Probab. Theory Relat. Fields (2007) 138:195–234
DOI 10.1007/s00440-006-0021-6

# On almost-sure versions of classical limit theorems for dynamical systems

**J.-R. Chazottes · S. Gouëzel**

Received: 16 January 2006 / Revised: 1 June 2006 /
Published online: 1 August 2006


**Abstract** The purpose of this article is to support the idea that "whenever we can prove a limit theorem in the classical sense for a dynamical system, we can prove a suitable almost-sure version based on an empirical measure with log-average". We follow three different approaches: martingale methods, spectral methods and induction arguments. Our results apply, among others, to Axiom A maps or flows, to systems inducing a Gibbs–Markov map, and to the stadium billiard.

**Keywords** Almost-sure central limit theorem · Almost-sure convergence to stable laws · Gibbs–Markov map · Inducing · Suspension flow · Martingales · hyperbolic flow · Stadium billiard

## Contents



J.-R. Chazottes (✉)
CPhT, CNRS-Ecole Polytechnique, 91128 Palaiseau Cedex, France
e-mail: jeanrene@cpht.polytechnique.fr

S. Gouëzel
IRMAR, Université de Rennes 1, Campus de Beaulieu, 35042 Rennes Cedex, France
e-mail: sebastien.gouezel@univ-rennes1.fr

## 1 Introduction

There has been recently a lively interest in probability theory concerning almost-sure versions of classical limit theorems. The prototype of such a theorem is the almost-sure central limit theorem: if $Z_n$ is an i.i.d. $L^2$ sequence with $\mathbb{E}(Z_i) = 0$ and $\mathbb{E}(Z_i^2) = 1$, then, almost surely,

$$\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \delta_{\sum_{j=0}^{k-1} Z_j/\sqrt{k}} \xrightarrow{\text{law}} \mathscr{N}(0,1) \tag{1}$$

where "$\xrightarrow{\text{law}}$" means weak convergence of probability measures on $\mathbb{R}$. Here and henceforth, $\delta_x$ is the Dirac mass at $x$. This result should be compared to the classical central limit theorem, which can be stated as follows:

$$\mathbb{E}\left[ \mathbb{1}_{\{\sum_{j=0}^{n-1} Z_j/\sqrt{n} \leqslant t\}} \right] \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{e}^{-x^2/2}\, \mathrm{d}x \tag{2}$$

for any $t \in \mathbb{R}$. To better compare these theorems, it is worth noticing that (1) implies that *almost surely*

$$\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{1}_{\{\sum_{j=0}^{k-1} Z_j/\sqrt{k} \leqslant t\}} \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{e}^{-x^2/2}\, \mathrm{d}x \tag{3}$$

for any $t \in \mathbb{R}$. So, instead of taking the expected value, we take a logarithmic average and obtain an almost-sure convergence.

In fact, whenever there is independence and a classical limit theorem, the corresponding almost-sure limit theorem also holds (under minor technical conditions), see [9] and references therein. The situation is more complicated for weakly dependent sequences, see [31] and references therein.

For dynamical systems $(X, T, m)$ given by the iteration of a map $T : X \circlearrowleft$ which preserves the probability measure $m$, we take $Z_j = f \circ T^j$, where $f : X \to \mathbb{R}$ is an observable. Here, the randomness only comes from the choice of the initial condition $x$ according to the invariant measure of the system. The sequence $Z_j$ is identically distributed (in fact stationary), but there is no independence in general. Nevertheless, it is well-known that many dynamical systems display a complicated behavior which can be adequately analyzed by probabilistic methods.

For some classes of systems, it is possible to use probabilistic techniques for weakly dependent sequences, and prove an almost sure invariance principle. That is, there exist $\varepsilon > 0$ and a Brownian motion $W$ (on a possibly extended space) such that, almost surely,

$$\sum_{j=0}^{n-1} f \circ T^j(x) = W(n)(x) + o(n^{1/2-\varepsilon}) \quad \text{when } n \to \infty. \tag{4}$$

This directly implies that the Birkhoff sums of $f$ satisfy an almost sure central limit theorem, by [28]. See e.g. [13,14,29] for examples of dynamical systems satisfying the almost sure invariance principle – these include Anosov maps as well as partially or non-uniformly hyperbolic transformations.

The goal of this article is to support the idea that "whenever we can prove a limit theorem in the classical sense for a dynamical system, we can prove a suitable almost-sure version". More precisely, we will investigate three methods that are used to prove limit theorems in dynamical systems: spectral methods, martingale methods, and induction arguments. We will show that whenever these methods apply, the corresponding limit theorem admits a suitable almost-sure version. Typically our statements will look like:

$$\frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \delta_{S_k f/B_k} \xrightarrow{\text{law}} \mathscr{W} \quad \text{almost-surely} \tag{5}$$

where $f : X \to \mathbb{R}$ is a "regular" observable, $B_k$ is a suitable normalizing sequence, $\mathscr{W}$ a suitable law, and $S_k f := f + f \circ T + \cdots + f \circ T^{k-1}$.

Let us give some more details about the methods for proving limit theorems we mentioned above:

- **Spectral methods:** If $T : X \to X$ is a probability preserving map, the corresponding *transfer operator* is defined on $L^2$ as the adjoint of the composition by $T$. Under suitable assumptions on the map $T$, it acts on spaces of regular functions, and has a spectral gap. This result, which implies in particular exponential decay of correlations, is a very useful tool to study limit theorems for $T$. In a specific setting, the so-called *Gibbs–Markov maps*, this good spectral behavior was used by Aaronson and Denker in [1,2] to prove that the suitably renormalized Birkhoff sums of a good observable converge to a Gaussian or stable law (see Theorem 2.8 for a precise statement of their results). Under the same assumptions, we will prove an almost sure limit theorem (Theorem 2.10). It will be derived from a more general theorem stated in terms of continuous perturbations of transfer operators, which applies in a large variety of settings (Theorem 2.11).

- **Martingale methods:** Let again $T$ be a probability preserving map on a space $X$. If $f : X \to \mathbb{R}$ is a function, it is sometimes possible to write it as $f = g - g \circ T + h$, where $g$ is a measurable function and the sequence $h \circ T^n$ is a reverse martingale difference for some filtration. The central limit theorem for reverse martingale differences then implies that the Birkhoff sums of $h$ satisfy a central limit theorem. This in turn yields the same conclusion for $f$. We will prove that a sequence of reverse martingale differences also satisfies an almost sure central limit theorem, by mimicking the proof in [26] for the direct martingale differences. As above, this gives an almost sure limit theorem in the dynamical systems setting, given in Theorem 2.16.
- **Induction methods:** Let $T$ be a probability preserving map on a space $X$, and let $Y$ be a positive measure subset of $X$. Let $T_Y$ be the induced map on $T$, and $\varphi$ the first return time. If the Birkhoff sums of a function $f$, for the transformation $T_Y$, satisfy a limit theorem, then it is well known (see e.g. [3, 30,34]) that, *under suitable additional assumptions*, the function $f$ also satisfies a limit theorem for the initial map $T$. In [30], the additional assumptions are formulated in terms of the return time function $\varphi$, which should essentially satisfy a central limit theorem. Our first goal when we started to write this paper was to extend this kind of result to almost sure limit theorems. We were surprised to realize that this extension was indeed possible, under *weaker* assumptions. Indeed, there is no need to assume anything on the return time function $\varphi$ (see Theorem 2.14).
  We also tried to eliminate the conditions on $\varphi$ in Melbourne and Török's classical limit theorem, and were only partially successful: this is possible under additional assumptions on the function $f$, which amount to a tightness condition for the maxima of the Birkhoff sums for $T_Y$ (see Definition 2.2). This condition can be checked in several practical cases, by a martingale argument. This yields new limit theorems which could not be proved by the previous variations around [30], see e.g. Theorem 2.19 in which there is no assumption on the return time $\varphi$.

We notice that almost-sure central limit theorems have been established for certain dynamical systems in [11,12] as a consequence of some concentration inequalities. In fact, a strengthening of the almost-sure central limit theorem is obtained in these papers. Moreover, [25] proves that, for *any* probability preserving dynamical system, there exists a function $f$ which satisfies the almost sure central limit theorem (but usually, this function is quite wild).

In Sect. 2, we give all the precise statements of the theorems, and the remaining sections are devoted to their proofs.

## 2 Statement of main results

### 2.1 Different notions of convergence

In this paragraph, we modify the classical notion of convergence of random variables in two different ways, by putting an additional condition which will prove very useful to induce limit theorems from a subset of the space to the whole space (see Theorem 2.12), or by studying almost sure convergence. We will see later that these new notions of convergence are satisfied in several cases by dynamical systems.

A continuous function $L : \mathbb{R}_+^* \to \mathbb{R}_+^*$ is *slowly varying* if, for all $\lambda > 0$, $L(\lambda x)/L(x) \to 1$ when $x \to \infty$. This implies that $L(x) = o(x^\varepsilon)$ for all $\varepsilon > 0$, as well as $1/L(x) = o(x^\varepsilon)$. Basic examples of functions with slow variation are constant functions and powers of the logarithm function.

A slowly varying function $L$ is said to be *normalized* if $L$ is $C^1$ and $L'(x) = o(L(x)/x)$. Every slowly varying function is asymptotically equivalent to a normalized slowly varying function, see [10; Theorem 1.3.3]. In particular, if one is only interested in the asymptotic behavior of slowly varying functions, one can without loss of generality restrict oneself to normalized slowly varying functions.

**Definition 2.1** *A* renormalization function *is a function $B : \mathbb{R}_+^* \to \mathbb{R}_+^*$ of the form $B(x) = x^d L(x)$ where $d > 0$ and $L$ is a normalized slowly varying function. The corresponding* renormalizing sequence *is $B_n := B(n)$.*

**Definition 2.2** *Let $S_n$ be a sequence of random variables on a probability space, and let $B_n$ be a renormalizing sequence. We say that $(S_n/B_n, B_n)$ converges with tight maxima to a random variable $\mathscr{W}$, if $S_n/B_n$ converges in law to $\mathscr{W}$, and the sequence $M_n = (\max_{1\leqslant k\leqslant n} |S_k|)/B_n$ is tight, i.e.*

$$\forall \varepsilon > 0, \exists\, c > 0 \text{ s.t. } \forall n \geqslant 1,\ \mathbb{P}\left\{\max_{1\leqslant k\leqslant n} |S_k|/B_n > c\right\} \leqslant \varepsilon\,. \tag{6}$$

Notice that this property is not a property of the sequence $S_n/B_n$ only, the renormalizing sequence $B_n$ plays a role in the definition of $M_n$. However, abusing notations, we will usually simply say that $S_n/B_n$ converges with tight maxima to $\mathscr{W}$.

*Example 2.3* Let $Z_0, Z_1 \dots$ be a sequence of reverse martingale differences. Let $S_n = \sum_{k=0}^{n-1} Z_k$. Assume that, for some renormalizing sequence $B_n, S_n/B_n$ converges in law to a random variable $\mathscr{W}$, and that $S_n/B_n$ is bounded in $L^1$. Then $S_n/B_n$ also converges with tight maxima to $\mathscr{W}$.

*Proof* The maximal inequality for reverse martingales shows that, for all $\alpha > 0$ and all $n \in \mathbb{N}$,

$$\mathbb{P}\left\{\max_{1\leqslant k\leqslant n} |S_k| \geqslant \alpha\right\} \leqslant \frac{C}{\alpha}\,\mathbb{E}(|S_n|) \tag{7}$$

where $C$ is a universal constant. In particular, for all $c > 0$,

$$\mathbb{P}\left\{\max_{1\leqslant k\leqslant n} |S_k|/B_n \geqslant c\right\} \leqslant \frac{C}{cB_n}\,\mathbb{E}(|S_n|) \tag{8}$$

which is bounded by $C'/c$ since $S_n/B_n$ is bounded in $L^1$. □

**Definition 2.4** *Let $S_n$ be a sequence of random variables on a probability space, and let $B_n$ be a renormalizing sequence. We say that $S_n/B_n$ satisfies an almost sure central limit theorem towards a random variable $\mathscr{W}$ if, for almost all $\omega$,*

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}\delta_{S_k(\omega)/B_k}\xrightarrow{\text{law}}\mathscr{W} \tag{9}$$

*where $\delta_x$ is the Dirac mass at $x$, and the convergence is the weak convergence for probability measures on $\mathbb{R}$.*

Contrary to Definition 2.2, this is a property of the sequence $S_n/B_n$ only.

## 2.2 Spectral arguments: Gibbs–Markov maps

### *2.2.1 Almost-sure limit theorems in the i.i.d. case*

Let us first recall the precise statements of almost-sure limit theorems for i.i.d. sequences of random variables in the domain of attraction of a Gaussian or stable law.

A function $f$, defined on a probability space $(\Omega,\mathscr{B},m)$, is said to *belong to a domain of attraction* if it satisfies one the following three conditions:

I. It belongs to $L^2(\Omega)$.
II. One has $\int \mathbb{1}_{\{|f|>x\}}\,\mathrm{d}m \sim x^{-2}\ell(x)$, for some function $\ell$ such that $L(x) := 2\int_1^x \frac{\ell(u)}{u}\,\mathrm{d}u$ is of slow variation and unbounded.
III. There exists $p \in (1,2)$ such that $\int \mathbb{1}_{\{f>x\}}\,\mathrm{d}m = (c_1+o(1))x^{-p}L(x)$ and $\int \mathbb{1}_{\{f<-x\}}\,\mathrm{d}m = (c_2+o(1))x^{-p}L(x)$, where $c_1, c_2$ are nonnegative real numbers such that $c_1+c_2>0$, and $L$ is of slow variation.

It is convenient to say that in conditions I and II we have $p=2$, and that $L(x)=1$ in condition I.

Let us briefly comment on these conditions. The second one is equivalent to the fact that $\widetilde{L}(x) = \int f^2\mathbb{1}_{\{|f|\leqslant x\}}\,\mathrm{d}m$ is of slow variation and unbounded. Moreover, the functions $L$ and $\widetilde{L}$ then are equivalent at $+\infty$. In that case, the function $f$ belongs to $L^q$ for all $q<2$, but not to $L^2$. Also $\ell(x) = o(L(x))$.

In condition III, the function $f$ belongs to $L^q$ for all $q<p$. It may or may not belong to $L^p$, according to the behavior of the function $L$. It never belongs to $L^q$ for $q>p$.

Note in particular that the three conditions are mutually exclusive.

The above definition of domain of attraction is motivated by the following well-known, classical result in Probability (see e.g. [21]):

**Theorem 2.5** *Let $Z$ be a random variable belonging to a domain of attraction. Let $Z_0, Z_1, \dots$ be a sequence of independent, identically distributed, random variables with the same law as $Z$. In all cases, we set $A_n = n\mathbb{E}(Z)$ and*

1. *If condition I holds, we set $B_n=\sqrt{n}$ and $\mathscr{W} = \mathscr{N}(0,\mathbb{E}(Z^2)-\mathbb{E}(Z)^2)$.*

2. *If condition II holds, we let* $B_n$ *be a renormalizing sequence with* $nL(B_n) \sim B_n^2$, *and* $\mathscr{W} = \mathscr{N}(0,1)$.
3. *If condition III holds, we let* $B_n$ *be a renormalizing sequence such that* $nL(B_n) \sim B_n^p$. *Define* $c = (c_1 + c_2)\Gamma(1-p)\cos\left(\frac{p\pi}{2}\right)$ *and* $\beta = \frac{c_1 - c_2}{c_1 + c_2}$. *Let* $\mathscr{W}$ *be the law with characteristic function*

$$\mathbb{E}(e^{it\mathscr{W}}) = e^{-c|t|^p\left(1 - i\beta\,\mathrm{sgn}(t)\tan\left(\frac{p\pi}{2}\right)\right)}. \tag{10}$$

*Then*

$$\frac{\sum_{i=0}^{n-1} Z_i - A_n}{B_n} \xrightarrow{\text{law}} \mathscr{W}. \tag{11}$$

The conditions put on the distribution of $Z$ are almost necessary and sufficient to get a convergence in law of that type, we only restricted the range of $p$'s, which could also be taken in the interval $(0,1]$.

