publishing, Ltd., 2013.
Original Russian Text © A. G. Kachurovskii, I. V. Podvigin, 2013, published in Matematicheskie Zametki, 2013, Vol. 94, No. 4, pp. 569–577.

# Large Deviations and the Rate of Convergence in the Birkhoff Ergodic Theorem

**A. G. Kachurovskii**[1*] **and I. V. Podvigin**[2**]

[1] *Sobolev Institute of Mathematics, Russian Academy of Sciences, Novosibirsk, Russia*
[2] *Novosibirsk State University, Novosibirsk, Russia*

Received February 27, 2012

**Abstract**—For bounded averaged functions, we prove the equivalence of the power-law and exponential rates of convergence in the Birkhoff individual ergodic theorem with the same asymptotics of the probability of large deviations in this theorem.

**DOI:** 10.1134/S0001434613090228

Keywords: *pointwise ergodic theorem, rates of convergence in ergodic theorems, large deviations, billiards, Anosov systems.*

## 1. INTRODUCTION

Suppose that $(\Omega, \mathfrak{F}, \lambda)$ is a space with probability measure, $T$ is its endomorphism, i.e., a mapping $T\colon \Omega \to \Omega$ such that, for all $A \in \mathfrak{F}$, the set $T^{-1}A$ belongs to $\mathfrak{F}$, and $\lambda(A) = \lambda(T^{-1}A)$. Also, let $\{T^t,\, t \in \mathbb{R}^+\}$ be a semiflow on $(\Omega, \mathfrak{F}, \lambda)$, i.e., a one-parameter semigroup of endomorphisms $T^t$ of this space such that, for any measurable function $f(\omega)$ on $\Omega$, the function $f(T^t\omega)$ is measurable on the direct product $\Omega \times \mathbb{R}^+$. For $f \in L_1(\Omega)$, $\omega \in \Omega$, $n \in \mathbb{N}$, and $t \in \mathbb{R}^+$, we denote

$$A_n f(\omega) = \frac{1}{n}\sum_{k=0}^{n-1} f(T^k\omega) \qquad \text{and} \qquad \overline{A}_t f(\omega) = \frac{1}{t}\int_0^t f(T^\tau\omega)\,d\tau.$$

Then the Birkhoff individual ergodic theorem asserts the existence of the $\lambda$-a.e. limits

$$f^* = \lim_{n\to\infty} A_n f \qquad \text{and} \qquad \overline{f}^* = \lim_{t\to\infty} \overline{A}_t f,$$

and the equalities

$$\int f^*\,d\lambda = \int f\,d\lambda = \int \overline{f}^*\,d\lambda.$$

For each $\varepsilon > 0$, the decrease of the probabilities

$$\mathsf{P}_n^\varepsilon = \lambda\Big\{\sup_{k\ge n} |A_k f - f^*| \ge \varepsilon\Big\} \qquad \text{as} \quad n \to \infty,$$
$$\overline{\mathsf{P}}_t^\varepsilon = \lambda\Big\{\sup_{s\ge t} |\overline{A}_s f - \overline{f}^*| \ge \varepsilon\Big\} \qquad \text{as} \quad t \to \infty,$$

determines the rate of convergence in that theorem.

Note that the validity of the condition

$$\lim_{n\to\infty} \mathsf{P}_n^\varepsilon = 0 \qquad \text{or} \qquad \lim_{t\to\infty} \overline{\mathsf{P}}_t^\varepsilon = 0$$

*E-mail: agk@math.nsc.ru
**E-mail: ivan_podvigin@ngs.ru

for all $\varepsilon > 0$ is equivalent to the convergence of $A_n f$ or $\overline{A}_t f$ a.e. as $n \to \infty$ or $t \to \infty$, respectively.

Along with this quantities, we shall consider the probabilities of large deviations

$$\mathsf{p}_n^\varepsilon = \lambda\{|A_n f - f^*| \ge \varepsilon\} \qquad \text{and} \qquad \overline{\mathsf{p}}_t^\varepsilon = \lambda\{|\overline{A}_t f - \overline{f}^*| \ge \varepsilon\}.$$

The validity of the condition $\lim_{n\to\infty} \mathsf{p}_n^\varepsilon = 0$ or $\lim_{t\to\infty} \overline{\mathsf{p}}_t^\varepsilon = 0$ for all $\varepsilon > 0$ is, by definition, a property of the convergence of $A_n f$ or $\overline{A}_t f$ in measure as $n \to \infty$ or $t \to \infty$, respectively.

