Proceedings of Symposia in Pure Mathematics
Volume **89**, 2015
http://dx.doi.org/10.1090/pspum/089/01487

Limit theorems in dynamical systems using the spectral method

Sébastien Gouëzel

Abstract. There are numerous techniques to prove probabilistic limit theorems for dynamical systems. These notes are devoted to one of these methods, the Nagaev-Guivarc'h spectral method, which extends to dynamical systems the usual proof of the central limit theorem relying on characteristic functions. We start with the simplest example (expanding maps of the interval), where everything is elementary. We then consider more recent (and more involved) applications of this method, on the one hand to get the convergence to stable laws in intermittent maps, on the other hand to obtain precise results on the almost sure approximation by a Brownian motion.

1. Introduction

There are many ways to prove the central limit theorem for square-integrable independent random variables, each of them having various advantages (and weaknesses) that make them generalizable to different situations. While the most versatile approaches are probably those relying on martingale arguments, this text is devoted to the approach that is generally used in first-year probability courses, relying on characteristic functions. Our goal in this text is to illustrate its effectiveness to prove the central limit theorem, or more general limit theorems, in deterministic dynamical systems. This powerful method, which we will call the Nagaev-Guivarc'h spectral method, was devised by Nagaev to study Markov chains [**Nag57**, **Nag61**], and reinvented by Guivarc'h for dynamical systems [**RE83**, **GH88**]. An excellent reference on this method is [**HH01**].

We start with the classical proof of the central limit theorem (see for instance [**Fel66**, Theorem XV.5.1]), that we will revisit later in dynamical situations.

Theorem 1.1. *Let X_i be a sequence of i.i.d. centered random variables in L^2. Write $S_n = X_1 + \cdots + X_n$. Then $S_n/\sqrt{n}$ converges in distribution to a Gaussian random variable $\mathcal{N}(0, \sigma^2)$, where $\sigma^2 = \mathbb{E}(X_i^2)$.*

Proof. Thanks to the independence, we can compute the characteristic function of S_n: for any real t,

$$\mathbb{E}(e^{\mathbf{i}tS_n}) = \lambda(t)^n, \tag{1.1}$$

2010 *Mathematics Subject Classification.* Primary 37A50; Secondary 37A30, 37D25.

Key words and phrases. Transfer operator, spectral method, intermittent map, almost sure invariance principle.

where $\lambda(t) = \mathbb{E}(e^{\mathbf{i}tX_1})$. Since X_1 is square integrable, one has for small t the asymptotics

$$\lambda(t) = 1 + \mathbf{i}t\mathbb{E}(X_1) - \frac{t^2}{2}\mathbb{E}(X_1^2) + o(t^2) = 1 - \sigma^2 t^2/2 + o(t^2). \tag{1.2}$$

Combining those two equations, we get for every fixed $t \in \mathbb{R}$

$$\mathbb{E}(e^{\mathbf{i}tS_n/\sqrt{n}}) = \bigl(1 - \sigma^2 t^2/(2n) + o(1/n)\bigr)^n \to e^{-\sigma^2 t^2/2}.$$

The function $e^{-\sigma^2 t^2/2}$ is the characteristic function of the Gaussian distribution $\mathcal{N}(0, \sigma^2)$. As pointwise convergence of characteristic functions implies convergence in distribution (this is Lévy's theorem, see [**Fel66**, Theorem XV.3.2]), the result follows. □

There are two main points in the above proof:

- First, in (1.1), one uses independence to write $\mathbb{E}(e^{\mathbf{i}tS_n})$ as $\lambda(t)^n$, for some function $\lambda(t)$ (here, it is simply $\mathbb{E}(e^{\mathbf{i}tX_1})$).
- Then, in (1.2), one finds the asymptotic behavior of $\lambda(t)$ for small t.

The convergence in distribution of $S_n/\sqrt{n}$ follows.