Notice that it is possible to construct the renormalizing sequence $B_n$ by taking $B_n = n^{1/p}\overline{L}(n)$, where $\overline{L}$ is a normalized slowly varying function built up from $L$, or more precisely from its *de Bruijn conjugate* (see [10]).

As a matter of fact, random variables $S_n = \sum_{i=0}^{n-1} Z_i - A_n$ as in the statement of the preceding theorem not only converge in law when properly rescaled, but they also converge with tight maxima (this is a consequence of Example 2.3). Moreover, the following theorem holds (see e.g. [9] for a proof).

**Theorem 2.6** *Under the same hypotheses and with the same notations as in Theorem 2.5,* $(\sum_{i=0}^{n-1} Z_i - A_n)/B_n$ *satisfies an almost sure limit theorem towards* $\mathscr{W}$.

### *2.2.2 Almost-sure limit theorems for Gibbs–Markov maps*

In this paragraph we give the analog of Theorem 2.6 for Gibbs–Markov maps, which are defined as follows.

**Definition 2.7** *Let* $T : X \circlearrowleft$ *be a nonsingular map on a probability, metric space* $(X, \mathscr{B}, m, d)$ *with bounded diameter, preserving the probability measure* $m$. *This map is said to be "*Gibbs–Markov*" if there exists a countable (measurable) partition* $\alpha$ *of* $X$ *such that:*

1. *For all* $a \in \alpha$, $T$ *is injective on* $a$ *and* $T(a)$ *is a union of elements of* $\alpha$.
2. *There exists* $\lambda > 1$ *such that, for all* $a \in \alpha$, *for all points* $x, y \in a$, $d(Tx, Ty) \geqslant \lambda d(x, y)$.
3. *Let* Jac *the inverse of the Jacobian of* $T$. *There exists* $C > 0$ *such that, for all* $a \in \alpha$, *for all points* $x, y \in a$, $\left|1 - \frac{\mathrm{Jac}(x)}{\mathrm{Jac}(y)}\right| \leqslant Cd(Tx, Ty)$.
4. *The map* $T$ *has the "big image property":* $\inf_{a\in\alpha} m(Ta) > 0$.

These properties say that $T$ is Markovian, uniformly expanding and with bounded distortion. In some sense, such maps have the strongest possible chaotic behavior, and are the first candidates when one wants to extend a probabilistic limit theorem to dynamical systems.

Let us define the separation time, $s(x,y)$, of two points $x,y \in X$ as the number of iterations of $T$ necessary for the orbit of $x$ and $y$ to fall into distinct atoms of the partition $\alpha$. For $\tau < 1$, define a new distance $d_\tau$ on $X$ by setting $d_\tau(x,y) = \tau^{s(x,y)}$. If $\tau$ is sufficiently close to 1, the map $T$ is still Gibbs–Markov for the distance $d_\tau$.

Let $f : X \to \mathbb{R}$ a function. For $X' \subset X$, we let

$$Df(X') = \sup\left\{ \frac{|f(x) - f(y)|}{d(x,y)} \,:\, x,y \in X', x \neq y \right\}. \tag{12}$$

This is the best Lipschitz constant of $f$ on $X'$.

We now state a theorem asserting that the convergence results of Theorem 2.5 extend from the i.i.d. case to the case of Gibbs–Markov maps. This result is proved in [1,2] and [18].

**Theorem 2.8** *Let $T : X \circlearrowleft$ be a Gibbs–Markov map for a partition $\alpha$, preserving the ergodic probability measure $m$. Consider $f : X \to \mathbb{R}$ such that $\sum_{a\in\alpha} m(a)Df(a) < \infty$ and such that the distribution of $f$ belongs to a domain of attraction as above. Assume also $\int f \,\mathrm{d}m = 0$. Then*

$$\frac{S_n f}{B_n} \xrightarrow{\text{law}} \mathscr{W} \tag{13}$$

*where $B_n = \sqrt{n}$ and $\mathscr{W} = \mathscr{N}(0,\sigma^2)$ for some $\sigma^2 \geqslant 0$ if $f \in L^2$; Otherwise $B_n$ and $\mathscr{W}$ are as in the i.i.d. case.*

We use the classical notation $S_n f := f + f \circ T + \cdots + f \circ T^{n-1}$.

*Remark 2.9* When $f \in L^2$, the value of $\sigma^2$ is not always $\int f^2 \,\mathrm{d}m$, due to the lack of independence. It is in fact equal to $\int f^2 \,\mathrm{d}m + 2\sum_{k=1}^{\infty} \int f \cdot f \circ T^k \,\mathrm{d}m$ (and this series is converging). On the other hand, when $f \notin L^2$, the sequence $f \circ T^k$ behaves really as if it were independent.

In this setting, we obtain the following result concerning almost-sure limit theorems.

**Theorem 2.10** *With the assumptions and notations of Theorem 2.8, $S_n f/B_n$ converges with tight maxima to $\mathscr{W}$. Moreover, it satisfies an almost sure limit theorem towards $\mathscr{W}$.*

### *2.2.3 A more general spectral result*

Theorem 2.10 will be derived from a more general spectral theorem which applies also to different settings. The spirit of this paragraph is close to the ideas of [23], with weaker continuity assumptions.

Let $T$ be a nonsingular map on the probability space $(X,m)$ (the probability measure $m$ is not assumed to be invariant), and let $f : X \to \mathbb{R}$ be measurable.

Let $\mathscr{G}, \mathscr{H}$ be two complex Banach spaces and let $\mathrm{i} : \mathscr{G} \to \mathscr{H}$ be a continuous linear map such that the image of the unit ball of $\mathscr{G}$ is relatively compact in $\mathscr{H}$. Assume that two elements $\alpha_0 \in \mathscr{G}$ and $\ell_0 \in \mathscr{G}'$ (the dual of $\mathscr{G}$) are given. Finally, consider some $0 < \varepsilon_0 < 1$, and assume that operators $\mathscr{L}_t : \mathscr{G} \to \mathscr{G}$ are given, for $|t| \leqslant \varepsilon_0$.

We assume the following properties:

1. For all $t, t' \in [-\varepsilon_0, \varepsilon_0]$ and all $n, p \in \mathbb{N}$,

$$\int \mathrm{e}^{itS_n f \circ T^p} \mathrm{e}^{it' S_p f} \, \mathrm{d}m = \langle \ell_0, \mathscr{L}_t^n \mathscr{L}_{t'}^p \alpha_0 \rangle. \tag{14}$$

2. There exist constants $C > 0, \eta < 1$ and $M \geqslant 1$ such that, for all $u \in \mathscr{G}$, for all $n \in \mathbb{N}$, for all $t \in [-\varepsilon_0, \varepsilon_0]$,

$$\left\| \mathscr{L}_t^n u \right\|_{\mathscr{G}} \leqslant C\eta^n \left\| u \right\|_{\mathscr{G}} + CM^n \left\| \mathrm{i}(u) \right\|_{\mathscr{H}}. \tag{15}$$

and

$$\left\| \mathrm{i}(\mathscr{L}_t^n u) \right\|_{\mathscr{H}} \leqslant CM^n \left\| \mathrm{i}(u) \right\|_{\mathscr{H}}. \tag{16}$$

3. The eigenvalues of modulus $\geqslant 1$ of the operator $\mathscr{L}_0$ are simple. Moreover, $\mathscr{L}_0' \ell_0 = \ell_0$.
4. There exists $\beta_0 > 0$ such that, for all $t \in [-\varepsilon_0, \varepsilon_0]$,

$$\| \mathrm{i} \circ (\mathscr{L}_t - \mathscr{L}_0) \|_{\mathscr{G} \to \mathscr{H}} \leqslant C|t|^{\beta_0}. \tag{17}$$

It is often possible to take for $\mathscr{G}$ a space of functions on $X$. The operator $\mathscr{L}_0$ is the transfer operator, $\alpha_0$ is the function 1 and $\ell_0$ is the integration against the measure $m$. The perturbed operator $\mathscr{L}_t$ is then usually given by $\mathscr{L}_t(u) = \mathscr{L}_0(\mathrm{e}^{itf} u)$ *if this can be defined*. The first assumption is then a formal consequence of the definition. To do this, one needs to be able to multiply an element of $\mathscr{G}$ by the function $\mathrm{e}^{itf}$, and still get an element of $\mathscr{G}$. This is not always the case. For example, when $T$ is a Gibbs–Markov map and the function $f$ is integrable and satisfies $\sum_{a \in \alpha} m(a) Df(a) < \infty$ (where $\alpha$ is the Markov partition of $T$), then it is possible to define $\mathscr{L}_t$ acting on the space $\mathscr{G}$ of locally Hölder functions, but not as naively as before: in general, if $u \in \mathscr{G}$, then $e^{itf} u \notin \mathscr{G}$. Nevertheless, the operator $\mathscr{L}_0$ is regularizing, and sends back $\mathrm{e}^{itf} u$ in $\mathscr{G}$, therefore $\mathscr{L}_t$ is well defined and satisfies the first assumption.

The more general setting given above is useful to treat more general dynamical systems where the convenient spaces to act on are not spaces of functions any more, such as in the hyperbolic setting (see [6,20]).

Notice that the second assumption is a uniform Lasota–Yorke inequality. By Hennion's Theorem [23], it ensures that $\mathscr{L}_t$ has a finite number of eigenvalues of modulus $\geqslant \rho$ for any $\rho > \eta$, and that these eigenvalues have finite multiplicity. The third assumption gives a more specific spectral description for $\mathscr{L}_0$.

The fourth assumption is a *weak* continuity assumption. It does not imply that $\|\mathscr{L}_t - \mathscr{L}_0\|_{\mathscr{G} \to \mathscr{G}} \to 0$ when $t \to 0$ (this would be a too strong assumption, which would not be satisfied in many interesting cases, see e.g. the case of the stadium billiard in Paragraph 2.5.3). However, together with the uniform Lasota–Yorke inequality, it is sufficient to get continuity properties for the spectrum of $\mathscr{L}_t$ by [7,24].

**Theorem 2.11** *Under the assumptions 1–4, let $B_n$ be a renormalizing sequence such that $S_n f/B_n$ converges in distribution to a random variable $\mathscr{W}$. Then $S_n f/B_n$ satisfies an almost sure limit theorem towards $\mathscr{W}$.*

## 2.3 Induction arguments

Melbourne and Török [30] have shown that under mild assumptions the central limit theorem for a map implies the central limit theorem for suspension flows over that map. In fact this holds for inducing: if an induced map satisfies a limit theorem, so does the map on the whole space, provided the return time is nice enough.

We will show that it is possible to replace this condition on the return time by a condition on tight maxima. To state this result, we need a few notations. Let $(X, \mathscr{B}, m, T)$ be an ergodic dynamical system, and let $Y \subset X$ be a subset with positive measure. For $y \in Y$, let

$$\varphi(y) = \inf\{n > 0 \,:\, T^n y \in Y\}. \tag{18}$$

This is the first return time of $y$ to $Y$. For a function $f : X \to \mathbb{R}$, define

$$f_Y(y) = \sum_{k=0}^{\varphi(y)-1} f(T^k y) \quad \text{if } y \in Y, \qquad f_Y(y) = 0 \quad \text{if } y \notin Y. \tag{19}$$

Denote by $T_Y : Y \to Y$ the induced map, that is, $T_Y y = T^{\varphi(y)} y$ for every $y \in Y$ such that $\varphi(y) < \infty$. Let $S_k^Y$ stand for the Birkhoff sums for $T_Y$. Finally set $m_Y = m(Y)^{-1} m_{|Y}$. The map $T_Y$ is defined almost everywhere on $Y$, and preserves the probability measure $m_Y$.

**Theorem 2.12** *Let $T$ be an ergodic endomorphism of a probability space $(X, \mathscr{B}, m)$. Let $Y \subset X$ be a set with positive measure, and let $f : X \to \mathbb{R}$ be an integrable function. Let $B : \mathbb{R}_+^* \to \mathbb{R}_+^*$ be a renormalization function. Assume that $S_n^Y f_Y / B(n/m(Y))$ converges with tight maxima to a random variable $\mathscr{W}$, for the measure $m_Y$. Then $S_n f/B(n)$ converges in law to $\mathscr{W}$, for the measure $m$.*

Under these assumptions, it is interesting to know when the convergence of $S_n f/B_n$ to $\mathscr{W}$ still has tight maxima, since it would make it possible to induce again and again. This is the case under a quite mild condition:

**Proposition 2.13** *Under the assumptions of Theorem 2.12, assume additionally that the function M defined on Y by* $M(y) = \max_{1 \leqslant k \leqslant \varphi(y)} |S_k f(y)|$ *satisfies:*

$$\sup_{n \in \mathbb{N}} n\, m\{y \in Y \,:\, M(y) \geqslant cB(n)\} \to 0 \quad \text{when } c \to +\infty. \tag{20}$$

*Then* $S_n f/B(n)$ *converges with tight maxima to* $\mathscr{W}$.

Note that the function $M$ is bounded by $|f|_Y$. For instance, if $B(n) = \sqrt{n}$ and $|f|_Y \in L^2(Y)$, then

$$nm\{M \geqslant c\sqrt{n}\} \leqslant nm\{|f|_Y \geqslant c\sqrt{n}\} \leqslant n\mathbb{E}(|f|_Y^2)/(nc^2) = O(1/c^2) \tag{21}$$

which shows that the assumption (20) is satisfied. More generally, if the tails of $f_Y$ and $|f|_Y$ are comparable, then this assumption is often satisfied.

For the almost-sure version of those limit theorems, we will need *weaker* assumptions, since no control on the maxima will be required:

**Theorem 2.14** *Let T be an ergodic endomorphism of a probability space* $(X, \mathscr{B}, m)$. *Let* $Y \subset X$ *be a set with positive measure, and let* $f : X \to \mathbb{R}$ *be an integrable function. Let B be a renormalizing function. We assume that* $S_n^Y f_Y / B(n/m(Y))$ *satisfies an almost sure limit theorem on Y, towards* $\mathscr{W}$. *Then* $S_n f/B(n)$ *also satisfies an almost sure limit theorem towards* $\mathscr{W}$, *on X.*

Analogues of the previous theorems hold for suspensions flows and Poincaré sections.

## 2.4 Martingale arguments

In this section we deal with the almost-sure version of the central limit theorem due to Gordin [16] (see also [27]):

**Theorem 2.15** *Let T be an ergodic endomorphism of a probability space* $(X, \mathscr{B}, m)$. *Let* $\mathscr{F} \subset \mathscr{B}$ *be a* $\sigma$*-algebra such that* $\mathscr{F} \subset T\mathscr{F}$. *Consider a square-integrable function* $f : X \to \mathbb{R}$ *such that* $\int f \,\mathrm{d}m = 0$ *and*

$$\sum_{n \geqslant 0} \left\| \mathbb{E}(f | T^n \mathscr{F}) - f \right\|_{L^2} < \infty \quad \text{and} \quad \sum_{n \geqslant 0} \left\| \mathbb{E}(f | T^{-n} \mathscr{F}) \right\|_{L^2} < \infty. \tag{22}$$

*Then there exists* $\sigma^2 \geqslant 0$ *such that* $S_n f/\sqrt{n} \xrightarrow{\text{law}} \mathscr{N}(0, \sigma^2)$.

We will prove in Sect. 8 the following theorem.

**Theorem 2.16** *Under the same assumptions,* $S_n f/\sqrt{n}$ *converges with tight maxima to* $\mathscr{N}(0, \sigma^2)$. *Moreover,* $S_n f/\sqrt{n}$ *also satisfies an almost sure limit theorem towards* $\mathscr{N}(0, \sigma^2)$.

The proof of the tight maxima is essentially a rephrasing of Example 2.3. On the other hand, the proof of the almost sure limit theorem will rely on an almost-sure limit theorem for reverse martingale differences. Since we are not aware of such a result in the literature, we will prove it, following closely the arguments in [26] for the direct martingale differences.