It turns out that if the averaged function $f$ is essentially bounded, then the rate of decrease of the quantities $\mathsf{P}_n^\varepsilon$ and $\mathsf{p}_n^\varepsilon$ ($\overline{\mathsf{P}}_t^\varepsilon$ and $\overline{\mathsf{p}}_t^\varepsilon$) has identical asymptotics as $n \to \infty$ ($t \to \infty$) for a wide class of rates of this decrease (see Theorem 1 below), including power-law and exponential rates, which are important for applications. This makes it possible to obtain estimates of the rate of convergence in the Birkhoff ergodic theorem, i.e., for each $\varepsilon > 0$, estimates of rates of decrease of $\mathsf{P}_n^{2\varepsilon}$ and $\overline{\mathsf{P}}_t^{2\varepsilon}$ as $n \to \infty$ and $t \to \infty$, respectively, can be determined from the well-known rates of decrease of the probabilities of large deviations $\mathsf{p}_n^\varepsilon$ and $\overline{\mathsf{p}}_t^\varepsilon$ (Theorems 2 and 3).

The application of the results obtained to dynamical systems modeled by the Young tower with polynomial tail readily gives power-law estimates of the rate of convergence in the Birkhoff ergodic theorem for systems such as the Bunimovich stadium and some classes of semiscattering billiards and billiards with cusps in the case of the Hölder property of the averaged function $f$. For such systems, power-law estimates of large deviations were established in [1], [2].

Again, in the case of the Hölder property of the averaged function, Theorem 3 readily gives exponential estimates of the rate of convergence in the individual ergodic theorem: for Anosov systems (large deviations were studied, for example, in [3]); for uniformly hyperbolic dynamical systems modeled by the Young tower with exponential tail, e.g., a classical scattering billiards such as a Lorentz gas on the two-dimensional torus (for estimates of large deviations, see in [1], [2]); as well as for Teichmüller flows (estimates of large deviations were established in [4]).

Note that, in the case of a square-integrable function $f$, power-law estimates of the rate of convergence in the Birkhoff theorem resulting from the well-known power-law rate of convergence in the von Neumann theorem (obtained, in turn, from the power-law decrease of the correlation coefficients or from the power-law singularity of the spectral measure in a neighborhood of zero) were recently established in [5], [6].

## 2. COINCIDING ASYMPTOTICS OF THE RATE OF CONVERGENCE AND OF LARGE DEVIATIONS

For essentially bounded functions $f$, the power-law rate of decrease as $n \to \infty$ of the probabilities of large deviations $\mathsf{p}_n^\varepsilon$ is equivalent to the power-law decrease, with the same exponent, of the rate of decrease of the probabilities $\mathsf{P}_n^\varepsilon$ [7, Theorem 12]. The following theorem establishes a similar result for more general rates of decrease for both endomorphisms and semiflows.

**Theorem 1.** *Let $f \in L_\infty(\Omega)$, and let the function $\varphi(x)\colon \mathbb{R}^+ \mapsto \mathbb{R}^+$ be monotone decreasing (for $x > x_0$) to zero so that*

$$\int_n^\infty \frac{\varphi(x)}{x}\, dx = O(\varphi(n)) \qquad as \quad n \to \infty. \tag{1}$$

*Then the following two assertions are equivalent:*

1) $\mathsf{p}_n^\varepsilon = O(\varphi(n))$ *as* $n \to \infty$ *for any* $\varepsilon > 0$;

2) $\mathsf{P}_n^\varepsilon = O(\varphi(n))$ *as* $n \to \infty$ *for any* $\varepsilon > 0$.

*In the case of continuous time, the following similar assertions are equivalent:*

3) $\overline{\mathsf{p}}_t^\varepsilon = O(\varphi(t))$ *as* $t \to \infty$ *for any* $\varepsilon > 0$;

4) $\overline{\mathsf{P}}_t^\varepsilon = O(\varphi(t))$ *as* $t \to \infty$ *for any* $\varepsilon > 0$.