One can abstract the above proof, to get the following statement:

THEOREM 1.2 (Nagaev-Guivarc'h method, naive version). *Let $X_1, X_2, \dots$ be a sequence of real random variables. Write S_n for their partial sums. Assume that there exist $\delta > 0$ and functions $c(t)$, $\lambda(t)$ and $d_n(t)$, defined on $[-\delta, \delta]$, such that for all $t \in [-\delta, \delta]$ and all $n \in \mathbb{N}$,*

$$\mathbb{E}(e^{\mathbf{i}tS_n}) = c(t)\lambda(t)^n + d_n(t). \tag{1.3}$$

Moreover, assume that:

1. *there exist A and σ^2 in $\mathbb{C}$ such that $\lambda(t) = \exp(\mathbf{i}At - \sigma^2 t^2/2 + o(t^2))$ when $t \to 0$;*
2. *the function c is continuous at 0;*
3. *the quantity $\|d_n\|_{L^\infty[-\delta,\delta]}$ tends to 0 when n tends to infinity.*

Then $A \in \mathbb{R}$, $\sigma^2 \geqslant 0$, and $(S_n - nA)/\sqrt{n}$ converges to a Gaussian distribution $\mathcal{N}(0, \sigma^2)$ when n tends to infinity.

PROOF. The proof is essentially the same. First, taking $t = 0$, one gets $c(0)\lambda(0)^n = 1 - d_n(0) \to 1$, hence $c(0) = 1$ and $\lambda(0) = 1$. Then, one uses the asymptotic expansion of λ to show that $\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/n})$ tends to 1. By Lévy's theorem, $S_n/n - A$ converges in distribution to 0. As S_n is real, it follows that $A \in \mathbb{R}$. Using again the asymptotic expansion of λ, we show that $\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/\sqrt{n}})$ converges to the function $e^{-\sigma^2 t^2/2}$. By Lévy's theorem again, this has to be the characteristic function of a real random variable, showing that $\sigma^2 \geqslant 0$. □

As such, the method does not seem of great interest: in concrete non-independent situations, how could one effectively construct functions c, λ and d_n that satisfy (1.3)? The main hindsight is that, in various contexts, these quantities will come for free from spectral arguments: in particular, $\lambda(t)$ will correspond to an eigenvalue of an operator depending on t.

Once spectral theory tools are available, one can prove the expansion (1.3) in several situations, that we will review in this paper. As we explained above, this easily implies the central limit theorem. However, it has much deeper consequences:

almost any proof in probability theory that relies on characteristic functions can be adapted to this setting. In a dynamical system setting, one can therefore prove, among other results (under suitable assumptions):

- the Berry-Esseen theorem, i.e., control the speed of convergence in the central limit theorem (see [**GH88**] or Theorem 3.7 below). This kind of estimates can sometimes also be proved using martingale arguments. However, with the Nagaev-Guivarc'h method, one can *characterize* the functions for which there is a given speed of convergence in the central limit theorem [**Gou10b**].
- the local limit theorem, i.e., control the probability that S_n belongs to a given interval of size of order 1 (while the central limit theorem deals with intervals of size of order $\sqrt{n}$). See for instance [**GH88**]. Note that this kind of fine control is impossible to prove using martingale arguments.
- deduce the local limit theorem from the central limit theorem [**Her05**].
- the convergence to other limit laws, for instance stable laws, see [**AD01**]. These results will be explained in Section 4, together with applications to intermittent maps.
- the vector-valued almost sure invariance principle [**Gou10a**], see Section 5 below.

One expects that virtually any result known for i.i.d. sequences should hold for dependent random variables if one can prove that (1.3) holds.

In this article, we concentrate on the applications of the Nagaev-Guivarc'h method to dynamical systems. Let us note, however, that the method applies equally well to Markov chains, and that most of the above statements have counterparts for Markov chains. In this respect, see [**HP10**] and references therein.

In the next section, we will give the precise statements of spectral theory that we will use all over this paper. Then, in Section 3, we will give full self-contained proofs of how the method can be used in the simplest situation, expanding maps of the interval (in this situation, there are numerous proofs of the central limit theorem, so this section is really meant as an illustration of how the method works, not of his full power). Sections 4 and 5 are then devoted to more complicated applications, for which the Nagaev-Guivarc'h method gives the best known results.

2. A bit of spectral theory

In this paragraph, we give a brief overview of spectral theory, or more precisely of the spectral theory we will need. A major reference on this topic is [**Kat66**].

Let $(\mathcal{B}, \|\cdot\|)$ be a complex Banach space, and let $\mathcal{L}$ be a continuous linear operator on $\mathcal{B}$ (i.e., there exists a constant C such that $\|\mathcal{L}u\| \leqslant C\|u\|$ for all $u \in \mathcal{B}$). The *spectrum* of $\mathcal{L}$, denoted by $\sigma(\mathcal{L})$, is the set of complex numbers z such that $zI - \mathcal{L}$ is not invertible.