## 2.5 Applications

In this paragraph, we describe various dynamical systems to which the previous results apply.

### *2.5.1 Axiom A maps and flows*

Let $T : X \to X$ be the restriction of an Axiom A map to one of its basic sets. We assume that $T$ is topologically mixing. Let $m$ be a Gibbs measure with respect to some Hölder continuous potential. It is well known that, if $f$ is Hölder continuous, then $S_nf/\sqrt{n}$ converges in distribution to $\mathscr{W} = \mathscr{N}(0, \sigma^2)$ for some $\sigma^2 \geqslant 0$. Since such a transformation satisfies the ASIP, it satisfies automatically an almost sure central limit theorem as explained in Sect. 1. We nevertheless give different proofs to show in this simple example how our theorems apply.

**Proposition 2.17** *The sequence $S_nf/\sqrt{n}$ satisfies an almost sure limit theorem towards $\mathscr{W}$.*

*Proof* The simplest proof of the central limit theorem for $S_nf$ is probably to show that the assumptions of Gordin's Theorem 2.15 are satisfied for some $\sigma$-algebra $\mathscr{F}$. This is the case if one constructs $\mathscr{F}$ as follows: fix some Markov partition of $T$, and define a set to be $\mathscr{F}$-measurable if it is a union of local stable leaves intersected with elements of the Markov partition.

Using this $\mathscr{F}$, we can apply Theorem 2.15 and get the classical central limit theorem. Moreover, Theorem 2.16 also applies, and we get the almost sure limit theorem (as well as tight maxima).

Notice that, by using $K$-partitions as in [27] or [14] instead of Markov partitions, this argument extends to much more general dynamical systems.

We could also have used Theorem 2.8 to prove this result, after coding and reduction to a subshift of finite type. This argument moreover shows that the assumption of topological mixing is not necessary, topological transitivity would suffice. □

Consider now a topologically transitive Axiom A flow $T_t$ on a basic set $X$. Let $m$ be a Gibbs measure with respect to a Hölder potential. Let $f$ be a Hölder continuous function with zero average. It is well known that $\frac{1}{\sqrt{T}} \int_0^T f \circ T_t \, \mathrm{d}t$ converges in distribution to a Gaussian random variable $\mathscr{N}(0, \sigma^2)$ (see e.g. [30]).

**Proposition 2.18** *Under the same assumptions, for almost every $x \in X$,*

$$\frac{1}{\log T} \int_1^T \mathrm{d}t\, \frac{1}{t} \delta_{\int_0^t f \circ T_s(x)\, \mathrm{d}s/\sqrt{t}} \xrightarrow{\text{law}} \mathscr{N}(0, \sigma^2). \tag{23}$$

*Proof* An Axiom A flow always admits a Markov partition, and can thus be written as a suspension over a subshift of finite type. For such a subshift, the almost sure limit theorem is a consequence of Proposition 2.17 (or directly of Theorem 2.16). The flow version of Theorem 2.14 then implies the desired result for the flow. □

### *2.5.2 Locally Gibbs–Markov maps*

Let $T : X \circlearrowleft$ be a nonsingular map on a probability, metric space $(X, \mathscr{B}, m, d)$, preserving the probability measure $m$. It is said to be *locally Gibbs–Markov* if it is Markovian for a partition $\alpha$ and if there exists $Y \subset X$ of positive measure, which is a union of elements in $\alpha$, such that:

- The induced map $T_Y$ is Gibbs–Markov for the partition $\alpha_Y = \alpha \cap Y$ and the measure $m_Y = m_{|Y}/m(Y)$.
- For all $a \in \alpha_Y$, the return-time function $\varphi$ is constant on $a$, equal to an integer $\varphi_a \geqslant 1$.
- There exists $C > 0$ such that, for all $a \in \alpha_Y$, for all $x, y \in a$, for all $0 \leqslant k < \varphi_a$, we have $d(T^k x, T^k y) \leqslant C d(T^{\varphi_a} x, T^{\varphi_a} y)$.

In the present setting, we have the analog of Theorem 2.6.

**Theorem 2.19** *Let $T : X \circlearrowleft$ be an ergodic, locally Gibbs–Markov map for a subset $Y \subset X$. Let $f : X \to \mathbb{R}$ be an integrable function such that*

$$\sum_{a \in \alpha} m(a) Df(a) < \infty \tag{24}$$

*and $\int f \,\mathrm{d}m = 0$. Let $f_Y : X \to \mathbb{R}$ be defined for $y \in Y$ by $f_Y(y) = \sum_{k=0}^{\varphi(y)-1} f(T^k y)$, where $\varphi(y)$ is the return time of $y$. If $y \notin Y$, we set $f_Y(y) = 0$.*

*We assume that $f_Y$ belongs to some domain of attraction, as defined in Paragraph 2.2.1. Then $S_n f / B_n$ converges to $\mathscr{W}$, and satisfies an almost sure limit theorem towards $\mathscr{W}$, where $B_n = \sqrt{n}$ and $\mathscr{W} = \mathscr{N}(0, \sigma^2)$ for some $\sigma^2 \geqslant 0$ if $f_Y \in L^2$, and $B_n$ and $\mathscr{W}$ are as in Theorem 2.5 if for $f_Y$ we are in the 2nd or 3rd case of that theorem.*

*Proof* The function $f_Y$ belongs by assumption to some domain of attraction. Moreover, it satisfies $\sum_{a \in \alpha_Y} m_Y(a) Df_Y(a) \leqslant \frac{1}{m(Y)} \sum_{a \in \alpha} m(a) Df(a) < \infty$. Hence, the assumptions of Theorem 2.8 are satisfied. Theorem 2.10 then shows that $f_Y$ satisfies a limit theorem with tight maxima, and an almost sure limit theorem. Theorems 2.12 and 2.14 make it possible to induce these limit theorems from $Y$ to $X$. □

The classical convergence result in this theorem is proved in [17], under suitable assumptions on $\varphi$. These assumptions can be removed here due to the notion of convergence with tight maxima.

Young towers with summable return times, as defined in [32] and [33], are locally Gibbs–Markov maps. More generally, several nonuniformly expanding maps have a unique invariant absolutely continuous probability measure and can be modeled by locally Gibbs–Markov maps. This is for example the case for the Pomeau-Manneville maps in dimension 1, or the Viana maps in dimension 2. We refer the reader to [4] and [19] for more details and general statements. Theorem 2.19 applies to all these examples.

### *2.5.3 The stadium billiard*

The stadium billiard, or Bunimovich billiard, has been introduced in [5]. It is constituted of two parallel segments of length $\ell$ and two semicircles of radius 1. The transformation is the usual billiard map in this billiard table. It preserves the Liouville measure and is ergodic. Let $f$ be a Hölder function with zero average. Let $I$ denote the average of $f$ along the trajectories that bounce perpendicularly to the segments of the billiard. It is shown in [8] that, if $I \neq 0$, then $S_n f/\sqrt{n \log n}$ converges to an explicit gaussian distribution, while if $I = 0$ then $S_n f/\sqrt{n}$ converges to a gaussian distribution. So, a nonstandard normalization is needed in the first case while a standard central limit theorem holds in the second case.

**Theorem 2.20** *In both cases, the limit theorem admits an almost sure counterpart.*

*Proof* In [8], the proof of the classical limit theorem is given in the first case by a spectral argument and then an induction. Using Theorems 2.11 and 2.14 together with the arguments of [8], we therefore obtain the desired almost sure limit theorem.

In the second case, the proof of the classical limit theorem relies on a martingale argument, and then on two inductions. Once again, we can use Theorems 2.16 and 2.14 to get the conclusion. □

The paper is organized as follows. In Sect. 3, we prove Theorem 2.14, which is the only nontrivial result of the paper concerning convergence with tight maxima. The rest of the paper is essentially devoted to almost sure limit theorems, with occasional complements on convergence with tight maxima in the different settings. More precisely, in Sect. 4, we establish some general results on almost-sure limit theorems in dynamical systems that we apply subsequently. The main result of that section, which may be of independent interest, is an almost-sure version of a result by Eagleson [15] about limit theorems to be “mixing”. In Sect. 5, we easily deduce Theorem 2.14 from the general results of Sect. 4. In Sect. 6, we prove Theorem 2.11, and we show in Sect. 7 how this implies the results concerning Gibbs–Markov maps, namely Theorem 2.10. Section 8 is devoted to the proof of Theorem 2.16, i.e. the almost sure central limit theorem under Gordin’s assumptions.

## 3 Inducing classical limit theorems

In this section, we prove Theorem 2.12 and Proposition 2.13, showing that a limit theorem with tight maxima for an induced map implies a classical limit theorem for the original map.

**Theorem 3.1** *Let $(X, \mathscr{B}, m, T)$ be an ergodic probability preserving dynamical system, and let $f : X \to \mathbb{R}$. Let $B_n$ be a renormalizing sequence such that $S_n f/B_n$ converges with tight maxima to a random variable $\mathscr{W}$. Let $t_1, t_2, \dots$ be a sequence of integer valued functions on $X$ such that $t_n/n$ converges to 1 in probability. Let also $m'$ be a probability measure on $X$ which is absolutely continuous with respect to $m$. Then $S_{t_n} f/B_n$ converges in distribution to $\mathscr{W}$, for the probability measure $m'$.*

*Proof* Fix $\varepsilon > 0, \delta > 0$. We will show that, if $n$ is large enough,

$$m\left\{x : \left|\frac{S_{t_n(x)}f(x) - S_n f(x)}{B_n}\right| \geqslant \varepsilon\right\} \leqslant 2\delta. \tag{25}$$

This will imply that $(S_{t_n} f - S_n f)/B_n$ tends in probability to 0 with respect also to the measure $m'$. Since $S_n f/B_n$ converges in distribution to $\mathscr{W}$, for the probability measure $m'$, by Eagleson's Theorem [15], this will conclude the proof.

Since $S_n f/B_n$ has tight maxima, there exists $c > 0$ such that, for all $n \in \mathbb{N}$,

$$m\left\{\max_{0 \leqslant j \leqslant n} |S_j f| \geqslant c B_n\right\} \leqslant \delta. \tag{26}$$

For $z \in \mathbb{R}$, let $\lceil z \rceil$ denote the smallest integer $\geqslant z$. Since $B_n$ is a renormalizing sequence, there exists $\gamma \in (0, 1)$ small enough that, for all large enough $n$, $B_{\lceil 2\gamma n \rceil} \leqslant \varepsilon B_n/(2c)$. We fix such a $\gamma$, and write $a_n = \lceil (1 - \gamma)n \rceil$.

If $n$ is large enough, $m\{x : |t_n(x) - n| > \gamma n\} \leqslant \delta$. Then

$$m\left\{\left|\frac{S_{t_n} f - S_n f}{B_n}\right| \geqslant \varepsilon\right\} \leqslant \delta + m\left\{\left|\frac{S_{t_n} f - S_n f}{B_n}\right| \geqslant \varepsilon, t_n \in [(1 - \gamma)n, (1 + \gamma)n]\right\}.$$

If $x$ belongs to this last set, there exists $j \in [(1 - \gamma)n, (1 + \gamma)n]$ such that $|S_n f(x) - S_j f(x)| \geqslant \varepsilon B_n$. In particular, $|S_k f(x) - S_{a_n} f(x)| \geqslant \varepsilon B_n/2$ for $k = j$ or $n$. Hence,

$$m\left\{\left|\frac{S_{t_n} f - S_n f}{B_n}\right| \geqslant \varepsilon\right\} \leqslant \delta + m\left\{\max_{0 \leqslant i \leqslant 2\gamma n} |S_{a_n + i} f - S_{a_n} f| \geqslant \varepsilon B_n/2\right\}. \tag{27}$$

Since $m$ is invariant and $\varepsilon B_n/2 \geqslant c B_{\lceil 2\gamma n \rceil}$, the measure of this last set is at most

$$m\left\{\max_{0 \leqslant i \leqslant \lceil 2\gamma n \rceil} |S_i f| \geqslant c B_{\lceil 2\gamma n \rceil}\right\}. \tag{28}$$

This quantity is bounded by $\delta$ by definition of $c$. This concludes the proof of (25). □

*Proof of theorem 2.12* This result is an easy consequence of Theorem 3.1 and the techniques of [30] and [17], as we will explain now. Without loss of generality, we can assume that $T$ is invertible, since otherwise we can work in the natural extension of $T$.

For $y \in Y$ and $N \in \mathbb{N}$, let $n(y,N)$ be the greatest integer $n$ such that $S_n^Y \varphi(y) \leqslant N$. For $x \in X$, let $\pi x$ denote its first preimage belonging to $Y$. The first two steps of the proof of [17; Theorem A.1] show that $S_N f(x)/B(N) - S_{n(\pi x,N)}^Y f_Y(\pi x)/B(N)$ converges to 0 in probability. Hence, it is sufficient to prove that $S_{n(y,N)}^Y f_Y(y)/B(N)$ converges in distribution to $\mathscr{W}$, for the measure $m'$ on $Y$ with density $\mathrm{d}m' = \mathbb{1}_Y \varphi \,\mathrm{d}m$.

By assumption, $S_{\lfloor Nm(Y)\rfloor}^Y f_Y/B(N)$ converges with tight maxima to $\mathscr{W}$, with respect to $m_Y$. Moreover, $\int \varphi \,\mathrm{d}m_Y = 1/m(Y)$ by Kač' Formula. Hence, by Birkhoff's ergodic Theorem, $n(y,N) \sim Nm(Y)$ for almost all $y \in Y$. Theorem 3.1 applies and shows that $S_{n(y,N)}^Y f_Y(y)/B(N)$ converges in distribution to $\mathscr{W}$ with respect to any probability measure which is absolutely continuous with respect to $m_Y$, and in particular for $m'$. This concludes the proof. □

*Proof of proposition 2.13* For $x \in X$, let $E(x) \geqslant 0$ denote its first entrance time in $Y$. Then

$$\max_{0\leqslant k\leqslant n} |S_k f(x)| \leqslant \max_{0\leqslant k\leqslant E(x)} |S_k f(x)| + \max_{0\leqslant k\leqslant n} |S_k^Y f_Y(T^{E(x)}x)| + \max_{0\leqslant k\leqslant n} |M \circ T_Y^k(T^{E(x)}x)|.$$

Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $m(E(x) \geqslant N) \leqslant \varepsilon$. Therefore, for $c > 0$,

$$\begin{aligned} m\left\{\max_{0\leqslant k\leqslant n} |S_k f(x)| \geqslant 3cB(n)\right\} \leqslant\ & \varepsilon + m\left\{\max_{0\leqslant k\leqslant E(x)} |S_k f(x)| \geqslant cB(n)\right\} \\ & + Nm\left\{y \in Y \,:\, \max_{0\leqslant k\leqslant n} |S_k^Y f_Y(y)| \geqslant cB(n)\right\} \\ & + Nm\left\{y \in Y \,:\, \max_{0\leqslant k\leqslant n} |M(T_Y^k y)| \geqslant cB(n)\right\}. \end{aligned}$$

In the upper bound, the second term is bounded by $N\varepsilon(c)$, where $\varepsilon(c)$ tends to 0 when $c \to \infty$, since $f_Y$ has tight maxima. The last term is also bounded by $N\varepsilon(c)$, by (20). We fix $c$ so that the second and third term are $\leqslant \varepsilon$. Then the first term tends to 0 when $n \to \infty$. For large enough $n$, we get $m(\max_{0\leqslant k\leqslant n} |S_k f(x)| \geqslant 3cB(n)) \leqslant 4\varepsilon$. □

## 4 General results for almost-sure limit theorems in dynamics

An almost-sure limit theorem in dynamics is a statement of the following type: Let $T : X \circlearrowleft$ be an ergodic map preserving a probability measure $m$. Let $f : X \to \mathbb{R}$. Under certain assumptions, there exists a renormalizing sequence $B_n$ such that, for almost every $x$,

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}\delta_{S_k f(x)/B_k} \tag{29}$$

converges weakly to a probability measure on $\mathbb{R}$.