The proof of the theorem is based on a technical lemma. For its formulation, we introduce some notation that be used throughout the paper. Let $f \in L_\infty(\Omega)$, $\Delta = \|f - f^*\|_\infty$ in the case of a discrete dynamical system and $\Delta = \|f - \overline{f}^*\|_\infty$ in the case a continuous system. In what follows we shall assume that the function $f$ is not invariant with respect to the dynamical system under consideration; otherwise, there is nothing to estimate. For each $\varepsilon > 0$, we set

$$r = r(\varepsilon) = 1 + \frac{\varepsilon}{\Delta}.$$

**Lemma 1.** *Let $f \in L_\infty(\Omega)$. Then, for any $\varepsilon > 0$,*

$$\mathsf{P}_n^{2\varepsilon} \le \sum_{k=0}^{\infty} \lambda\Big\{ \sup_{nr^k \le m < nr^{k+1}} |A_m f - f^*| \ge 2\varepsilon \Big\} \le \sum_{k=0}^{\infty} \mathsf{p}_{n_k}^{\varepsilon} \qquad \textit{for all} \quad n \ge 1,$$

$$\overline{\mathsf{P}}_t^{2\varepsilon} \le \sum_{k=0}^{\infty} \lambda\Big\{ \sup_{tr^k \le s < tr^{k+1}} |\overline{A}_s f - \overline{f}^*| \ge 2\varepsilon \Big\} \le \sum_{k=0}^{\infty} \overline{\mathsf{p}}_{t_k}^{\varepsilon} \qquad \textit{for all} \quad t > 0,$$

*where $n_k$ is the minimal integer not less than $nr^k$ and $t_k = tr^k$.*

**Proof of Lemma 1.** For discrete time, the proof was obtained in [7] (in the proof of Theorem 12). Here we present a similar proof for the continuous case. Replacing $f$ by $f - \overline{f}^*$, we can assume that $\overline{f}^* \equiv 0$ a.e.

Since $|\overline{A}_t f| \le \Delta$ for all $t > 0$, it follows that, for all $\varepsilon > \Delta/2$, the assertion of Lemma 1 is obviously valid. Let us prove it for $\varepsilon \le \Delta/2$. Suppose that $|\overline{A}_s f| < \varepsilon$ for some $s$; then

$$|\overline{A}_{s+t} f| = \frac{1}{s+t}\left| \int_0^s f \circ T^\tau \, d\tau + \int_s^{s+t} f \circ T^\tau \, d\tau \right| \le \frac{s|\overline{A}_s f| + t\|f\|_\infty}{s+t} < \frac{s\varepsilon + t\|f\|_\infty}{s+t} < 2\varepsilon$$

for all $t \in [0, s\varepsilon/\Delta]$. Hence we have

$$\Big\{ \sup_{tr^k \le s < tr^{k+1}} |\overline{A}_s f| \ge 2\varepsilon \Big\} \subseteq \{|\overline{A}_{t_k} f| \ge \varepsilon\}, \qquad t_k = tr^k,$$

and

$$\overline{\mathsf{P}}_t^{2\varepsilon} \le \sum_{k=0}^{\infty} \lambda\Big\{ \sup_{tr^k \le s < tr^{k+1}} |\overline{A}_s f| \ge 2\varepsilon \Big\} \le \sum_{k=0}^{\infty} \overline{\mathsf{p}}_{t_k}^{\varepsilon}.$$

Lemma 1 is proved. □

**Remark 1.** The quantity $\mathsf{P}_n^{\varepsilon}$ (and all quantities similar to it) depends on the measure $\lambda$ in two ways, because it is a $\lambda$-measure of the set

$$\Omega_\lambda = \{\sup_{k \ge n} |A_k f - f^*| \ge \varepsilon\}$$

whose dependence on $\lambda$ is "hidden" in the function $f^*$. The invariance of the measure is used only in this second dependence. Therefore, the assertion of Lemma 1 remains valid if, instead of the quantity $\mathsf{P}_n^{\varepsilon}$ (and similar quantities), we consider the quantity $\mathsf{P}_n^{\varepsilon}(\mu) = \mu\{\Omega_\lambda\}$ (and similar quantities) for an arbitrary probability measure $\mu$, not necessarily invariant.

**Proof of Theorem 1.** Assertion 2) immediately implies 1), because always $\mathsf{p}_n^{\varepsilon} \le \mathsf{P}_n^{\varepsilon}$. Let us obtain 2) from 1). Suppose that, for each $\varepsilon > 0$, the following relation holds: $\mathsf{p}_n^{\varepsilon} = O(\varphi(n))$ as $n \to \infty$, i.e., $\mathsf{p}_n^{\varepsilon} \le C(\varepsilon)\varphi(n)$ for some constant $C(\varepsilon) > 0$ for all $n > n_0(\varepsilon)$. By Lemma 1, for $n > n_0(\varepsilon)$, we obtain the estimate

$$\mathsf{P}_n^{2\varepsilon} \le \sum_{k=0}^{\infty} \mathsf{p}_{n_k}^{\varepsilon} \le C(\varepsilon) \sum_{k=0}^{\infty} \varphi(n_k).$$