Example 2.1 (An example to keep in mind). Let $\mathcal{B}$ be the set of continuous functions $f : [0,1] \to \mathbb{C}$ with $f(0) = 0$. Define an operator $\mathcal{L} : \mathcal{B} \to \mathcal{B}$ by $\mathcal{L}f(x) = \int_0^x f(y)\,\mathrm{d}y$. It is a continuous linear operator. If $\mathcal{L}f = 0$, then $f = 0$ (just differentiate), hence $\mathcal{L}$ is injective. However, it is not surjective, since its image is made of C^1 functions. Hence, $0 \in \sigma(\mathcal{L})$. For any $z \neq 0$, the operator $zI - \mathcal{L}$ is invertible (one should just solve a differential equation to invert it), hence

$\sigma(\mathcal{L}) = \{0\}$. Note that $\mathcal{L}$ has no nonzero eigenfunction at all. This shows that the behavior of operators can be very different from the finite-dimensional situation.

Let z be an isolated point in $\sigma(\mathcal{L})$. The corresponding spectral projection Π_z is defined by $\Pi_z \coloneqq \frac{1}{2i\pi}\int_{\mathcal{C}}(wI - \mathcal{L})^{-1}\,\mathrm{d}w$, where $\mathcal{C}$ is a small circle around z. This definition is independent from the choice of $\mathcal{C}$, by holomorphy of $w \mapsto (wI - \mathcal{L})^{-1}$ outside of $\sigma(\mathcal{L})$. When $\mathcal{L}$ is finite-dimensional, it is easy to check (for instance by putting $\mathcal{L}$ in upper-triangular form and using the Cauchy formula for integrals) that Π_z is the projection on the generalized eigenspace associated to z, with kernel the direct sum of the other generalized eigenspaces. In infinite dimension, an analogous result is true: the operator Π_z is a projection, its image and kernel are invariant under $\mathcal{L}$, and the spectrum of the restriction of $\mathcal{L}$ to the image is $\{z\}$, while the spectrum of the restriction of $\mathcal{L}$ to the kernel if $\sigma(\mathcal{L}) - \{z\}$. See [**Kat66**, Theorem III.6.17].

We say that $z \in \mathbb{C}$ is an *isolated eigenvalue of finite multiplicity* of $\mathcal{L}$ if z is an isolated point of $\sigma(\mathcal{L})$, and the range of Π_z is finite-dimensional. In this case, $\ker(zI - \mathcal{L})^j$ is independent of j for large enough j, and coincides with $\operatorname{Im}\Pi_z$. Note that the converse is not true: in Example 2.1, we have $\ker(0I - \mathcal{L})^j = \{0\}$ for all j, but 0 is *not* an isolated eigenvalue of finite multiplicity (the spectral projection Π_0 is the identity). If z is an isolated eigenvalue of finite multiplicity of an operator $\mathcal{L}$, its *multiplicity* is the dimension of the range of Π_z.

We denote by $\sigma_{\mathrm{ess}}(\mathcal{L})$ the *essential spectrum* of $\mathcal{L}$, i.e., the set of points in $\sigma(\mathcal{L})$ that are not isolated eigenvalues of finite multiplicity. One can think of $\mathcal{L}$ as a finite matrix outside of $\sigma_{\mathrm{ess}}(\mathcal{L})$, and as something more complicated on $\sigma_{\mathrm{ess}}(\mathcal{L})$. This will make it possible to understand the asymptotics of $\mathcal{L}^n$ if the dominating elements of the spectrum (i.e., the points in $\sigma(\mathcal{L})$ with large modulus) do not belong to $\sigma_{\mathrm{ess}}(\mathcal{L})$. We say that such an operator is *quasi-compact*, or that it has a *spectral gap*.