Let $g_n$ be a sequence of real, Lipschitz functions with compact support which are dense (for the topology of uniform convergence) in the space of continuous functions with compact support. The convergence of (29) is then equivalent to the convergence, for each $n$, of the sequence

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}\,g_n\left(\frac{S_k f(x)}{B_k}\right) \tag{30}$$

as $N \to \infty$. For technical commodity, we will be mainly interested in convergences like in (30).

The first important observation is that the convergence in (30) does not depend on the asymptotic class of $B_k$:

**Lemma 4.1** *Let $x_k$ be a real sequence and let $g : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function with compact support. Assume that $\frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}g(x_k)$ converges to a limit $E$. Then, for any sequence $\rho_k$ which tends to 1 when $k \to \infty$,*

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}g(\rho_k x_k) \to E. \tag{31}$$

*Proof* It is sufficient to prove that $g(\rho_k x_k) - g(x_k) \to 0$ when $k \to \infty$. Thus it suffices to prove that there exists a constant $C$ such that

$$\forall x \in \mathbb{R}, \forall \rho \in \mathbb{R},\ |g(x) - g(\rho x)| \leqslant C|1-\rho|\,. \tag{32}$$

Let $K$ be such that $g$ is equal to zero off $[-K, K]$. If $|x| \leqslant 2K$, we have $|g(x) - g(\rho x)| \leqslant \|g\|\,|x-\rho x| \leqslant 2K\,\|g\|\,|1-\rho|$. If $|x| \geqslant 2K$ and $\rho \geqslant 1/2$, we have $|\rho x| \geqslant |x|/2 \geqslant K$. Therefore, $|g(x) - g(\rho x)| = 0 \leqslant |1-\rho|$. Finally, if $|x| \geqslant 2K$ and $\rho \leqslant 1/2$, we have $|g(x) - g(\rho x)| \leqslant \|g\|_{L^\infty} \leqslant 2\,\|g\|_{L^\infty}\,|1-\rho|$. This proves (32) in all cases. □

The next step is to prove that the convergence in (30) is equivalent, under mild assumptions, to the convergence of more general sums, where the normalization factor $1/k$ is replaced by a factor of the form $\varphi(T^k x)/k$. This is an analog for almost-sure limit theorems of a result by Eagleson [15], which states that the convergence in law of $S_nf/B_n$ for the invariant measure $\mathrm{d}m$ is equivalent to the same convergence for a measure $\varphi\,\mathrm{d}m$ with $\varphi \geqslant 0$ and $\int \varphi\,\mathrm{d}m = 1$.

**Theorem 4.2** *Let $T : X \circlearrowleft$ be an ergodic map preserving a probability measure $m$ and let $f \in L^1(X)$. Let $B_n$ be a renormalizing sequence. Let $g$ be a bounded, Lipschitz function on $\mathbb{R}$. Then the following two conditions are equivalent:*

1. *There exist a function $\varphi \in L^1(X)$ with non-zero integral and a set $A \subset X$ with positive measure such that, for all $x \in A$, the quantity*

$$\nu_{N,\varphi,g}(x) = \frac{1}{\log N}\sum_{k=1}^{N}\frac{\varphi(T^k x)}{k}\, g\left(\frac{S_k f(x)}{B_k}\right) \tag{33}$$

   *converges to a limit $I(x)$, which may depend on $x$, when $N \to \infty$.*
2. *There exists $I \in \mathbb{R}$ such that, for any function $\varphi \in L^1(X)$, for almost every $x \in X$, $\nu_{N,\varphi,g}(x)$ converges to $I \int \varphi\,\mathrm{d}m$ when $N \to \infty$.*

This theorem applies in particular when $f$ satisfies an almost-sure limit theorem, since the first condition is then satisfied for $\varphi \equiv 1$.

The proof of this theorem relies on several technical lemmas. In the remaining part of this section, $T$ will be an ergodic endomorphism on a probability space $(X, \mathscr{B}, m)$, $g$ will be a bounded, Lipschitz function on $\mathbb{R}$, and $B_n$ will be a renormalizing sequence.

**Lemma 4.3** *Let $\varphi \in L^1(X)$ and $\psi \in L^1(X)$. Then, for almost every $x \in X$,*

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{\varphi(T^k x)}{k}\min\left(1, \frac{|\psi(T^k x)|}{B_k}\right) \to 0. \tag{34}$$

*Proof* Let us first prove the lemma for $\varphi = 1$. Let

$$u_N(x) = \frac{1}{\log N}\sum_{k=1}^{N-1}\frac{1}{k}\min\left(1, \frac{|\psi(T^k x)|}{B_k}\right). \tag{35}$$

We have $\int u_N\,\mathrm{d}m = O(1/\log N)$ since $\sum 1/(kB_k) < +\infty$. Letting $N_p = \lfloor \exp(p^2) \rfloor$, we get $\sum \left\| u_{N_p} \right\|_{L^1} < \infty$. Consequently, for almost every $x$, $u_{N_p}(x) \to 0$ when $p \to \infty$. Moreover, if $N_p \leqslant N < N_{p+1}$, the error made by replacing $u_{N_p}(x)$ by $u_N(x)$ tends uniformly to 0. Hence, $u_N(x)$ tends almost everywhere to 0. This proves (34) for $\varphi = 1$, and consequently for any bounded $\varphi$.

If $\varphi$ belongs only to $L^1$, notice that $v_k(x) = \sum_{i=0}^{k-1} |\varphi(T^i x)|$ satisfies $v_k(x) \sim k \|\varphi\|_{L^1}$ for almost every $x$. Moreover,

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{|\varphi(T^k x)|}{k} = \frac{1}{\log N}\sum_{k=1}^{N}\frac{v_{k+1}(x)-v_k(x)}{k}$$
$$= \frac{1}{\log N}\left(\frac{v_{N+1}(x)}{N} - v_1(x) + \sum_{k=2}^{N} v_k(x)\left(\frac{1}{k-1}-\frac{1}{k}\right)\right). \tag{36}$$

Hence, the limsup of this quantity is at most $\|\varphi\|_{L^1}$, for almost every $x$.

Finally, decompose $\varphi$ as $\varphi_1 + \varphi_2$ where $\varphi_1$ is bounded and $\|\varphi_2\|_{L^1} \leqslant \varepsilon$. Using the convergence (34) for $\varphi_1$ and the previous argument for $\varphi_2$, we get that, for almost every $x$,

$$\limsup_{N\to\infty}\frac{1}{\log N}\sum_{k=1}^{N}\frac{\varphi(T^k x)}{k}\min\left(1, \frac{|\psi(T^k x)|}{B_k}\right) \leqslant \varepsilon. \tag{37}$$

Letting $\varepsilon$ tend to 0 concludes the proof. □

Let us note the following consequence of (36), which will be used several times in the sequel.

**Lemma 4.4** *If $\varphi \in L^1(X)$ then for almost every $x \in X$*

$$\limsup_{N\to\infty} |\nu_{N,\varphi,g}(x)| \leqslant \|g\|_{L^\infty} \|\varphi\|_{L^1}. \tag{38}$$

We now come to a more important invariance lemma.

**Lemma 4.5** *If $\varphi \in L^1(X)$ then for almost every $x \in X$*

$$\limsup_{N\to\infty} |\nu_{N,\varphi,g}(x) - \nu_{N,\varphi\circ T,g}(x)| = 0. \tag{39}$$

*Proof* We first prove (39) under the additional assumption that $\varphi$ is bounded. We have

$$\nu_{N,\varphi,g}(x) - \nu_{N,\varphi\circ T,g}(x)$$
$$= \frac{1}{\log N}\left[\sum_{k=1}^{N}\frac{\varphi(T^k x)}{k}\, g\left(\frac{S_k f(x)}{B_k}\right) - \sum_{k=1}^{N}\frac{\varphi(T^{k+1} x)}{k}\, g\left(\frac{S_k f(x)}{B_k}\right)\right]$$
$$= \frac{1}{\log N}\left[\varphi(Tx)\, g\left(\frac{f(x)}{B_1}\right) - \frac{\varphi(T^{N+1}x)}{N}\, g\left(\frac{S_N f(x)}{B_N}\right)\right]$$

$$+\frac{1}{\log N}\sum_{k=2}^{N}\frac{\varphi(T^k x)}{k}\left[g\left(\frac{S_k f(x)}{B_k}\right)-\frac{k}{k-1}\,g\left(\frac{S_{k-1}f(x)}{B_{k-1}}\right)\right].$$

The first term tends to 0 when $N\to\infty$, so it suffices to estimate the second term. We have

$$\begin{aligned}
&\left|g\left(\frac{S_k f(x)}{B_k}\right)-\frac{k}{k-1}\,g\left(\frac{S_{k-1}f(x)}{B_{k-1}}\right)\right|\\
&\quad\leqslant \frac{\|g\|_{L^\infty}}{k-1}+\left|g\left(\frac{S_k f(x)}{B_k}\right)-g\left(\frac{S_{k-1}f(x)}{B_{k-1}}\right)\right|\\
&\quad\leqslant \frac{\|g\|_{L^\infty}}{k-1}+\left|g\left(\frac{S_k f(x)}{B_k}\right)-g\left(\frac{S_k f(x)}{B_{k-1}}\right)\right|\\
&\qquad+\left|g\left(\frac{S_k f(x)}{B_{k-1}}\right)-g\left(\frac{S_{k-1}f(x)}{B_{k-1}}\right)\right|.
\end{aligned}$$

We will separately estimate the contribution of each of these three terms. First,

$$\frac{1}{\log N}\sum_{k=2}^{N}\frac{\varphi(T^k x)}{k}\,\frac{\|g\|_{L^\infty}}{k-1}=O\left(\frac{1}{\log N}\right)=o(1). \tag{40}$$

Then, for almost every $x, S_k f(x)=O(k)$. Hence,

$$\left|g\left(\frac{S_k f(x)}{B_k}\right)-g\left(\frac{S_k f(x)}{B_{k-1}}\right)\right|\leqslant C|S_k f(x)|\left|\frac{1}{B_{k-1}}-\frac{1}{B_k}\right|\leqslant Ck\left|\frac{1}{B_{k-1}}-\frac{1}{B_k}\right|.$$

The sequence $B_k$ is eventually increasing, say, from the index $K$ on. Hence

$$\begin{aligned}
&\frac{1}{\log N}\sum_{k=2}^{N}\frac{\varphi(T^k x)}{k}\left|g\left(\frac{S_k f(x)}{B_k}\right)-g\left(\frac{S_k f(x)}{B_{k-1}}\right)\right|\\
&\quad\leqslant \frac{1}{\log N}\sum_{k=2}^{N}C\left|\frac{1}{B_{k-1}}-\frac{1}{B_k}\right|\leqslant\frac{C}{\log N}\left(\sum_{k=2}^{K}\left|\frac{1}{B_{k-1}}-\frac{1}{B_k}\right|+\frac{1}{B_K}\right)=o(1).
\end{aligned}$$

Finally

$$\left|g\left(\frac{S_k f(x)}{B_{k-1}}\right)-g\left(\frac{S_{k-1}f(x)}{B_{k-1}}\right)\right|\leqslant C\min\left(1,\frac{|f(T^{k-1}x)|}{B_{k-1}}\right). \tag{41}$$

Hence, the contribution of the corresponding term tends almost everywhere to 0, by Lemma 4.3. This concludes the proof when $\varphi$ is bounded.

To handle the case of a general $\varphi\in L^1$, write $\varphi=\varphi_1+\varphi_2$ where $\varphi_1$ is bounded and $\|\varphi_2\|_{L^1}\leqslant\varepsilon$. Applying the previous result to $\varphi_1$ and Lemma 4.4 to $\varphi_2$, we get almost everywhere

$$\begin{aligned}&\limsup |\nu_{N,\varphi-\varphi\circ T,g}(x)|\\&\quad\leqslant \limsup |\nu_{N,\varphi_1-\varphi_1\circ T,g}(x)| + \limsup |\nu_{N,\varphi_2-\varphi_2\circ T,g}(x)| \leqslant 0 + 2\varepsilon \, \|g\|_{L^\infty}\,.\end{aligned}$$

The conclusion of the lemma is obtained by letting $\varepsilon$ tend to 0. □

**Lemma 4.6** *If $\varphi \in L^1(X)$ then for almost every $x \in X$*

$$\limsup_{N\to\infty} |\nu_{N,\varphi\circ T,g}(x) - \nu_{N,\varphi,g}(Tx)| = 0. \tag{42}$$

*Proof* We have

$$\begin{aligned}&|\nu_{N,\varphi\circ T,g}(x) - \nu_{N,\varphi,g}(Tx)|\\&\quad= \frac{1}{\log N}\left|\sum_{k=1}^{N}\frac{\varphi(T^{k+1}x)}{k}\left[g\left(\frac{S_kf(x)}{B_k}\right) - g\left(\frac{S_kf(Tx)}{B_k}\right)\right]\right|\\&\quad\leqslant \frac{C}{\log N}\sum_{k=1}^{N}\frac{\varphi(T^{k+1}x)}{k}\min\left(1, \frac{|f(T^kx) - f(x)|}{B_k}\right).\end{aligned}$$

By Lemma 4.3, this term converges to 0 almost everywhere. The lemma is proved. □

*Proof of theorem 4.2* Let us suppose that there exists $\varphi \in L^1$ whose integral is non-zero, and such that $\nu_{N,\varphi,g}(x)$ converges on a set of positive measure. We can suppose that $\int \varphi \, dm = 1$. Otherwise, replace $\varphi$ by $\varphi/\int \varphi \, dm$. By Lemmas 4.5 and 4.6 we have, for almost every $x$,

$$\limsup_{N\to\infty} |\nu_{N,\varphi,g}(x) - \nu_{N,\varphi,g}(Tx)| = 0\,. \tag{43}$$

In particular, the set of $x$'s for which $\nu_{N,\varphi,g}(x)$ converges is invariant. Hence, by ergodicity, it is of measure one. Moreover, the limit is an invariant function, hence a constant one. Denote it by $I$.