Since $\varphi$ is monotone decreasing for $x > x_0$, and $r > 1$, it follows that, for all $n > [x_0]$,

$$\begin{aligned}\sum_{k=0}^{\infty}\varphi(n_k) &\le \sum_{k=0}^{\infty}\varphi(nr^k) = \varphi(n) + \sum_{k=1}^{\infty}\varphi(nr^k)\\ &\le \varphi(n) + \int_0^{\infty}\varphi(nr^x)\,dx = \varphi(n) + \frac{1}{\ln r}\int_n^{\infty}\frac{\varphi(y)}{y}\,dy\\ &\le \varphi(n) + \frac{M}{\ln r}\varphi(n) = \biggl(1+\frac{M}{\ln r}\biggr)\varphi(n).\end{aligned}$$

Here we have used the change of variable of integration $nr^x = y$ and the asymptotic property (1) of the function $\varphi$ written as the inequality

$$\int_n^{\infty}\frac{\varphi(x)}{x}\,dx \le M\varphi(n),$$

which holds for some constant $M > 0$ for all $n > n_1$, where $n_1$ can be taken as $[x_0]$. Indeed, if $n_1 > [x_0]$, then, for all $n > [x_0]$, using the monotonicity of $\varphi(x)$, we obtain

$$\begin{aligned}\int_n^{\infty}\frac{\varphi(x)}{x}\,dx &= \int_n^{n_1+1}\frac{\varphi(x)}{x}\,dx + \int_{n_1+1}^{\infty}\frac{\varphi(x)}{x}\,dx\\ &\le \varphi(n)\ln\biggl(\frac{n_1+1}{n}\biggr) + M\varphi(n_1+1) \le \varphi(n)\biggl(\ln\biggl(\frac{n_1+1}{[x_0]+1}\biggr) + M\biggr).\end{aligned}$$

For all $n > \max\{[x_0], n_0(\varepsilon)\}$, combining the resulting estimates, we obtain the inequality

$$\mathsf{P}_n^{2\varepsilon} \le C(\varepsilon)\biggl(1+\frac{M}{\ln r}\biggr)\varphi(n).$$

Since Lemma 1, also holds for continuous time, we obtain the proof in the case under consideration by replacing the quantities $\mathsf{P}_n^{\varepsilon}$ and $\mathsf{p}_n^{\varepsilon}$ by their continuous analogs $\overline{\mathsf{P}}_t^{\varepsilon}$ and $\overline{\mathsf{p}}_t^{\varepsilon}$. □

In applications, it often occurs that the probabilities of large deviations $\mathsf{p}_n^{\varepsilon}$ do not decrease to zero with the same rate $\varphi(n)$ for all positive $\varepsilon$, and this rate may depend on $\varepsilon$. Therefore, instead of one function $\varphi(x)$, we must consider a one-parameter family of functions $\{\varphi_\varepsilon(x)\}_{\varepsilon>0}$. Let us refine Theorem 1 in this respect (estimates for $\mathsf{P}_n^{\varepsilon}$ from $\mathsf{p}_n^{\varepsilon}$).

**Theorem 2.** *Suppose that $f \in L_\infty(\Omega)$, and, for some (fixed) $\varepsilon > 0$, the function $\varphi_\varepsilon(x)\colon \mathbb{R}^+ \mapsto \mathbb{R}^+$ is monotone decreasing (for $x > x_0(\varepsilon)$) to zero so that relation (1) holds. Then if there exist constants $C(\varepsilon) > 0$ and $n_0(\varepsilon) \in \mathbb{Z}^+$ such that, for all $n > n_0(\varepsilon)$, the inequality*

$$\mathsf{p}_n^{\varepsilon} \le C(\varepsilon)\varphi_\varepsilon(n),$$

*holds, then, for some constant $M(\varepsilon) > 0$ and for all $n > \max\{[x_0(\varepsilon)]; n_0(\varepsilon)\}$, the following inequality will be valid:*

$$\mathsf{P}_n^{2\varepsilon} \le C(\varepsilon)\biggl(1+\frac{M(\varepsilon)}{\ln r(\varepsilon)}\biggr)\varphi_\varepsilon(n). \tag{2}$$

*In the case of continuous time, a similar assertion is valid with time $n$ replaced by $t$ and the quantities $\mathsf{p}_n^{\varepsilon}$, $\mathsf{P}_n^{2\varepsilon}$ replaced by $\overline{\mathsf{p}}_t^{\varepsilon}$, $\overline{\mathsf{P}}_t^{2\varepsilon}$.*

**Proof of Theorem 2.** The asymptotic relation (1) for the function $\varphi_\varepsilon(x)$ can be rewritten as inequality

$$\int_n^{\infty}\frac{\varphi_\varepsilon(x)}{x}\,dx \le M(\varepsilon)\varphi_\varepsilon(n),$$

valid for some constant $M(\varepsilon) > 0$ for all $n > [x_0(\varepsilon)]$. Repeating the part of the proof of Theorem 1 from 2) to 1) word for word, we obtain estimate (2).