The *spectral radius* $r(\mathcal{L})$ of $\mathcal{L}$ is $\sup\{|z| \,:\, z \in \sigma(\mathcal{L})\}$, and the *essential spectral radius* $r_{\mathrm{ess}}(\mathcal{L})$ is $\sup\{|z| \,:\, z \in \sigma_{\mathrm{ess}}(\mathcal{L})\}$. An operator is quasi-compact if $r_{\mathrm{ess}}(\mathcal{L}) < r(\mathcal{L})$. These quantities can be computed as follows: $r(\mathcal{L}) = \inf_{n>0}\|\mathcal{L}^n\|^{1/n}$, and

$$r_{\mathrm{ess}}(\mathcal{L}) = \inf\|\mathcal{L}^n - K\|^{1/n}, \tag{2.1}$$

where the infimum is over all integers $n > 0$ and all compact operators K. In particular, we get that the essential spectral radius of a compact operator vanishes. This corresponds to the classical fact that the spectrum of a compact operator is a sequence of eigenvalues of finite multiplicity tending to 0. The above formula also shows that the essential spectral radius is not altered by the addition of a compact operator. Intuitively, adding a compact operator only amounts to adding (or perturbing) isolated eigenvalues of finite multiplicity.

To estimate the essential spectral radius of an operator, the formula (2.1) is not very convenient, since one should exhibit good compact operators to use it. A more efficient technique relies on inequalities named in the literature after Doeblin-Fortet or after Lasota-Yorke, that we will call DFLY inequalities. This technique is formalized in the following lemma, essentially due to [**Hen93**] (the following formulation can be found in [**BGK07**]).

LEMMA 2.2. *Consider a continuous linear operator $\mathcal{L}$ on a complex Banach space $\mathcal{B}$. Let $M > 0$. Suppose that, for some $n > 0$ and for all $x \in \mathcal{B}$,*

$$\|\mathcal{L}^n x\| \leqslant M^n\|x\| + \|x\|_w, \tag{2.2}$$

where $\|\cdot\|_w$ is a seminorm on $\mathcal{B}$ such that the unit ball of $\mathcal{B}$ (for $\|\cdot\|$) is relatively compact for $\|\cdot\|_w$. Then $r_{\mathrm{ess}}(\mathcal{L}) \leqslant M$.

If one could decompose $\mathcal{L}$ as a sum of two operators $\mathcal{L}_1 + \mathcal{L}_2$ bounded respectively by the first and second term of (2.2), the operator $\mathcal{L}_1$ would have a spectral radius (and therefore an essential spectral radius) at most M, while $\mathcal{L}_2$ would be compact. Hence, (2.1) would imply that the essential spectral radius of $\mathcal{L}$ would be at most M, giving the claim of the lemma. This heuristic argument motivates the lemma, but the rigorous proof is completely different.

We will also need to describe the evolution of the spectrum when one perturbs operators. Consider a family of operators $t \mapsto \mathcal{L}_t$, depending continuously on t (i.e., $\|\mathcal{L}_t - \mathcal{L}_s\|$ tends to 0 when t tends to s). There are such situations where the spectrum varies discontinuously: one can for instance have $\sigma(\mathcal{L}_0)$ equal to the unit disk in $\mathbb{C}$, and $\sigma(\mathcal{L}_t)$ equal to the unit circle for $t \neq 0$. However, the situation is better for eigenvalues of finite multiplicity (where everything can be reduced to finite-dimensional arguments), and even better for eigenvalues of multiplicity one. We only give a precise statement for this case, since this is what we will need later on, see [**Kat66**, IV.3.6 and Theorem VII.1.8] for more general statements.

PROPOSITION 2.3. *Let z_0 be an isolated eigenvalue of multiplicity one of an operator $\mathcal{L}_0$. Then any operator $\mathcal{L}$ close enough to $\mathcal{L}_0$ has a unique eigenvalue $z(\mathcal{L})$ close to z_0. Moreover, if $t \mapsto \mathcal{L}_t$ is a family that depends on t in a C^0, or C^k, or analytic way, then the eigenvalue z_t and the corresponding eigenprojection Π_t and eigenvector ξ_t depend on t in the same way.*

Assume moreover that the rest of the spectrum of $\mathcal{L}_0$ is contained in a disk of strictly smaller radius $B(0, |z_0| - \varepsilon)$. Write $Q_t = (I - \Pi_t)\mathcal{L}_t$ for the part of $\mathcal{L}_t$ corresponding to $\sigma(\mathcal{L}_t) - \{z_t\}$, so that $\mathcal{L}_t = z_t\Pi_t + Q_t$. For any $r > |z_0| - \varepsilon$, these operators satisfy $\|Q_t^n\| \leqslant Cr^n$, for small enough t and for all $n \in \mathbb{N}$, where the constant C is independent of t and n.