Lemma 4.5 also gives that, for all $k \in \mathbb{N}^*$, for almost every $x \in X$,

$$\nu_{N,S_k\varphi/k,g}(x) \to I. \tag{44}$$

Let $\varepsilon > 0$. Choose $k$ such that $\|S_k\varphi/k - 1\|_{L^1} \leqslant \varepsilon$. Then, for almost every $x$,

$$\begin{aligned}&\limsup_{N\to\infty} |\nu_{N,1,g}(x) - I|\\&\quad\leqslant \limsup_{N\to\infty} |\nu_{N,1,g}(x) - \nu_{N,S_k\varphi/k,g}(x)| + \limsup_{N\to\infty} |\nu_{N,S_k\varphi/k,g}(x) - I|.\end{aligned}$$

The first term is at most $\varepsilon \, \|g\|_{L^\infty}$, by Lemma 4.4. The second one goes to 0. Finally, by letting $\varepsilon$ tend to 0, we end up with: for almost every $x$,

$$\nu_{N,1,g}(x) \to I. \tag{45}$$

Now let $\psi \in L^1(X)$ be an arbitrary function. Let $\varepsilon > 0$, choose $k \in \mathbb{N}$ such that $\left\| S_k\psi/k - \int \psi \,\mathrm{d}m \right\|_{L^1} \leqslant \varepsilon$. Then, for almost every $x$,

$$\begin{aligned}\limsup_{N\to\infty} \left| \nu_{N,\psi,g}(x) - I \int \psi \,\mathrm{d}m \right| \leqslant\ & \limsup_{N\to\infty} |\nu_{N,\psi,g}(x) - \nu_{N,S_k\psi/k,g}(x)| \\ &+ \limsup_{N\to\infty} |\nu_{N,S_k\psi/k,g}(x) - \nu_{N,\int \psi \,\mathrm{d}m,g}(x)| \\ &+ \limsup_{N\to\infty} \left| \nu_{N,\int \psi \,\mathrm{d}m,g}(x) - I \int \psi \,\mathrm{d}m \right|.\end{aligned}$$

The first term tends to 0 by Lemma 4.5. We already proved that the third term goes to 0. Finally, the second one is at most $\varepsilon \, \|g\|_{L^\infty}$, by Lemma 4.4. We conclude the proof by sending $\varepsilon$ to 0. □

## 5 Inducing almost sure limit theorems

In this section we prove Theorem 2.14. For this purpose, it suffices, according to the discussion at the beginning of Sect. 4, to establish the following theorem:

**Theorem 5.1** *Let $T : X \circlearrowleft$ be an ergodic map preserving a probability measure $m$. Let $Y \subset X$ be a set of positive measure and denote by $T_Y : Y \circlearrowleft$ the map induced by $T$, and by $\varphi$ the first return-time function. Let $f : X \to \mathbb{R}$ be integrable, and define $f_Y : Y \to \mathbb{R}$ by*

$$f_Y(y) = \sum_{k=0}^{\varphi(y)-1} f(T^k y). \tag{46}$$

*We will write $S_k^Y$ for the Birkhoff sums for the map $T_Y$.*

*Let $g : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function with compact support. Let $B$ be a renormalizing function. Assume that for almost every $y \in Y$,*

$$\frac{1}{\log N} \sum_{k=1}^{N} \frac{1}{k}\, g\left( \frac{S_k^Y f_Y(y)}{B(k/m(Y))} \right) \to E \tag{47}$$

*for some constant $E$. Then, for almost every $x \in X$,*

$$\frac{1}{\log N} \sum_{k=1}^{N} \frac{1}{k}\, g\left( \frac{S_k f(x)}{B(k)} \right) \to E. \tag{48}$$

*Proof* By Theorem 4.2, it is sufficient to prove that, for almost every $x \in Y$,

$$\frac{1}{\log N} \sum_{q=1}^{N} \frac{1}{q}\, g\left( \frac{S_q f(x)}{B(q)} \right) \to E. \tag{49}$$

Let $x \in Y$. Set $t_k(x) = \sum_{i=0}^{k-1} \varphi(T_Y^i x)$: these are the successive return times of $x$ to $Y$. For almost every $x$, we have $t_k \sim k/m(Y)$ by Birkhoff's ergodic Theorem applied to $T_Y$ and for the $T_Y$-invariant measure $m_Y = m_{|Y}/m(Y)$. Since $\log t_k \sim \log t_{k+1} \sim \log k$, we are left to prove (49) for times $N$ of the form $t_k$. We have

$$\frac{1}{\log t_k} \sum_{q=1}^{t_k} \frac{1}{q} g\left(\frac{S_q f(x)}{B(q)}\right) = \frac{1}{\log t_k} \sum_{p=0}^{k-1} \sum_{q=t_p+1}^{t_{p+1}} \frac{1}{q} g\left(\frac{S_q f(x)}{B(q)}\right). \tag{50}$$

For $q \in \mathbb{N}^*$, let $p = p(x,q)$ be the largest integer such that $S_p^Y \varphi(x) < q$. For almost all $x$, we have $S_n^Y \varphi(x) \sim n/m(Y)$, which yields $p(x,q) \sim qm(Y)$. In particular, $1/q \sim m(Y)/p$, and $B(q) \sim B(p/m(Y))$. Using Lemma 4.1, we get

$$\frac{1}{\log t_k} \sum_{p=0}^{k-1} \sum_{q=t_p+1}^{t_{p+1}} \frac{1}{q} g\left(\frac{S_q f(x)}{B(q)}\right) = \frac{1}{\log k} \sum_{p=1}^{k-1} \frac{m(Y)}{p} \sum_{q=t_p+1}^{t_{p+1}} g\left(\frac{S_q f(x)}{B(p/m(Y))}\right) + o(1).$$

For $y \in Y$, let $F(y) = \sum_{k=0}^{\varphi(y)-1} |f(T^k y)|$. Then $|S_q f(x) - S_p^Y f_Y(x)| \leqslant F(T_Y^p x)$. Therefore,

$$\frac{1}{\log k} \left| \sum_{p=1}^{k-1} \frac{m(Y)}{p} \sum_{q=t_p+1}^{t_{p+1}} g\left(\frac{S_q f(x)}{B(p/m(Y))}\right) - \sum_{p=1}^{k-1} \frac{m(Y)\varphi(T_Y^p x)}{p} g\left(\frac{S_p^Y f_Y(x)}{B(p/m(Y))}\right)\right|$$
$$\leqslant \frac{C}{\log k} \sum_{p=1}^{k-1} \frac{\varphi(T_Y^p x)}{p} \min\left(1, \frac{F(T_Y^p x)}{B(p/m(Y))}\right).$$

By Lemma 4.3, this term tends almost everywhere to 0. We have proved that, for almost every $x \in Y$,

$$\frac{1}{\log t_k} \sum_{q=1}^{t_k} \frac{1}{q} g\left(\frac{S_q f(x)}{B(q)}\right) = \frac{1}{\log k} \sum_{p=1}^{k-1} \frac{m(Y)\varphi(T_Y^p x)}{p} g\left(\frac{S_p^Y f_Y(x)}{B(p/m(Y))}\right) + o(1).$$

The assumption (47) together with Theorem 4.2 show that this last term converges almost everywhere to $E$. □

*Remark 5.2* It is possible to give a quicker proof of Theorem 5.1 by proving instead of (49) that, for almost every $x \in Y$,

$$\frac{1}{\log N} \sum_{q=1}^{N} \frac{\mathbb{1}_Y(T^q x)}{q} g\left(\frac{S_q f(x)}{B(q)}\right) \to Em(Y). \tag{51}$$

This is sufficient to conclude, by Theorem 4.2. And there are less computations to check (51) than (49). However, the problem of this new proof is that it can not be generalized easily to the case of flows, contrary to the proof given above.

## 6 Almost sure limit theorems by spectral methods

In this section, we prove Theorem 2.11. Hence, we will consider a dynamical system $(X, T, m)$ and a function $f : X \to \mathbb{R}$, and assume that the assumptions 1–4 of Paragraph 2.2.3 hold.

**Lemma 6.1** *Let $(\nu_j)_{1\leqslant j\leqslant N}$ be two by two distinct complex numbers, and let $(a_j)_{1\leqslant j\leqslant N}$ be complex numbers. Assume that*

$$\sup_{n\in\mathbb{N}}\left|\sum_{j=1}^{N} a_j\nu_j^n\right| < \infty. \tag{52}$$

*Then, for all j, either $a_j = 0$ or $|\nu_j| \leqslant 1$.*

*Proof* Define on the open unit disk in $\mathbb{C}$ an analytic function

$$\varphi(z) = \sum_{n=1}^{\infty}\left(\sum_{j=1}^{N} a_j\nu_j^n\right) z^n. \tag{53}$$

It coincides with the function $\sum \frac{a_j}{1-\nu_j z}$ on a small neighborhood of zero, hence on the whole unit disk. In particular, it has no pole there. This implies that $|\nu_j| \leqslant 1$ whenever $a_j \neq 0$. □

**Lemma 6.2** *There exist $\varepsilon_1 \leqslant \varepsilon_0, 0 < \beta \leqslant \beta_0, C > 0$, $0 < \rho < 1$ and a function $\lambda : [-\varepsilon_1, \varepsilon_1] \to \mathbb{C}$ such that $|\lambda(t)| \leqslant 1$ and, for all $t, t' \in [-\varepsilon_1, \varepsilon_1]$, for all $n, p \in \mathbb{N}$,*

$$\left|\mathbb{E}(e^{itS_nf\circ T^p}e^{it'S_pf}) - \lambda(t)^n\lambda(t')^p\right| \leqslant C|t|^\beta + C|t'|^\beta + C\rho^n + C\rho^p. \tag{54}$$

*Moreover,*

$$|\lambda(t) - 1| \leqslant C|t|^\beta. \tag{55}$$

*In particular,*

$$\left|\mathbb{E}(e^{itS_nf}) - \lambda(t)^n\right| \leqslant C|t|^\beta + C\rho^n. \tag{56}$$

*Proof* The inequality (56) (for $t'$ and $p$ instead of $t$ and $n$) is a consequence of (54) by taking $t = 0$ and letting $n$ tend to infinity. Hence, we just have to prove (54). Estimate (55) will be proved along the way.

The operator $\mathscr{L}_0$ acting on $\mathscr{G}$ has a simple eigenvalue at 1, and possibly other eigenvalues $\nu_1, \dots, \nu_k$ of modulus $\geqslant 1$. The assumptions 2 and 4 of Paragraph 2.2.3 yield the following spectral description of $\mathscr{L}_t$ for small enough $t$, by Theorem 1 and Corollary 1 in [24].

The operator $\mathscr{L}_t$ has an eigenvalue $\lambda(t)$ close to 1, and eigenvalues $\nu_j(t)$ close to $\nu_j$. Denoting by $P(t)$ and $Q_j(t)$ the corresponding spectral projections, we can write

$$\mathscr{L}_t = \lambda(t)P(t) + \sum \nu_j(t)Q_j(t) + N(t), \tag{57}$$

where $N(t)$ satisfies $\|N(t)^n\| \leqslant C\rho^n$ uniformly in $t$, for some $\rho < 1$. Moreover, for any small enough $\beta > 0$, we have

$$|\lambda(t) - 1| \leqslant C|t|^\beta, \quad |\nu_j(t) - \nu_j| \leqslant C|t|^\beta. \tag{58}$$

Moreover,

$$\|\mathrm{i} \circ (P(t) - P(0))\|_{\mathscr{G}\to\mathscr{H}} \leqslant C|t|^\beta, \quad \left\|\mathrm{i} \circ (Q_j(t) - Q_j(0))\right\|_{\mathscr{G}\to\mathscr{H}} \leqslant C|t|^\beta. \tag{59}$$

Finally, the norms $\|P(t)\|_{\mathscr{G}\to\mathscr{G}}$ and $\left\|Q_j(t)\right\|_{\mathscr{G}\to\mathscr{G}}$ are uniformly bounded, and these operators satisfy

$$\|P(t)u\|_{\mathscr{G}} \leqslant C\,\|\mathrm{i}(P(t)u)\|_{\mathscr{H}}, \quad \left\|Q_j(t)u\right\|_{\mathscr{G}} \leqslant C\left\|\mathrm{i}(Q_j(t)u)\right\|_{\mathscr{H}}. \tag{60}$$

To simplify the notations, we will write $\nu_0(t) = \lambda(t)$ and $Q_0(t) = P(t)$. Let us check some algebraic consequences of this spectral description. First, we have for any $u \in \mathscr{G}$

$$\left\|Q_0(0)(Q_j(t) - Q_j(0))u\right\|_{\mathscr{G}} \leqslant C|t|^\beta. \tag{61}$$

Indeed, by (60),

$$\left\|Q_0(0)(Q_j(t) - Q_j(0))u\right\|_{\mathscr{G}} \leqslant \left\|\mathrm{i}(Q_0(0)(Q_j(t) - Q_j(0))u)\right\|_{\mathscr{H}}. \tag{62}$$

If $j = 0$, this quantity is equal to

$$\|\mathrm{i}(Q_0(0)(Q_0(t) - \mathrm{Id})u)\|_{\mathscr{H}} = \|\mathrm{i}((Q_0(0) - Q_0(t))(Q_0(t) - \mathrm{Id})u)\|_{\mathscr{H}} \leqslant C|t|^\beta,$$

by (59) and the uniform boundedness of $\|Q_0(t)\|_{\mathscr{G}\to\mathscr{G}}$. On the other hand, if $j \neq 0$, then (62) is equal to

$$\left\|\mathrm{i}(Q_0(0)Q_j(t)u)\right\|_{\mathscr{H}} = \left\|\mathrm{i}((Q_0(0) - Q_0(t))Q_j(t)u)\right\|_{\mathscr{H}}, \tag{63}$$

which is again bounded by $C|t|^\beta$. This proves (61)

Let us now prove

$$\langle \ell_0, Q_j(t)Q_{j'}(t')\alpha_0\rangle = \delta_{j0}\delta_{j'0} + O(|t|^\beta) + O(|t'|^\beta). \tag{64}$$

Since $\ell_0$ is the fixed point of $\mathscr{L}_0'$, we have

$$\langle \ell_0, Q_j(t)Q_{j'}(t')\alpha_0\rangle = \langle \ell_0, Q_0(0)Q_j(t)Q_{j'}(t')\alpha_0\rangle. \tag{65}$$

Moreover,

$$\begin{aligned} Q_0(0)Q_j(t)Q_{j'}(t') = {} & Q_0(0)(Q_j(t) - Q_j(0))Q_{j'}(t') \\ & + Q_0(0)Q_j(0)(Q_{j'}(t') - Q_{j'}(0)) + Q_0(0)Q_j(0)Q_{j'}(0). \end{aligned}$$

The last term is equal to $\delta_{j0}\delta_{j'0}Q_0(0)$, while the other ones are bounded by $C|t|^\beta$ and $C|t'|^\beta$ by (61). This proves (64).

We can now compute. We have

$$\begin{aligned}\mathbb{E}(e^{itS_nf\circ T^p}e^{itS_pf}) &= \langle \ell_0, \mathscr{L}_t^n\mathscr{L}_{t'}^p\alpha_0\rangle \\ &= \sum_{j,j'=0}^{k} \nu_j(t)^n\nu_{j'}(t')^p\langle \ell_0, Q_j(t)Q_{j'}(t')\alpha_0\rangle \\ &\quad+\sum_j \nu_j(t)^n\langle \ell_0, Q_j(t)N(t')^p\alpha_0\rangle \\ &\quad+\sum_{j'}\nu_{j'}(t')^p\langle \ell_0, N(t)^nQ_{j'}(t')\alpha_0\rangle \\ &\quad+\langle \ell_0, N(t)^nN(t')^p\alpha_0\rangle. \end{aligned} \tag{66}$$

We will show that, in this formula, whenever there is a coefficient $\nu_j(t)$ or $\nu_{j'}(t')$ of modulus $> 1$, then the corresponding factor vanishes. By symmetry, it suffices to do that for $\nu_j(t)$.

Fix $p\in\mathbb{N}$. The previous formula implies that

$$\sum_j\left(\langle \ell_0, Q_j(t)N(t')^p\alpha_0\rangle+\sum_{j'}\nu_{j'}(t')^p\langle \ell_0, Q_j(t)Q_{j'}(t')\alpha_0\rangle\right)\nu_j(t)^n \tag{67}$$

is uniformly bounded, independently of $n$. Let $j$ be such that $|\nu_j(t)| > 1$. Lemma 6.1 then shows that

$$\langle \ell_0, Q_j(t)N(t')^p\alpha_0\rangle+\sum_{j'}\nu_{j'}(t')^p\langle \ell_0, Q_j(t)Q_{j'}(t')\alpha_0\rangle = 0. \tag{68}$$

Multiply this equation by $\rho^{-p/2}$. Then $\rho^{-p/2}\langle \ell_0, Q_j(t)N(t')^p\alpha_0\rangle$ is still tending to 0, while $\rho^{-1/2}\nu_{j'}(t')$ has modulus $> 1$ for any $j'$, if $t'$ is small enough. Applying once again Lemma 6.1 (but varying $p$ this time), this shows that, for all $j'$, $\langle \ell_0, Q_j(t)Q_{j'}(t')\alpha_0\rangle = 0$. In turn, we obtain $\langle \ell_0, Q_j(t)N(t')^p\alpha_0\rangle = 0$. We have shown that, whenever $|\nu_j(t)| > 1$, all the corresponding factors vanish in (66).