In the continuous case, estimate (2) becomes valid for $t > \max\{x_0(\varepsilon); t_0(\varepsilon)\}$, where $t_0(\varepsilon) \in \mathbb{R}^+$ is defined in the same way as $n_0(\varepsilon)$. □

In the following lemma, we determine how wide is the class of functions for which the asymptotic relation (1) holds.

**Lemma 2.** *Let $\psi_1(x)\colon \mathbb{R}^+ \mapsto \mathbb{R}^+$ be a nonincreasing (for $x > x_0$) function. Then, for each $\alpha > 0$, the function $\varphi_1(x) = \psi_1(x)/x^\alpha$ satisfies condition (1).*

*If the function $\psi_2(x)\colon \mathbb{R}^+ \mapsto \mathbb{R}^+$ satisfies condition (1), then, for any $\delta > 0$, the function $\varphi_2(x) = \psi_2(x^\delta)$ also satisfies (1).*

**Proof of Lemma 2.** Let $\psi_1(x)$ be a nonnegative nonincreasing (for $x > x_0$) function; then, for $n > [x_0]$, we obtain

$$\int_n^\infty \frac{\varphi_1(x)}{x}\,dx = \int_n^\infty \frac{\psi_1(x)}{x^{1+\alpha}}\,dx \le \psi_1(n)\int_n^\infty \frac{dx}{x^{1+\alpha}} = \frac{\psi_1(n)}{\alpha n^\alpha} = \frac{1}{\alpha}\,\varphi_1(n),$$

which proves the first assertion of the lemma.

If $\psi_2(x)$ satisfies (1), then, for any $\delta > 0$ for some constant $M > 0$, for all sufficiently large $n$, we obtain

$$\int_n^\infty \frac{\varphi_2(x)}{x}\,dx = \int_n^\infty \frac{\psi_2(x^\delta)}{x}\,dx = \frac{1}{\delta}\int_{n^\delta}^\infty \frac{\psi_2(y)}{y}\,dy \le \frac{1}{\delta}\,M\psi_2(n^\delta) = \frac{M}{\delta}\,\varphi_2(n),$$

which concludes the proof of the lemma. □

## 3. ESTIMATES OF THE RATE OF CONVERGENCE VIA ESTIMATES OF LARGE DEVIATIONS

As easily follows from Lemma 2, except for power-law functions $x^{-\alpha}$, $\alpha > 0$, the asymptotic relation (1) is satisfied by power-law functions with logarithmic multiplier,

$$x^{-\alpha}\ln^\beta x, \qquad \alpha > 0, \quad \beta \in \mathbb{R}, \quad x > 1,$$

exponential functions $e^{-\gamma x}$, $\gamma > 0$, and exponential functions with power-law multiplier $x^\alpha e^{-\gamma x}$, $\alpha \in \mathbb{R}$, $\gamma > 0$. In addition, from an exponential function, using the power-type change of variable, we can obtain stretched exponential decay functions $e^{-\gamma x^\delta}$, where $\gamma, \delta > 0$ (which also satisfy condition (1)).

Let us refine Theorem 2 for all the functions described above. Consider the incomplete gamma-function $\Gamma(a,x)$ and its particular case, the integral exponential $E_1(x)$

$$\Gamma(a,x) = \int_x^\infty e^{-t}t^{a-1}\,dt, \qquad x > 0, \quad a \in \mathbb{R},$$

$$E_1(x) = \Gamma(0,x) = \int_x^\infty \frac{e^{-t}}{t}\,dt, \qquad x > 0.$$

Also consider the functions

$$A(a,x) = \frac{e^x}{x^a}\,\Gamma(a,x), \qquad x > 0, \quad a \in \mathbb{R},$$

$$D(a,x) = \begin{cases} \dfrac{1}{a}, & x \le 0, \\[2mm] \dfrac{1}{a} + \dfrac{x}{a^2} + \dfrac{x(x-1)}{a^3} + \dots + \dfrac{x(x-1)\cdots(x-[x])}{a^{[x]+2}}, & x > 0, \end{cases} \qquad a > 0.$$