We can now reformulate the Nagaev-Guivarc'h spectral method of Theorem 1.2 in a spectral setting:

THEOREM 2.4 (Nagaev-Guivarc'h method, spectral version). *Let $X_1, X_2, \dots$ be a sequence of real random variables, with partial sums denoted by S_n. Assume that there exist a complex Banach space $\mathcal{B}$ and a family of operators $\mathcal{L}_t$ acting on $\mathcal{B}$ (for $|t| \leqslant \delta$) and $\xi \in \mathcal{B}$, $\nu \in \mathcal{B}^*$ such that:*

(1) *coding: for all $n \in \mathbb{N}$, for all $|t| \leqslant \delta$, $\mathbb{E}(e^{\mathbf{i}tS_n}) = \langle \nu, \mathcal{L}_t^n \xi \rangle$.*
(2) *spectral description: $r_{\mathrm{ess}}(\mathcal{L}_0) < 1$, and $\mathcal{L}_0$ has a single eigenvalue of modulus $\geqslant 1$, located at 1. It is an isolated eigenvalue, of multiplicity one.*
(3) *regularity: The family $t \mapsto \mathcal{L}_t$ is C^2.*

Then there exist $A \in \mathbb{R}$ and $\sigma^2 \geqslant 0$ such that $(S_n - nA)/\sqrt{n}$ converges in distribution to a Gaussian $\mathcal{N}(0, \sigma^2)$.

PROOF. The first part of Proposition 2.3 ensures that, for small enough t, the operator $\mathcal{L}_t$ has a unique eigenvalue $\lambda(t)$ close to 1. Moreover, the second part of this proposition ensures that the rest of the spectrum of $\mathcal{L}_t$ is contained in a disk of radius $r < 1$ (with uniform bounds for the iterates of the restricted operators). We get

$$\mathcal{L}_t^n = \lambda(t)^n \Pi_t + Q_t^n,$$

where Π_t is the eigenprojection of $\mathcal{L}_t$ corresponding to the eigenvalue $\lambda(t)$, and $Q_t = (I-\Pi_t)\mathcal{L}_t$ satisfies $\|Q_t^n\| \leqslant Cr^n$. It follows that $\mathbb{E}(e^{\mathbf{i}tS_n}) = \lambda(t)^n\langle \nu, \Pi_t\xi\rangle + \langle \nu, Q_t^n\xi\rangle$. This is a decomposition of the form $\lambda(t)^n c(t) + d_n(t)$, where d_n tends to 0 with n, and c is continuous at 0. To apply Theorem 1.2, it remains to see that the function $\lambda(t)$ has an asymptotic expansion of the form $e^{\mathbf{i}At-\sigma^2t^2/2+o(t^2)}$ for small t. This follows from the fact that this function is C^2, by Proposition 2.3 and the regularity assumption. □

In a lot of applications of the above theorem, the space $\mathcal{B}$ will be a space of functions, ξ will be the function 1, and ν the integration with respect to a fixed measure. In particular, the sequence X_n belongs to the space $\mathcal{B}$ (see for instance the illustration with one-dimensional expanding maps, in Section 3). However, there are more exotic applications, such as [**BGK07**]. In this article, the method is applied to a lattice of coupled expanding maps: the Banach space $\mathcal{B}$ is a set of "projective limits of sequences of compatible measures on increasing sequences of boxes" (it is therefore very far from a space of functions, or even from a space of distributions). This justifies the seemingly abstract formulation of Theorem 2.4, where we have not insisted that the Banach space $\mathcal{B}$ should be related in any sense to the sequence $X_1, X_2, \dots$, except for the relation $\mathbb{E}(e^{\mathbf{i}tS_n}) = \langle \nu, \mathcal{L}_t^n\xi\rangle$.

3. The Nagaev-Guivarc'h spectral method for expanding maps of the interval

In this paragraph, we illustrate how Theorem 2.4 can be applied to prove a central limit theorem, in one of the simplest possible situations: uniformly expanding maps of the interval which are piecewise onto.

We consider a map $T : I \to I$, where $I = [0,1]$ is written as the union of two disjoint intervals $I_1 = [0,a)$ and $I_2 = [a,1]$. We assume that the restriction of T to I_i admits a C^2 extension to $\overline{I_i}$, which is a diffeomorphism between $\overline{I_i}$ and I, and satisfies $T' \geqslant \alpha > 1$. Since the boundaries will not play any role in what follows, we will abusively write I_i instead of $\overline{I_i}$.