By (64), the factor of $\lambda(t)^n\lambda(t')^p$ is $1 + O(|t|^\beta) + O(|t'|^\beta)$, which is nonzero if $t$ and $t'$ are small enough. This yields $|\lambda(t)| \leqslant 1, |\lambda(t')| \leqslant 1$. The factors of the other terms in (66) are $O(|t|^\beta) + O(|t'|^\beta)$ by (64). Hence, we have proved

$$\mathbb{E}(e^{itS_nf\circ T^p}e^{itS_pf}) = \lambda(t)^n\lambda(t')^p + O(\rho^n) + O(\rho^p) + O(|t|^\beta) + O(|t'|^\beta).$$

□

**Corollary 6.3** *Let $B_n \to \infty$. The random variables $S_n f/B_n$ converge in distribution towards a random variable $\mathscr{W}$ if and only if, for all $t \in \mathbb{R}$,*

$$\lambda(t/B_n)^n \to \mathbb{E}(\mathrm{e}^{it\mathscr{W}}). \tag{69}$$

*Proof* The convergence in distribution of random variables is equivalent to the pointwise convergence of the characteristic functions. That is, $S_n f/B_n$ converges to $\mathscr{W}$ if and only if, for all $t \in \mathbb{R}$,

$$\mathbb{E}(e^{itS_n f/B_n}) \to \mathbb{E}(\mathrm{e}^{it\mathscr{W}}). \tag{70}$$

By (56), this is equivalent to (69). □

This corollary has been proved and used by Hervé in [22].

*Proof of theorem 2.11* We will prove that, for all $t \geqslant 0$, for almost all $x \in X$,

$$\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\exp(itS_k f(x)/B_k) \to \mathbb{E}(\mathrm{e}^{it\mathscr{W}}) \tag{71}$$

and

$$\int_{|s|\leqslant t}\frac{1}{\log n}\sum_{k=1}^{n}\frac{1}{k}\exp(isS_k f(x)/B_k)\,\mathrm{d}s \to \int_{|s|\leqslant t}\mathbb{E}(\mathrm{e}^{is\mathscr{W}})\,\mathrm{d}s. \tag{72}$$

By [26; Lemma 6.7], this will imply the desired almost sure limit theorem. We will only prove (71), since the other equation follows from the same estimates. So, let us fix $t \in \mathbb{R}$ until the end of the proof.

We will need the following abstract lemma.

**Lemma 6.4** *There exists a constant $C > 0$ such that, for any $N \in \mathbb{N}$ and any $z \in \mathbb{C}$, if $|z^j - 1| \leqslant 1/2$ for all $j = 0, \dots, N$, then $|z - 1| \leqslant \frac{C}{N}$.*

*Proof* Write $z = r\mathrm{e}^{i\theta}$ with $r > 0$ and $\theta \in [-\pi/2, \pi/2]$. Since $r^N \in [1/2, 2]$, we get $|r - 1| \leqslant C/N$ for some constant $C$. If $\theta \neq 0$, let $n \in \mathbb{N}$ be minimal such that $|\theta n| > \pi/2$ (with $n \geqslant 2$ by assumption). Then $|\theta n| \leqslant 2|\theta(n-1)| \leqslant 2 \cdot \pi/2 = \pi$. Hence, $\theta n \in [-\pi, -\pi/2) \cup (\pi/2, \pi]$. In particular, $z^n \notin B(1, 1/2)$. This yields $n > N$. In particular, $|\theta| \leqslant \pi/(2N)$. □

**Lemma 6.5** *There exists $C > 0$ such that, for all integers $k, l$ with $l \geqslant Ck$, for all $t' \in \mathbb{R}$ with $|t'| \leqslant |t|$,*

$$|\lambda(t'/B_l)^k - 1| \leqslant C\frac{k}{l}\cdot \tag{73}$$

*Proof* The random variables $S_n f/B_n$ converge in distribution to $\mathscr{W}$. Hence, their characteristic functions converge, uniformly on every compact subinterval of $\mathbb{R}$. In particular, there exist $N > 0$ and $A > 0$ such that, for all $u \in [-A, A]$ and all $n \geqslant N$,

$$|\mathbb{E}(e^{iuS_n f/B_n}) - 1| \leqslant 1/10. \tag{74}$$

Let $M$ be a large constant, and consider $l \geqslant 2MN$. Let $j \in [N, l/M]$. If $M$ is large enough, then $|tB_j/B_l| \leqslant A$. Write $u = t'B_j/B_l$. Then, by (56),

$$\lambda(t'/B_l)^j = \lambda(u/B_j)^j = \mathbb{E}(e^{iuS_jf/B_j}) + O(|u|^\beta) + O(\rho^j). \tag{75}$$

Increasing $N$ and $M$ if necessary, we can ensure that the $O$ terms are bounded by 1/10. We get

$$\left|\lambda(t'/B_l)^j - \mathbb{E}(e^{iuS_jf/B_j})\right| \leqslant 1/10. \tag{76}$$

Given (74), this yields

$$\left|\lambda(t'/B_l)^j - 1\right| \leqslant 1/5. \tag{77}$$

Consider now $j \in [0, N)$. Since $N, j+N \in [N, l/M]$, (77) applies to these two numbers. Therefore,

$$\left|\lambda\left(\frac{t'}{B_l}\right)^j - 1\right| = \left|\frac{\left(\lambda(t'/B_l)^{j+N} - 1\right) - \left(\lambda(t'/B_l)^N - 1\right)}{\lambda(t'/B_l)^N}\right| \leqslant \frac{1/5 + 1/5}{4/5} = \frac{1}{2}.$$

Hence, for all $j \in [0, l/M]$,

$$|\lambda(t'/B_l)^j - 1| \leqslant 1/2. \tag{78}$$

By Lemma 6.4, $|\lambda(t'/B_l) - 1| \leqslant C/l$. Finally, if $k \leqslant l/M$,

$$\left|\lambda(t'/B_l)^k - 1\right| = \left|\lambda(t'/B_l) - 1\right| \left|\sum_{i=0}^{k-1} \lambda(t'/B_l)^i\right| \leqslant \frac{C}{l}\sum_{i=0}^{k-1}\frac{3}{2} \leqslant C\frac{k}{l}.$$

□

Let $\xi_k(x) = \exp(itS_kf(x)/B_k) - \mathbb{E}(\exp(itS_kf/B_k))$.

**Lemma 6.6** *For any $k, l$, we have $|\mathbb{E}(\xi_k\xi_l)| \leqslant C$. Besides, there exist $0 < \rho < 1, \delta > 0$ and $K \geqslant 1$ such that, if $l \geqslant Kk$,*

$$|\mathbb{E}(\xi_k\xi_l)| \leqslant C\frac{B_k}{B_l} + C\rho^k + C\rho^{l-k} + \frac{C}{B_k^\delta} + C\left(\frac{k}{l}\right)^{1/2}. \tag{79}$$

*Proof* Since the functions $\xi_k$ are all bounded by 2, the first estimate is trivial. If $k$ remains bounded, the result of the lemma is also trivial. Hence, we may assume that $k$ is as large as needed in the course of proof. For the rest of the proof, we will denote by $\varepsilon(k,l)$ an error term which is compatible with (79), and we will say that such an error term is admissible.

We have

$$\begin{aligned}&\mathbb{E}\big(e^{itS_kf(x)/B_k}e^{itS_lf(x)/B_l}\big)\\&\quad=\mathbb{E}\left[\exp\left(i\left(\frac{t}{B_k}+\frac{t}{B_l}\right)S_kf\right)\cdot\exp\left(i\frac{t}{B_l}S_{l-k}f\circ T^k\right)\right].\end{aligned}$$

Let $a=t/B_l$ and $b=t/B_k+t/B_l$. If $k$ is large enough, $a$ and $b$ are small enough so that (54) applies. We get

$$\mathbb{E}\big(e^{itS_kf(x)/B_k}e^{itS_lf(x)/B_l}\big)=\lambda\left(\frac{t}{B_l}\right)^{l-k}\lambda\left(\frac{t}{B_k}+\frac{t}{B_l}\right)^{k}+\varepsilon(k,l),\tag{80}$$

where $\varepsilon(k,l)$ is an admissible error term.

On the other hand, by (56),

$$\begin{aligned}&\mathbb{E}(e^{itS_kf/B_k})\mathbb{E}(e^{itS_lf/B_l})\\&\quad=\left(\lambda\left(\frac{t}{B_k}\right)^{k}+O\left(\frac{|t|^\beta}{B_k^\beta}\right)+O(\rho^k)\right)\left(\lambda\left(\frac{t}{B_l}\right)^{l}+O\left(\frac{|t|^\beta}{B_l^\beta}\right)+O(\rho^l)\right)\\&\quad=\lambda\left(\frac{t}{B_k}\right)^{k}\lambda\left(\frac{t}{B_l}\right)^{l}+\varepsilon(k,l).\end{aligned}$$

Subtracting these two expressions, we have to show that

$$\lambda\left(\frac{t}{B_l}\right)^{l-k}\lambda\left(\frac{t}{B_k}+\frac{t}{B_l}\right)^{k}-\lambda\left(\frac{t}{B_k}\right)^{k}\lambda\left(\frac{t}{B_l}\right)^{l}\tag{81}$$

is an admissible error term to conclude. Since $xx'-yy'=x(x'-y')+(x-y)y'$ and $\lambda$ is bounded by 1, it is even sufficient to prove that $\lambda(t/B_l)^{l-k}-\lambda(t/B_l)^{l}$ is admissible, as well as $\lambda(t/B_k+t/B_l)^k-\lambda(t/B_k)^k$.

Let us first study $\lambda(t/B_l)^{l-k}-\lambda(t/B_l)^{l}$. Since $\lambda(t/B_l)^l$ is uniformly bounded, it suffices to prove that $\lambda(t/B_l)^{-k}-1$ is admissible. It even suffices to prove that $\lambda(t/B_l)^{k}-1$ is admissible. This is a consequence of Lemma 6.5 if $l\geqslant Kk$ for some large enough $K$.

Let us now turn to $\lambda(t/B_k+t/B_l)^k-\lambda(t/B_k)^k$. We have

$$\begin{aligned}&\left|\lambda\left(\frac{t}{B_k}+\frac{t}{B_l}\right)^{k}-\lambda\left(\frac{t}{B_k}\right)^{k}\right|\\&\quad=\left|\mathbb{E}\left(e^{i\left(\frac{t}{B_k}+\frac{t}{B_l}\right)S_kf}\right)-\mathbb{E}\left(e^{i\frac{t}{B_k}S_kf}\right)\right|+\varepsilon(k,l)\\&\quad\leqslant\mathbb{E}\left|e^{i\frac{t}{B_l}S_kf}-1\right|+\varepsilon(k,l).\end{aligned}$$

Moreover,

$$\left(\mathbb{E}\left|\mathrm{e}^{i\frac{t}{B_l}S_kf}-1\right|\right)^2 \leqslant \mathbb{E}\left(\left|\mathrm{e}^{i\frac{t}{B_l}S_kf}-1\right|^2\right) = \mathbb{E}\left(1-\mathrm{e}^{i\frac{t}{B_l}S_kf}+1-\mathrm{e}^{-i\frac{t}{B_l}S_kf}\right)$$
$$= 1-\lambda(t/B_l)^k+1-\lambda(-t/B_l)^k+\varepsilon(k,l).$$

By Lemma 6.5, $\left|1-\lambda(t/B_l)^k\right| \leqslant Ck/l$ and $\left|1-\lambda(-t/B_l)^k\right| \leqslant Ck/l$. This concludes the proof of the lemma. □

Let

$$T_N = \frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}\left(\mathrm{e}^{itS_kf/B_k}-\mathbb{E}(\mathrm{e}^{itS_kf/B_k})\right). \tag{82}$$

Then $T_N = \frac{1}{\log N}\sum \frac{\xi_k}{k}$. Using Lemma 6.6, and the estimates $\sum_{l=k}^{\infty}\frac{1}{lB_l} \sim \frac{C}{B_k}$ and $\sum_{l=k}^{\infty}\frac{1}{l\sqrt{l}} \sim \frac{C}{\sqrt{k}}$ (see [30; Proposition 1.5.10]), we get

$$\mathbb{E}(|T_N|^2) \leqslant \frac{C}{(\log N)^2}\sum_{k=1}^{N}\frac{1}{k}\left(\sum_{k\leqslant l\leqslant Kk}\frac{1}{l}+\sum_{l=Kk}^{N}\frac{1}{l}\left(\frac{B_k}{B_l}+\rho^k+\rho^{l-k}+\frac{1}{B_k^{\delta}}+\frac{\sqrt{k}}{\sqrt{l}}\right)\right)$$
$$\leqslant \frac{C}{(\log N)^2}\sum_{k=1}^{N}\frac{1}{k}\left(\log K+\frac{B_k}{B_k}+\rho^k\log N+C+\frac{1}{B_k^{\delta}}\log N+\frac{\sqrt{k}}{\sqrt{k}}\right).$$

Since $\frac{1}{kB_k^{\delta}}$ is summable, as well as $\frac{\rho^k}{k}$, we obtain

$$\mathbb{E}(|T_N|^2) \leqslant \frac{C}{\log N}. \tag{83}$$

Let $N_p = \lfloor \exp(p^2)\rfloor$. Since $\mathbb{E}(|T_{N_p}|^2)$ is summable, $T_{N_p}(x)$ converges to 0 almost everywhere. That is, for almost all $x$,

$$\frac{1}{\log N_p}\sum_{k=1}^{N_p}\frac{1}{k}\left(\mathrm{e}^{itS_kf(x)/B_k}\right) \to \mathbb{E}(\mathrm{e}^{it\mathscr{W}}). \tag{84}$$

For a general $N$, we choose $p$ such that $N_p \leqslant N < N_{p+1}$ and check that the difference between the previous sums for $N$ and for $N_p$ converges to 0. This concludes the proof of Theorem 2.11. □

## 7 Limit theorems for Gibbs–Markov maps

In this section, we prove Theorem 2.10, by applying Theorem 2.11. Let us first recall some useful facts on Gibbs–Markov maps.

Denote by $\mathscr{G}$ the set of bounded, locally Lipschitz functions $u$ (i.e. functions $u$ satisfying $\sup_{a\in\alpha} Du(a) < \infty$ and $\|u\|_{L^\infty} < \infty$), endowed with its canonical norm

$$\|u\|_{\mathscr{G}} = \sup_{a\in\alpha} Du(a) + \|u\|_{L^\infty} . \tag{85}$$

The transfer operator $\mathscr{L}$ associated to $T$ acts on $\mathscr{G}$ and satisfies a Lasota–Yorke inequality

$$\left\|\mathscr{L}^n u\right\|_{\mathscr{G}} \leqslant C\eta^n \|u\|_{\mathscr{G}} + C\|u\|_{L^1} . \tag{86}$$

Let $f$ be a function satisfying the assumptions of Theorem 2.10. By [1,2,18], it is possible to define an operator $\mathscr{L}_t$ acting on $\mathscr{G}$ by $\mathscr{L}_t(u) = \mathscr{L}(e^{itf}u)$. Moreover, it satisfies the assumptions 1–4 of Paragraph 2.2.3 for $\mathscr{H} = L^1$ and i the canonical inclusion. Hence, Theorems 2.8 and 2.11 together show that $S_nf/B_n$ satisfies an almost sure central limit theorem.

To prove the tight maxima statement of Theorem 2.11, we will need a more precise description of the mixing properties of $\mathscr{L}$. Let $r$ be the gcd of the return times of an atom of the partition into itself. If $r = 1$, the map $T$ is mixing and its correlations decrease exponentially fast: for every function $u \in \mathscr{G}$, $\left\|\mathscr{L}^n u - \int u \,\mathrm{d}m\right\|_{\mathscr{G}} \leqslant C\eta^n \|u\|_{\mathscr{G}}$, for some $\eta < 1$. When $r > 1$, there exists a partition of $X$, say $X_0, \dots, X_{r-1}$, each $X_i$ being a union of elements of $\alpha$, such that $T$ maps $X_i$ to $X_{i+1}$, for every $i \in \mathbb{Z}/r\mathbb{Z}$. Let $\mathscr{B}_0$ be the (finite) $\sigma$-algebra generated by $\{X_0, \dots, X_{r-1}\}$, and $\Pi : u \mapsto \mathbb{E}(u|\mathscr{B}_0)$. The operator $\Pi$ is the projector on the eigenvalues with modulus 1 of $\mathscr{L}$. In particular, for every function $u \in \mathscr{G}$,

$$\left\|\mathscr{L}^n(u - \Pi u)\right\|_{\mathscr{G}} \leqslant C\eta^n \|u\|_{\mathscr{G}} \tag{87}$$

for some $\eta < 1$. Since $\Pi$ is a conditional–expectation operator, it satisfies $\int \Pi(u)v \,\mathrm{d}m = \int \Pi(u)\Pi(v) \,\mathrm{d}m$. Lastly, $\Pi$ and $\mathscr{L}$ commute.