**Theorem 3.** *Let $f \in L_\infty(\Omega)$. Then, for each (fixed) $\varepsilon > 0$, the following assertions are valid.*

1. *If $\mathsf{p}_n^\varepsilon \le C(\varepsilon)n^{-\alpha(\varepsilon)}$ for some constants $C(\varepsilon) > 0$, $\alpha(\varepsilon) > 0$ for all $n > n_0(\varepsilon)$, $n_0(\varepsilon) \in \mathbb{Z}^+$, then, for all $n > n_0(\varepsilon)$,*

$$\mathsf{P}_n^{2\varepsilon} \le C(\varepsilon)\biggl(1 + \frac{1}{r(\varepsilon)^{\alpha(\varepsilon)} - 1}\biggr)n^{-\alpha(\varepsilon)} < C(\varepsilon)\biggl(1 + \frac{1}{\alpha(\varepsilon)\ln r(\varepsilon)}\biggr)n^{-\alpha(\varepsilon)}.$$

2. *If* $\mathsf{p}_n^\varepsilon \le C(\varepsilon)n^{-\alpha(\varepsilon)}\ln^{\beta(\varepsilon)} n$ *for some constants* $\alpha(\varepsilon) > 0$, $\beta(\varepsilon) \in \mathbb{R}$ *for all* $n > n_0(\varepsilon)$, $n_0(\varepsilon) \in \mathbb{N}$, *then, for all* $n > \max\{[e^{\beta(\varepsilon)/\alpha(\varepsilon)}]; n_0(\varepsilon); 2\}$,

$$\mathsf{P}_n^{2\varepsilon} \le C(\varepsilon)\biggl(1 + \frac{D(\alpha(\varepsilon), \beta(\varepsilon))}{\ln r(\varepsilon)}\biggr)n^{-\alpha(\varepsilon)}\ln^{\beta(\varepsilon)} n.$$

3. *If* $\mathsf{p}_n^\varepsilon \le C(\varepsilon)e^{-\gamma(\varepsilon)n^{\delta(\varepsilon)}}$ *for some constants* $\gamma(\varepsilon) > 0$, $\delta(\varepsilon) > 0$ *for all* $n > n_0(\varepsilon)$, $n_0(\varepsilon) \in \mathbb{Z}^+$, *then, for all* $n > n_0(\varepsilon)$,

$$\begin{aligned}\mathsf{P}_n^{2\varepsilon} &\le C(\varepsilon)\frac{1}{\delta(\varepsilon)}\biggl(1 + \frac{e^{\gamma(\varepsilon)}E_1(\gamma(\varepsilon))}{\ln r(\varepsilon)}\biggr)e^{-\gamma(\varepsilon)n^{\delta(\varepsilon)}} \\ &< C(\varepsilon)\frac{1}{\delta(\varepsilon)}\biggl(1 + \frac{\ln(1 + 1/\gamma(\varepsilon))}{\ln r(\varepsilon)}\biggr)e^{-\gamma(\varepsilon)n^{\delta(\varepsilon)}}.\end{aligned}$$

4. *If* $\mathsf{p}_n^\varepsilon \le C(\varepsilon)n^{\alpha(\varepsilon)}e^{-\gamma(\varepsilon)n}$ *for some constants* $\alpha(\varepsilon) \in \mathbb{R}$, $\gamma(\varepsilon) > 0$ *for all* $n > n_0(\varepsilon)$, $n_0(\varepsilon) \in \mathbb{Z}^+$, *then, for all* $n > \max\{[\alpha(\varepsilon)/\gamma(\varepsilon)]; n_0(\varepsilon)\}$,

$$\mathsf{P}_n^{2\varepsilon} \le C(\varepsilon)\biggl(1 + \frac{A(\alpha(\varepsilon), \gamma(\varepsilon))}{\ln r(\varepsilon)}\biggr)n^{\alpha(\varepsilon)}e^{-\gamma(\varepsilon)n}.$$

*In the case of continuous time, similar assertions are valid with time* $n$ *replaced by* $t$ *and the quantities* $\mathsf{p}_n^\varepsilon$, $\mathsf{P}_n^{2\varepsilon}$ *by* $\overline{\mathsf{p}}_t^\varepsilon$, $\overline{\mathsf{P}}_t^{2\varepsilon}$.