THEOREM 3.1. *Let $f : I \to \mathbb{R}$ be a C^1 function. Write $S_nf = \sum_{k=0}^{n-1} f\circ T^k$. There exist $A \in \mathbb{R}$ and $\sigma^2 \geqslant 0$ such that, on the probability space* (I, Leb), *the sequence $(S_nf - nA)/\sqrt{n}$ converges in distribution to a Gaussian $\mathcal{N}(0,\sigma^2)$.*

There are a lot of comments to be made about this theorem, various methods of proof, various extensions, and so on. Since the main emphasis here is to show a simple application of the spectral method, we first start with the proof, and will give the comments afterwards.

To apply Theorem 2.4, we should introduce an operator $\mathcal{L} = \mathcal{L}_0$ acting on a Banach space, related to the composition with T, and with good spectral properties (it should be quasi-compact). The first idea is to use the composition with T (also called the Koopman operator) but there are difficulties to do so, since $f\circ T^n$ is usually much wilder than f (for instance, its derivative is $\geqslant C\alpha^n \to \infty$). These difficulties can be resolved by working in distribution spaces, but it is more elementary (and, often, more efficient) to consider the dual of the Koopman operator, and work with smooth functions. Let $h_i : I \to I_i$ be the inverse of $T_{|I_i}$. We define an operator $\mathcal{L}$ by

$$\mathcal{L}u(x) = \sum_i h_i'(x) u(h_i x) = \sum_{T(y)=x} \frac{1}{T'(y)} u(y). \tag{3.1}$$

This operator satisfies

$$\begin{aligned}\int \mathcal{L}u \cdot v \,\mathrm{dLeb} &= \sum_i \int_I h_i'(x) u(h_i x) v(x) \,\mathrm{d}x = \sum_i \int_{I_i} u(y) v(Ty) \,\mathrm{d}y \\ &= \int u \cdot v \circ T \,\mathrm{dLeb}.\end{aligned} \tag{3.2}$$

Hence, $\mathcal{L}$ is the adjoint of the Koopman operator, as desired. It is called the *transfer operator*, or the *Ruelle-Perron-Frobenius operator*.

Let us now define perturbed transfer operators, as follows. Let f be a C^1 function for which we want to prove a central limit theorem. We define

$$\mathcal{L}_t u = \mathcal{L}(e^{\mathbf{i}tf} u).$$

LEMMA 3.2. *The operators $\mathcal{L}_t$ satisfy the identity*

$$\int \mathcal{L}_t^n u \cdot v = \int u \cdot e^{\mathbf{i}tS_n f} \cdot v \circ T^n \,\mathrm{dLeb}.$$

PROOF. We start from the right hand side, and use the duality property (3.2):

$$\begin{aligned}\int u \cdot e^{\mathbf{i}tS_n f} \cdot v \circ T^n \,\mathrm{dLeb} &= \int (ue^{\mathbf{i}tf}) \cdot (e^{\mathbf{i}tS_{n-1} f} v \circ T^{n-1}) \circ T \,\mathrm{dLeb} \\ = \int \mathcal{L}(ue^{\mathbf{i}tf}) \cdot e^{\mathbf{i}tS_{n-1} f} v \circ T^{n-1} \,\mathrm{dLeb} &= \int \mathcal{L}_t(u) \cdot e^{\mathbf{i}tS_{n-1} f} v \circ T^{n-1} \,\mathrm{dLeb}.\end{aligned}$$

Therefore, an induction gives the result of the lemma. □

In particular, $\mathbb{E}(e^{\mathbf{i}tS_n f}) = \int \mathcal{L}_t^n 1 \cdot 1 \,\mathrm{dLeb}$. This shows that the coding assumption of Theorem 2.4 is satisfied, taking for ξ the constant function 1, and for ν the integration with respect to Lebesgue measure. Note that we have not yet specified the Banach space $\mathcal{B}$ on which the operators $\mathcal{L}_t$ will act. One could try to use the spaces L^∞ or L^2, but there would be no quasi-compactness. We will rather use $\mathcal{B} = C^1$. Note that $t \mapsto e^{\mathbf{i}tf}$ is analytic from $\mathbb{R}$ to C^1 (just use the series expansion $e^{\mathbf{i}tf} = \sum (\mathbf{i}tf)^k / k!$), hence $t \mapsto \mathcal{L}_t$ is analytic, and in particular C^2.