If $u$ is integrable and $\sum m(a)Du(a) < \infty$, then

$$\mathscr{L}u \in \mathscr{G}, \text{ and } \left\|\mathscr{L}u\right\|_{\mathscr{G}} \leqslant C\left(\sum_{a\in\alpha} m(a)Du(a) + \int |u| \,\mathrm{d}m\right). \tag{88}$$

**Lemma 7.1** *Assume that $f$ belongs to some domain of attraction with $\int f \,\mathrm{d}m = 0$. Let $B_n$ be the renormalizing sequence given by Theorem 2.5 for this domain of attraction. Then $S_nf/B_n$ is bounded in $L^1$.*

*Proof* The result is clear for $\Pi f$, since its Birkhoff sums are bounded. So, without loss of generality, we can replace $f$ by $f - \Pi f$ and suppose that $\Pi f = 0$.

We will use the following estimates on the probabilities. They are the consequence of the slow variation of $L$ and of the choice of $B_n$. They are easy to verify for the three types of domain of attraction defined in Paragraph 2.2.1.

$$\int 1\!\!1_{\{|f|>B_n\}} \,\mathrm{d}m \leqslant \frac{C}{n}; \quad \int |f| 1\!\!1_{\{|f|>B_n\}} \,\mathrm{d}m \leqslant \frac{CB_n}{n}; \quad \int f^2 1\!\!1_{\{|f|\leqslant B_n\}} \,\mathrm{d}m \leqslant \frac{CB_n^2}{n}. \tag{89}$$

We define a function $\varphi_n$ by $\varphi_n(x) = f(x)$ if $|f(x)| \leqslant B_n$, and $\operatorname{sgn}(f(x))B_n$ otherwise. It satisfies $D\varphi_n(a) \leqslant Df(a)$. Besides, since $\int f \, \mathrm{d}m = 0$, (89) implies that $\left|\int \varphi_n \, \mathrm{d}m\right| \leqslant \frac{CB_n}{n}$. Then let $\psi_n = \varphi_n - \Pi\varphi_n$ and $\chi_n = f - \psi_n$. Since $\Pi f = 0$, they satisfy $\|\chi_n\|_{L^1} = O(B_n/n)$ and

$$\|\psi_n\|_{L^2}^2 \leqslant \int \varphi_n^2 \, \mathrm{d}m \leqslant \int f^2 \mathbb{1}_{\{|f| \leqslant B_n\}} \, \mathrm{d}m + B_n^2 \int \mathbb{1}_{\{|f| > B_n\}} \, \mathrm{d}m \leqslant C\frac{B_n^2}{n}. \tag{90}$$

We have $\int |S_n f| \, \mathrm{d}m \leqslant \int |S_n \psi_n| \, \mathrm{d}m + \int |S_n \chi_n| \, \mathrm{d}m$. Since $\|\chi_n\|_{L^1} = O(B_n/n)$, we have

$$\int |S_n \chi_n| \, \mathrm{d}m \leqslant CB_n. \tag{91}$$

Moreover,

$$\int |S_n \psi_n|^2 \, \mathrm{d}m = n \int (\psi_n)^2 \, \mathrm{d}m + 2\sum_{k=1}^{n} (n-k) \int \psi_n \cdot \psi_n \circ T^k \, \mathrm{d}m. \tag{92}$$

The first term is bounded by (90). For the second one,

$$\left|\int \psi_n \cdot \psi_n \circ T^k \, \mathrm{d}m\right| = \left|\int \mathscr{L}^k \psi_n \cdot \psi_n \, \mathrm{d}m\right| \leqslant C\eta^{k-1} \|\mathscr{L}\psi_n\|_{\mathscr{G}} \tag{93}$$

by (87) and because $\Pi\psi_n = 0$. The $L^1$ norm of $\psi_n$ is bounded, as well as $\sum_{a \in \alpha} m(a) D\psi_n(a)$. Hence (88) shows that $\mathscr{L}\psi_n$ is uniformly bounded in $\mathscr{G}$. By gathering the preceding equations we end up with

$$\int |S_n \psi_n|^2 \, \mathrm{d}m \leqslant Cn\frac{B_n^2}{n} + C\sum_{k=1}^{n} (n-k)\eta^k \leqslant CB_n^2. \tag{94}$$

Equations (91) and (94) show that $S_n f/B_n$ is bounded in $L^1$. This proves the lemma. □

*Proof of theorem 2.10* We have already proved the almost sure central limit theorem, it only remains to check the tight maxima statement. This statement is trivial for $\Pi f$, since its Birkhoff sums are uniformly bounded. Hence, we can without loss of generality replace $f$ with $f - \Pi f$, and assume that $\Pi f = 0$.

Let $g = \sum_{n=1}^{\infty} \mathscr{L}^n f$. This series is convergent in $\mathscr{G}$ by the spectral gap property (87), and (88). The function $h = f + g - g \circ T$ then satisfies $\mathscr{L}h = 0$, i.e. $\mathbb{E}(h | T^{-1}\mathscr{B}) = 0$. The sequence $h \circ T^n$ is therefore a reverse martingale difference for the filtration $\mathscr{F}_n = T^{-n}\mathscr{B}$. Moreover, $S_n h/B_n = S_n f/B_n + (g - g \circ T^n)/B_n$. By Lemma 7.1, $S_n f/B_n$ is bounded in $L^1$, hence $S_n h/B_n$ is also bounded in $L^1$. Consequently, Example 2.3 shows that $S_n h/B_n$ has tight maxima. To conclude, we have to show that the sequence $(g - g \circ T^n)/B_n$ also has tight maxima. This is a consequence of the boundedness of $g$. □

## 8 Almost sure limit theorems by martingale arguments

### 8.1 Almost-sure limit theorem for reverse martingale differences

**Theorem 8.1** *Let $\mathscr{F}_n$ be a decreasing sequence of $\sigma$-algebras on a probability space, and let $Z_n$ be a $\mathscr{F}_n$-measurable square-integrable random variable such that $\mathbb{E}(Z_n|\mathscr{F}_{n+1}) = 0$. Let $B_n \in \mathbb{R}_+$ increase to infinity, let $\zeta$ be a non-negative random variable, and let $b_k \in \mathbb{R}_+$ be a bounded sequence with $\sum b_k = +\infty$. Assume that*

1. *Almost surely, $Z_n/B_n \to 0$.*
2. *Almost surely, $\frac{1}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z_j^2(\omega)/B_k^2}$ converges weakly to $\delta_{\zeta(\omega)}$.*
3. *The sequence $b_k$ satisfies $b_k = O\left(\frac{B_k - B_{k-1}}{B_k}\right)$.*
4. *We have*
$$\sup_k \mathbb{E}\left[\max_{1\leqslant j\leqslant k} \frac{|Z_j|^2}{B_k^2}\right] < \infty. \tag{95}$$

*Then, for almost all $\omega$, the real measure*

$$\frac{1}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z_j(\omega)/B_k} \tag{96}$$

*converges weakly to the measure $\mathscr{N}(0, \zeta(\omega))$.*

The proof will follow closely [26], except that we deal with reverse martingales instead of martingales, which means we have to reverse all the stopping time arguments.

*Proof* Let $M > 1$. Define a stopping time

$$\tau_k = \max\left[1, \sup\{1 \leqslant l \leqslant k : \sum_{j=l}^k Z_j^2 > 2MB_k^2\}\right]. \tag{97}$$

The set $\{\tau_k = j\}$ is $\mathscr{F}_j$-measurable. Let now

$$Z'_{jk} = Z_j \mathbb{1}_{\{\tau_k \leqslant j\}}. \tag{98}$$

Since $\mathbb{1}_{\{\tau_k \leqslant j\}}$ is $\mathscr{F}_{j+1}$-measurable, $\mathbb{E}(Z'_{jk}|\mathscr{F}_{j+1}) = 0$. We will prove that, for almost all $\omega$ with $\zeta(\omega) \leqslant M$,

$$\frac{1}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z'_{jk}/B_k} \to \mathscr{N}(0, \zeta(\omega)). \tag{99}$$

Let us show that this convergence implies the theorem. The difference between $\frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z'_{jk}/B_k}$ and $\frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z_j/B_k}$ has total mass at most

$$\frac{2}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \mathbb{1}_{\{\tau_k>1\}} \leqslant \frac{2}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \mathbb{1}_{\{\sum_{j=1}^k Z_j^2 > 2MB_k^2\}}. \tag{100}$$

Let $f:\mathbb{R}\to\mathbb{R}$ be the piecewise affine function equal to 0 on $\{x\leqslant M\}$ and to 1 on $\{x\geqslant M+1\}$. Then this total mass is at most

$$2\int_{\mathbb{R}} f(x)\,\mathrm{d}\left[\frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z_j^2(\omega)/B_k^2}\right](x). \tag{101}$$

For almost all $\omega$, this quantity converges by assumption to $2\int_{\mathbb{R}} f(x)\,\mathrm{d}[\delta_{\zeta(\omega)}](x)$, which is zero when $\zeta(\omega)\leqslant M$. Hence, the conclusion of the theorem holds for almost all $\omega$ with $\zeta(\omega)\leqslant M$. Taking a sequence $M_n = n$ and since $\zeta$ is finite almost surely, we obtain the full conclusion of the theorem.

So, we just have to prove (99). We will rather prove that, for all $t\geqslant 0$, for almost all $\omega$ with $\zeta(\omega)\leqslant M$,

$$\frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \exp\left(it\sum_{j=1}^k Z'_{jk}(\omega)/B_k\right) \to \exp(-\zeta(\omega)t^2/2) \tag{102}$$

and

$$\int_{|s|\leqslant t} \frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \exp\left(is\sum_{j=1}^k Z'_{jk}(\omega)/B_k\right)\mathrm{d}s \to \int_{|s|\leqslant t}\exp(-\zeta(\omega)s^2/2)\,\mathrm{d}s. \tag{103}$$

By [26; Lemma 6.7], this will imply the desired convergence (99) almost surely. In fact, we will only prove (102), since (103) follows from the same estimates.

There exists a function $r:\mathbb{R}\to\mathbb{C}$ with $|r(x)|\leqslant C|x|^3$ such that

$$\exp(itx) = \exp(-t^2x^2/2)(1+itx)\exp(r(x)). \tag{104}$$

We obtain

$$\exp\left(it\sum_{j=1}^k Z'_{jk}/B_k\right)$$
$$= \exp\left(-t^2\sum_{j=1}^k {Z'_{jk}}^2/2B_k^2\right)\prod_{j=1}^k(1+itZ'_{jk}/B_k)\exp\left(\sum_{j=1}^k r(tZ'_{jk}/B_k)\right).$$

We will denote this last product by $\mathscr{E}_k(t)\Pi_k(t)R_k(t)$. Writing

$$\begin{aligned}\mathscr{E}_k(t)\Pi_k(t)R_k(t) - \exp(-\zeta(\omega)t^2/2) &= (\mathscr{E}_k(t) - \exp(-\zeta(\omega)t^2/2))\Pi_k(t)R_k(t)\\ &\quad + \exp(-\zeta(\omega)t^2/2)\Pi_k(t)(R_k(t)-1)\\ &\quad + \exp(-\zeta(\omega)t^2/2)(\Pi_k(t)-1),\end{aligned}$$

we get

$$\begin{aligned}&\left|\frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \exp\left(it\sum_{j=1}^k Z'_{jk}(\omega)/B_k\right) - \exp(-\zeta(\omega)t^2/2)\right|\\ &\quad\leqslant \frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k|\mathscr{E}_k(t) - \exp(-\zeta(\omega)t^2/2)||\Pi_k(t)R_k(t)|\\ &\quad\quad + \frac{1}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k \exp(-\zeta(\omega)t^2/2)|\Pi_k(t)(R_k(t)-1)|\\ &\quad\quad + \frac{1}{\sum_{k=1}^n b_k}\left|\sum_{k=1}^n b_k \exp(-\zeta(\omega)t^2/2)(\Pi_k(t)-1)\right|. \end{aligned} \tag{105}$$

If we can prove that these three terms tend to 0 for almost all $\omega$ with $\zeta(\omega) \leqslant M$, we will have proved (102) and the proof will be complete.

Write $N_k = \max_{1\leqslant j\leqslant k}|Z_j|$. We have

$$\begin{aligned}\left|\sum_{j=1}^k r(tZ'_{jk}/B_k)\right| &\leqslant \sum_{j=1}^k Ct^3|Z'_{jk}|^3/B_k^3 \leqslant Ct^3\left(Z^2_{\tau(k)} + \sum_{j=\tau(k)+1}^k Z_j^2\right)N_k/B_k^3\\ &\leqslant Ct^3\left(N_k^2 + 2MB_k^2\right)N_k/B_k^3 = Ct^3\left(N_k^2/B_k^2 + 2M\right)N_k/B_k.\end{aligned}$$

For almost all $\omega, N_k/B_k \to 0$ by assumption. Hence, almost surely, $R_k(t)$ tends to 1 (and is in particular bounded). In the same way,

$$\begin{aligned}|\Pi_k(t)|^2 = \prod_{j=1}^k(1 + t^2 {Z'_{jk}}^2/B_k^2) &\leqslant \exp\left(t^2\left(Z^2_{\tau(k)} + \sum_{j=\tau(k)+1}^k Z_j^2\right)/B_k^2\right)\\ &\leqslant \exp(t^2(N_k^2/B_k^2 + 2M)).\end{aligned}$$

Consequently, $\Pi_k(t)$ is almost surely bounded. This proves that the second term in (105) tends almost surely to 0. Moreover, almost surely, the first term in (105) is bounded by $\frac{C}{\sum_{k=1}^n b_k}\sum_{k=1}^n b_k|\mathscr{E}_k(t) - \exp(-\zeta(\omega)t^2/2)|$, which is at most

$$\frac{C}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \left[ \mathbb{1}_{\{\tau_k>1\}} + \left| \exp\left(-t^2 \sum_{j=1}^k Z_j^2/2B_k^2\right) - \exp\left(-\zeta(\omega)t^2/2\right) \right| \right]$$
$$= \frac{C}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \mathbb{1}_{\{\tau_k>1\}}$$
$$+C \int_{\mathbb{R}} |\exp(-xt^2/2) - \exp\left(-\zeta(\omega)t^2/2\right)| \,\mathrm{d}\left[\frac{1}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \delta_{\sum_{j=1}^k Z_j^2/B_k^2}\right](x).$$

We have seen that the first term tends to 0 for almost all $\omega$ such that $\zeta(\omega) \leqslant M$. Moreover, the second term converges almost surely to $\int_{\mathbb{R}} |\exp(-xt^2/2) - \exp(-\zeta(\omega)t^2/2)| \,\mathrm{d}\delta_{\zeta(\omega)}(x) = 0$. This proves that the first term in (105) tends almost surely to 0 on $\{\zeta \leqslant M\}$.