**Proof of Theorem 3.** It suffices to apply inequality (2) with a suitable constant $M(\varepsilon)$. To simplify the notation, we shall drop the parameter $\varepsilon$. Consider the function

$$F(n) = \frac{1}{\varphi(n)}\int_n^\infty \frac{\varphi(x)}{x}\,dx, \qquad n \ge 1,$$

and find its upper bound. Let us prove assertions 1, 3, and 4 of the theorem together. To do this, consider the function

$$\varphi(x) = x^{-\alpha}e^{-\gamma x}, \qquad \alpha \in \mathbb{R}, \quad \gamma > 0 \quad \text{and if } \gamma = 0, \text{ then } \alpha > 0.$$

Then it is easy to see that, for nonnegative $\alpha$ and $\gamma$, the function $\varphi$ decreases on the whole positive half-line (i.e., $x_0 = 0$). For $\alpha < 0$, the decrease begins from $x_0 = -\alpha/\gamma$. Further, it is easy to verify that, for $\gamma \ne 0$,

$$F(n) = \int_0^\infty e^{-\gamma n x}(x+1)^{-\alpha-1}\,dx$$

is a decreasing function; therefore, for all $n \ge 1$, the following inequality is valid:

$$F(n) \le F(1) = \frac{e^\gamma}{\gamma^{-\alpha}}\Gamma(-\alpha, \gamma) = A(-\alpha, \gamma),$$

whence we immediately obtain the estimate in assertion 3. If, further, $\alpha = 0$, then we obtain the estimate in assertion 2 with $\delta = 1$. For $\delta \ne 1$, the estimate follows from the argument for $\delta = 1$ and the power-type change of variable examined in Lemma 2. The second inequality in this estimate follows from the inequality (see, for example, [8, relation 5.1.20])

$$\frac{1}{2}\ln\biggl(1 + \frac{2}{x}\biggr) < e^x E_1(x) < \ln\biggl(1 + \frac{1}{x}\biggr), \qquad x > 0.$$

If we set $\gamma = 0$, then $\varphi(x) = x^{-\alpha}$, $\alpha > 0$, and $F(n) = 1/\alpha$. This yields the second inequality in assertion 1 of the theorem. The first inequality in this assertion is obtained by a direct calculation, using Lemma 1:

$$\mathsf{P}_n^{2\varepsilon} \le \sum_{k=0}^\infty \mathsf{p}_{n_k}^\varepsilon \le \sum_{k=0}^\infty C n_k^{-\alpha} \le C\sum_{k=0}^\infty (nr^k)^{-\alpha}$$

$$= Cn^{-\alpha}\sum_{k=0}^{\infty}(r^{-\alpha})^k = Cn^{-\alpha}\frac{1}{1-r^{-\alpha}} = C\biggl(1+\frac{1}{r^{\alpha}-1}\biggr)n^{-\alpha}.$$

We now pass to the proof of the second assertion of the theorem. Consider the function

$$\varphi(x) = x^{-\alpha}\ln^{\beta}x, \qquad \alpha>0, \quad \beta<0.$$

Obviously, for all $x\in[0,1]$, this function is not defined; therefore, to apply results of Theorem 2, let us extend it on this closed interval by zero. Then, for all $x>1$ ($x_0=1$), this function is positive and monotone decreasing to zero. The function $F(n)$ corresponding to it is monotone increasing; therefore, for all $n>1$,

$$F(n) = \int_1^{\infty}\frac{dx}{x^{\alpha+1}(1+\ln x/\ln n)^{-\beta}} < F(\infty) = \lim_{n\to\infty}F(n) = \int_1^{\infty}\frac{dx}{x^{\alpha+1}} = \frac{1}{\alpha},$$

whence we obtain $M=1/\alpha=D(\alpha,\beta)$. Now let $\varphi(x)=x^{-\alpha}\ln^{\beta}x$, $\alpha>0$, $\beta>0$ for $x>1$, and let $\varphi(x)=0$ for $x\in[0,1]$. It is easy to verify that this function decreases for $x>e^{\beta/\alpha}$ ($x_0=e^{\beta/\alpha}$). Denote by $m$ the integer part of the number $\beta$, i.e., $m=[\beta]$. For fixed $n$ and $\alpha$, we set

$$I(\beta) = \int_n^{\infty}\frac{\varphi(x)}{x}\,dx = \int_n^{\infty}\frac{\ln^{\beta}x}{x^{\alpha+1}}\,dx.$$

Integrating by parts, we obtain the formula for decreasing the value of $\beta$:

$$I(\beta) = \frac{\ln^{\beta}n}{\alpha n^{\alpha}} + \frac{\beta}{\alpha}I(\beta-1).$$