The third term in (105) is more complicated to deal with. Notice that $\mathbb{E}(\Pi_k(t)|\mathscr{F}_2) = \prod_{j=2}^k (1 + itZ'_{jk}/B_k)\mathbb{E}(1 + Z'_{1k}|\mathscr{F}_2) = \prod_{j=2}^k (1 + itZ'_{jk}/B_k)$ since $\mathbb{E}(Z'_{jk}|\mathscr{F}_{j+1}) = 0$. By induction, we get $\mathbb{E}(\Pi_k(t)|\mathscr{F}_m) = \prod_{j=m}^k (1 + itZ'_{jk}/B_k)$. In particular,

$$\mathbb{E}(\Pi_k(t)) = 1. \tag{106}$$

For $l \leqslant k$, let us estimate $\mathbb{E}(\Pi_k(t)\overline{\Pi_l(t)})$. For $p \geqslant 1$, write

$$A_p = \prod_{j=p}^l (1 - itZ'_{jl}/B_l) \prod_{j=p}^k (1 + itZ'_{jk}/B_k). \tag{107}$$

For $p > l$, there is no $Z'_{jl}$ term. In particular, the same argument as above shows that $\mathbb{E}(A_p) = 1$. Consider now $p \leqslant l$. Then

$$\begin{aligned}\mathbb{E}(A_p|\mathscr{F}_{p+1}) &= A_{p+1}\mathbb{E}\big((1 - itZ'_{pl}/B_l)(1 + itZ'_{pk}/B_k)|\mathscr{F}_{p+1}\big)\\ &= A_{p+1}\Big[1 - it\mathbb{E}(Z'_{pl}/B_l|\mathscr{F}_{p+1}) + it\mathbb{E}(Z'_{pk}/B_k|\mathscr{F}_{p+1})\\ &\quad + \frac{t^2}{B_kB_l}\mathbb{E}(Z'_{pk}Z'_{pl}|\mathscr{F}_{p+1})\Big]\\ &= A_{p+1} + A_{p+1}\frac{t^2}{B_kB_l}\mathbb{E}(Z'_{pk}Z'_{pl}|\mathscr{F}_{p+1}).\end{aligned}$$

Taking expectations, we get

$$|\mathbb{E}(A_{p+1}) - \mathbb{E}(A_p)| = \frac{t^2}{B_kB_l}\Big|\mathbb{E}(A_{p+1}\mathbb{1}_{\{\tau_k\leqslant p\}}\mathbb{1}_{\{\tau_l\leqslant p\}}\mathbb{E}(Z_p^2|\mathscr{F}_{p+1}))\Big|. \tag{108}$$

If $\tau_k \leqslant p$ and $\tau_l \leqslant p$, we have

$$
\begin{aligned}
|A_{p+1}|^2 &\leqslant \prod_{j=\tau_k+1}^{k} (1+t^2 Z_j^2/B_k^2) \prod_{j=\tau_l+1}^{l} (1+t^2 Z_j^2/B_l^2) \\
&\leqslant \exp\left(t^2 \sum_{j=\tau_k+1}^{k} Z_j^2/B_k^2\right) \exp\left(t^2 \sum_{j=\tau_l+1}^{l} Z_j^2/B_l^2\right) \\
&\leqslant \exp(2Mt^2)\exp(2Mt^2).
\end{aligned}
$$

Hence,

$$
|\mathbb{E}(A_{p+1}) - \mathbb{E}(A_p)| \leqslant \frac{t^2 \exp(2Mt^2)}{B_k B_l}\, \mathbb{E}(Z_p^2 1\!\!1_{\{\tau_l \leqslant p\}}). \tag{109}
$$

Summing for $p$ from 1 to $l$, we obtain:

$$
\begin{aligned}
|\mathbb{E}(\Pi_k(t)\overline{\Pi_l(t)}) - 1| &\leqslant \frac{t^2 \exp(2Mt^2)}{B_k B_l}\, \mathbb{E}\left(\sum_{p=\tau_l}^{l} Z_p^2\right) \\
&\leqslant \frac{t^2 \exp(2Mt^2)}{B_k B_l}\, \mathbb{E}\left(Z_{\tau_l}^2 + \sum_{p=\tau_l+1}^{l} Z_p^2\right) \\
&\leqslant \frac{t^2 \exp(2Mt^2)}{B_k B_l}\, \mathbb{E}\left(\max_{1\leqslant j\leqslant l} Z_j^2 + 2MB_l^2\right).
\end{aligned}
$$

Since $\mathbb{E}(\max_{1\leqslant j\leqslant l} Z_j^2/B_l^2)$ is uniformly bounded by assumption, we obtain finally:

$$
|\mathbb{E}(\Pi_k(t)\overline{\Pi_l(t)}) - 1| \leqslant C\frac{B_l}{B_k}\cdot \tag{110}
$$

Write $\Pi_n^b(t) = \frac{1}{\sum_{k=1}^n b_k} \sum_{k=1}^n b_k \Pi_k(t)$. Then

$$
\begin{aligned}
\mathbb{E}(|\Pi_n^b(t) - 1|^2) &\leqslant \frac{2}{\left(\sum_{k=1}^n b_k\right)^2} \sum_{k=1}^n b_k \sum_{l=1}^k b_l |\mathbb{E}(\Pi_k(t)\overline{\Pi_l(t)}) - 1| \\
&\leqslant \frac{C}{\left(\sum_{k=1}^n b_k\right)^2} \sum_{k=1}^n b_k \sum_{l=1}^k b_l \frac{B_l}{B_k}\cdot
\end{aligned}
$$

By assumption, $b_l B_l \leqslant C(B_l - B_{l-1})$. Summing from 1 to $k$, we get $\sum_{l=1}^k b_l B_l \leqslant CB_k$. Finally,

$$
\mathbb{E}(|\Pi_n^b(t) - 1|^2) \leqslant \frac{C}{\left(\sum_{k=1}^n b_k\right)^2} \sum_{k=1}^n b_k \frac{B_k}{B_k} \leqslant \frac{C}{\sum_{k=1}^n b_k}\cdot \tag{111}
$$

Since $b_k$ is bounded and $\sum b_k = +\infty$, there exists a sequence $u_n$ such that $\sum_{k=1}^{u_n} b_k - n^2 = O(1)$. Equation (111) shows that $\mathbb{E}(|\Pi^b_{u_n}(t) - 1|^2)$ is summable. In particular, for almost every $\omega$, $\Pi^b_{u_n}(t)$ converges to 1.

Consider now an arbitrary $m$, and choose $n$ with $u_n \leqslant m < u_{n+1}$. Then

$$\Pi^b_m(t) = \frac{\sum_{k=1}^{u_n} b_k}{\sum_{k=1}^{m} b_k} \Pi^b_{u_n}(t) + \frac{1}{\sum_{k=1}^{m} b_k} \sum_{k=u_n+1}^{m} b_k \Pi_k(t). \tag{112}$$

Since $\frac{\sum_{k=1}^{u_n} b_k}{\sum_{k=1}^{m} b_k} \to 1$, $\Pi^b_{u_n}(t) \to 1, \Pi_k(t)$ is bounded and $\frac{\sum_{k=u_n+1}^{m} b_k}{\sum_{k=1}^{m} b_k} \to 0$, this shows that $\Pi^b_m(t)$ converges to 1. Hence, the third term of (105) tends almost surely to 0. This concludes the proof. □

## 8.2 Dynamical application

*Proof of theorem 2.16* Under the assumptions of Gordin's Theorem, there exist two functions $g, h \in L^2$ such that $f = g - g \circ T + h$, and the sequence $h \circ T^n$ is a reverse martingale difference for the filtration $\mathscr{F}_n = T^{-n}\mathscr{F}$, i.e., $h \circ T^n$ is $\mathscr{F}_n$-measurable and $\mathbb{E}(h \circ T^n | \mathscr{F}_{n+1}) = 0$. Moreover, a variance computation using (22) shows that $S_n f/\sqrt{n}$ is bounded in $L^2$.

Let us first show that $S_n f/\sqrt{n}$ has tight maxima. Since $S_n h/\sqrt{n} = S_n f/\sqrt{n} + (g - g \circ T^n)/\sqrt{n}$, the sequence $S_n h/\sqrt{n}$ is bounded in $L^2$, and therefore in $L^1$. Example 2.3 thus shows that $S_n h/\sqrt{n}$ has tight maxima. To conclude the proof, we just have to check that $(g - g \circ T^n)/\sqrt{n}$ has tight maxima. We have

$$\frac{\max_{1 \leqslant k \leqslant n} |g - g \circ T^k|}{\sqrt{n}} \leqslant \frac{|g|}{\sqrt{n}} + \frac{\max_{1 \leqslant k \leqslant n} |g \circ T^k|}{\sqrt{n}}. \tag{113}$$

Moreover, for any $c > 0$,

$$\begin{aligned} \mathbb{P}\left\{ \max_{1 \leqslant k \leqslant n} |g \circ T^k|/\sqrt{n} \geqslant c \right\} &\leqslant \mathbb{P}\left\{ \max_{1 \leqslant k \leqslant n} g^2 \circ T^k/n \geqslant c^2 \right\} \\ &\leqslant c^{-2}\mathbb{E}(S_n g^2/n) = c^{-2}\mathbb{E}(g^2). \end{aligned}$$

Hence, this sequence is also tight. This concludes the proof of the tightness of maxima of $S_n f/\sqrt{n}$.

Let us now turn to the proof of the almost sure central limit theorem. Set $b_k = \frac{1}{k}, B_n = \sqrt{n}, \zeta = \int h^2 \, \mathrm{d}m$ and $Z_n = h \circ T^{n+1}$. We check the assumptions of Theorem 8.1. Birkhoff's ergodic Theorem applied to $h^2$ shows that $h^2 \circ T^n = o(n)$, hence $Z_n/\sqrt{n} \to 0$ almost everywhere. Moreover, $\sum_{k=1}^{n} Z_j^2/n$ tends almost everywhere to $\mathbb{E}(h^2)$, and the second condition of Theorem 8.1 follows. The third condition is trivial. Finally,

$$\mathbb{E}\big[\max_{1\leqslant j\leqslant k}|Z_j|^2/B_k^2\big] \leqslant \mathbb{E}(S_k h^2/k) = \mathbb{E}(h^2) \tag{114}$$

hence this sequence is bounded.

Therefore, Theorem 8.1 applies and proves that, almost everywhere,

$$\frac{1}{\log N}\sum_{k=1}^{N}\frac{1}{k}\delta_{S_k h(x)/\sqrt{k}} \xrightarrow{\text{law}} \mathcal{N}(0,\mathbb{E}(h^2))\,. \tag{115}$$

Moreover, $S_k f = S_k h + g - g\circ T^k$. Again by Birkhoff's ergodic Theorem applied to $g^2$, $(g - g\circ T^k)/\sqrt{k}$ tends almost surely to 0. The result follows. □

**Acknowledgements** We thank the anonymous referees for their very useful comments and suggestions.

## References


1. Aaronson, J., Denker, M.: A local limit theorem for stationary processes in the domain of attraction of a normal distribution. In: Balakrishnan, N., Ibragimov, I.A., Nevzorov, V.B. (eds.) Asymptotic Methods in Probability and Statistics with Applications, papers from the international conference, St. Petersburg, Russia, 1998. pp. 215–224, Birkhäuser (2001)
2. Aaronson, J., Denker, M.: Local limit theorems for partial sums of stationary sequences generated by Gibbs–Markov maps. Stoch. Dyn. **1**, 193–237 (2001)
3. Aaronson, J., Denker, M., Urbański, M.: Ergodic theory for Markov fibred systems and parabolic rational maps. Trans. Am. Math. Soc. **337**, 495–548 (1993)
4. Alves J.F., Luzzatto, S., Pinheiro, V.: Markov structures and decay of correlations for non-uniformly expanding dynamical systems. Ann. Inst. H. Poincaré Anal. Non Linéaire **22**, 817–839 (2005)
5. Bunimovich, L.: On the ergodic properties of nowhere dispersing billiards. Comm. Math. Phys. **65**, 295–312 (1979)
6. Baladi, V., Tsujii, M.: Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms. Ann. Inst. Fourier (in press) (2005)
7. Baladi, V., Young, L.-S.: On the spectra of randomly perturbed expanding maps. Comm. Math. Phys. **156**, 355–385 (1993)
8. Bálint, P., Gouëze, S.: Limit theorems in the stadium billiard. Comm. Math. Phys. **263**, 461–512 (2006)
9. Berkes, I., Csáki, E.: A universal result in almost sure central limit theory. Stoch. Process. Appl. **94**(1), 105–134 (2001)
10. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation. In: Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1987
11. Chazottes, J.-R., Collet, P.: Almost-sure central limit theorems and the Erdös-Rényi law for expanding maps of the interval. Ergodic Theory Dyn. Syst. **25**(2), 419–441 (2005)
12. Chazottes, J.-R., Collet, P., Schmitt, B.: Statistical consequences of Devroye inequality for processes. Applications to a class of non-uniformly hyperbolic dynamical systems. Nonlinearity **18**(5), 2341–2364 (2005)
13. Denker, M., Philipp, W.: Approximation by Brownian motion for Gibbs measures and flows under a function. Ergodic Theory Dyn. Syst. **4**(4), 541–552 (1984)
14. Dolgopyat, D.: Limit theorems for partially hyperbolic systems. Trans. Am. Math. Soc. **356**, 1637–1689 (2004)
15. Eagleson, G.K.: Some simple conditions for limit theorems to be mixing. Teor. Verojatnost. i Primenen. **21**(3), 653–660 (1976)
16. Gordin, M.: The central limit theorem for stationary processes. Dokl. Akad. Nauk SSSR **188**, 739–741 (1969)

17. Gouëzel., S.: Statistical properties of a skew-product with a curve of neutral points. Ergodic Theory Dyn. Syst. (in press) (2003)
18. Gouëzel, S.: Central limit theorem and stable laws for intermittent maps. Probab. Theory Relat. Fields **128**, 82–122 (2004)
19. Gouëzel, S.: Decay of correlations for nonuniformly expanding systems. Bull. Soc. Math. Fr. **134**(1), 1–31 (2006)
20. Gouëzel, S., Liverani, C.: Banach spaces adapted to Anosov systems. Ergodic Theory Dyn. Syst. **26**, 189–217 (2006)
21. Gnedenko, B.V., Kolmogorov, A.N.: Limit distributions for sums of independent random variables. Translated from the Russian, annotated, and revised by K. L. Chung. With appendices by J. L. Doob and P. L. Hsu. Revised edition. Addison-Wesley, Reading, London, Don Mills. (1968)
22. Hervé, L.: Théorème local pour chaînes de Markov de probabilité de transition quasi-compacte. Applications aux chaînes $V$-géométriquement ergodiques et aux modèles itératifs. Ann. Inst. H. Poincaré Probab. Statist. **41**(2), 179–196 (2005)
23. Hennion, H., Hervé, L.: Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness. Lecture Notes in Mathematics, vol. 1766. Springer, Berlin Heidelberg New York (2001)
24. Keller, C., Liverani, C.: Stability of the spectrum for transfer operators. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), **28**(1), 141–152 (1999)
25. Lesigne, E.: Almost sure central limit theorem for strictly stationary processes. Proc. Am. Math. Soc. **128**, 1751–1759 (2000)
26. Lifshits., M.A.: Almost sure limit theorem for martingales. In: Limit theorems in probability and statistics, Balatonlelle, 1999, vol. II, pp. 367–390. János Bolyai Math. Soc., Budapest (2002)
27. Liverani, C.: Central limit theorems for deterministic systems. In: international conference on dynamical systems, Montevideo 1995. Pitman Research Notes in Mathematics, vol. 362 (1996)
28. Lacey, M.T., Philipp, W.: A note on the almost sure central limit theorem. Statist. Probab. Lett. **9**(3), 201–205 (1990)
29. Melbourne, I., Nicol, M.: Almost sure invariance principle for nonuniformly hyperbolic systems. Comm. Math. Phys. **260**(1), 131–146 (2005)
30. Melbourne, I., Török, A.: Statistical limit theorems for suspension flows. Israel J. Math. **144**, 191–210 (2004)
31. Yoshihara, K.-I.: Weakly dependent stochastic sequences and their applications. In: Recent Topics on Weak and Strong Limit Theorems, Vol. XIV. Sanseido Co. Ltd., Chiyoda (2004)
32. Young, L.-S.: Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. **147** (2), 585–650 (1998)
33. Young, L.S.: Recurrence times and rates of mixing. Israel J. Math. **110**, 153–188 (1999)
34. Zweimüller, R.: Stable limits for probability preserving maps with indifferent fixed points. Stoch. Dyn. **3**, 83–99 (2003)