Applying it several times until the exponent becomes negative, in the case $\beta\neq m$, we obtain

$$\begin{aligned}I(\beta) = {}&\frac{\ln^{\beta}n}{\alpha n^{\alpha}} + \frac{\beta\ln^{\beta-1}n}{\alpha^2 n^{\alpha}} + \frac{\beta(\beta-1)\ln^{\beta-2}n}{\alpha^3 n^{\alpha}} + \cdots \\ &+ \frac{\beta(\beta-1)\cdots(\beta-m+1)\ln^{\beta-m}n}{\alpha^{m+1}n^{\alpha}} + \frac{\beta(\beta-1)\cdots(\beta-m)}{\alpha^{m+1}}I(\beta-m-1).\end{aligned}$$

In the case of an integer $\beta$, i.e., $\beta=m$, the last summand will be absent on the right-hand side of this equality. In view of the inequality $\beta-m-1<0$, using the result already proved above, we obtain

$$I(\beta-m-1) = \varphi(n)F(n) \le \frac{1}{\alpha}\varphi(n) = \frac{\ln^{\beta-m-1}n}{\alpha n^{\alpha}}.$$

Using this estimate and the obvious inequalities

$$\ln^{\beta-j}x < \ln^{\beta}x \qquad \text{for all} \quad x>e, \quad j>0,$$

we obtain the estimate

$$I(\beta) \le \biggl(\frac{1}{\alpha} + \frac{\beta}{\alpha^2} + \frac{\beta(\beta-1)}{\alpha^3} + \cdots + \frac{\beta(\beta-1)\dots(\beta-m)}{\alpha^{m+2}}\biggr)\frac{\ln^{\beta}n}{n^{\alpha}},$$

which yields the constant $M=D(\alpha,\beta)$ and the required (in assertion 2) estimate valid for all $n>\max\{[e^{\beta/\alpha}];n_0;2\}$. Since, for continuous time, Theorem 2 and inequality (2) hold, it follows that, for this case, the proof is completely similar to the one given above; we must only replace time $n$ by $t$. □

**Remark 2.** If ergodic means are generated not by the Koopman operator $Uf(\omega)=f(T\omega)$, $f\in L_1(\Omega)$ (the semigroup of operators $U_tf(\omega)=f(T^t\omega)$, $t\ge0$, in the continuous case), but by the Dunford–Schwartz operator in $L_1(\Omega)$ (the semigroup of such operators in the continuous case), i.e., by a linear contraction operator in both $L_1(\Omega)$, and $L_\infty(\Omega)$ (a more general case, because the Koopman operator in our case is an isometry in $L_1(\Omega)$ and a contraction in $L_\infty(\Omega)$), then, obviously, the analogs of the assertions of all three theorems proved in this section will be valid for this more general case as well.

## ACKNOWLEDGMENTS

This work was supported by the program "Leading Scientific Schools" (grant no. NSh-5998.2012.1).

## REFERENCES

1. I. Melbourne and M. Nicol, "Large deviations for nonuniformly hyperbolic systems," Trans. Amer. Math. Soc. **360** (12), 6661–6676 (2008).
2. I. Melbourne, "Large and moderate deviations for slowly mixing dynamical systems," Proc. Amer. Math. Soc. **137** (5), 1735–1741 (2009).
3. S. Waddington, "Large deviation asymptotics for Anosov flows," Ann. Inst. H. Poincaré Anal. Non Linéaire **13** (4), 445–484 (1996).
4. V. Araújo and A. I. Bufetov, "A large deviations bound for the Teichmüller flow on the moduli space of abelian differentials," Ergodic Theory Dynam. Systems **31** (4), 1043–1071 (2011).
5. A. G. Kachurovskii and V. V. Sedalishchev, "Constants in estimates for the rates of convergence in von Neumann's and Birkhoff's ergodic theorems," Mat. Sb. **202** (8), 21–40 (2011) [Sb. Math. **202** (8), 1105–1125 (2011)].
6. A. G. Kachurovskii and V. V. Sedalishchev, "On the constants in the estimates of the rate of convergence in the Birkhoff ergodic theorem," Mat. Zametki **91** (4), 624–628 (2012) [Math. Notes **91** (4), 582–587 (2012)].
7. A. G. Kachurovskii, "The rate of convergence in ergodic theorems," Uspekhi Mat. Nauk **51** (4), 73–124 (1996) [Russian Math. Surveys **51** (4), 653–703 (1996)].
8. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Ed. by M. Abramowitz and I. Stegun (National Bureau of Standards, Washington, D. C., 1964; Nauka, Moscow, 1979).