As is often the case, the only difficulty to apply Theorem 2.4 is the quasi-compactness assumption. This is where the assumption of uniform expansion $T' \geqslant \alpha > 1$ will play a role (without this assumption, all we have said until now remains true, but the theorem is false even if $T'(x_0) = 1$ at a single point x_0, as we will see later in Section 4).

The main point is that the iterates of T have a small distortion, i.e., if a set has relatively small measure then its images also have relatively small measure. This is the content of the following technical lemma. For $\underline{i} = (i_1, \dots, i_n) \in \{1, 2\}^n$, let $h_{\underline{i}} = h_{i_1} \circ \dots \circ h_{i_n}$. These functions are the inverse branches of T^n.

LEMMA 3.3. *There exists C such that, for all n, for all $\underline{i}$ of length n,*

$$|h_{\underline{i}}''(x)| \leqslant C h_{\underline{i}}'(x).$$

This lemma shows that $(\log h'_{\underline{i}})'$ is uniformly bounded. Hence, $|\log h'_{\underline{i}}(x) - \log h'_{\underline{i}}(y)| \leqslant C|x-y|$. In particular, $h'_{\underline{i}}(x)/h'_{\underline{i}}(y)$ is bounded away from zero and infinity. This implies that $\operatorname{Leb}(h_{\underline{i}}U)/\operatorname{Leb}(h_{\underline{i}}V)$ is equal to $\operatorname{Leb}(U)/\operatorname{Leb}(V)$ up to a uniform multiplicative constant. This justifies the affirmation that the lemma proves a bounded distortion property of T^n.

Proof. Writing the formula for the derivative of a composition, and taking the logarithm, we get

$$\log(h'_{\underline{i}}) = \sum_{k=1}^{n} (\log h'_{i_k}) \circ h_{i_{k+1}} \circ \cdots \circ h_{i_n}.$$

We differentiate again this equality, getting an expression for $h''_{\underline{i}}/h'_{\underline{i}}$. On the right hand side, the derivative of the k-th term has a factor h''_{i_k}/h'_{i_k} (bounded by a constant C), multiplied by derivatives of the functions h'_{i_j} for $j > k$. Each of these is $\leqslant \alpha^{-1} < 1$, hence we get

$$|h''_{\underline{i}}/h'_{\underline{i}}| \leqslant \sum_{k=1}^{n} C\alpha^{-(n-k)} \leqslant C'. \qquad \square$$

Corollary 3.4. *There exists a constant C such that, for all $n \in \mathbb{N}$, for all C^1 function u,*

$$\|\mathcal{L}^n u\|_{C^1} \leqslant C\alpha^{-n}\|u\|_{C^1} + C\|u\|_{C^0}.$$

Proof. We have $\mathcal{L}^n u = \sum h'_{\underline{i}}(x) u \circ h_{\underline{i}}(x)$, where the sum is over all $\underline{i}$ of length n. Writing $I_{\underline{i}} = h_{\underline{i}}(I)$, the bounded distortion lemma implies that $|h'_{\underline{i}}(x)| \leqslant C \operatorname{Leb}(I_{\underline{i}})$. Therefore,

$$|\mathcal{L}^n u(x)| \leqslant C \sum \operatorname{Leb}(I_{\underline{i}})\|u\|_{C^0} = C\|u\|_{C^0},$$

as $\sum \operatorname{Leb}(I_{\underline{i}}) = \operatorname{Leb}(I) = 1$.

Let us now control the derivative of $\mathcal{L}^n u$, i.e.,

$$\sum h''_{\underline{i}}(x) u \circ h_{\underline{i}}(x) + \sum h'_{\underline{i}}(x) u' \circ h_{\underline{i}}(x) h'_{\underline{i}}(x).$$

In the first term, $|h''_{\underline{i}}| \leqslant C|h'_{\underline{i}}|$ by Lemma 3.3. Hence, this term is bounded by $C\|u\|_{C^0}$ as above. In the second term, we bound the last $h'_{\underline{i}}(x)$ by α^{-n}, and the remaining part is bounded by $C\|u'\|_{C^0}$ as above. $\square$

The above corollary is a DFLY inequality. Together with Lemma 2.2 and the compactness of the inclusion of C^1 in C^0, this implies that the essential spectral radius of $\mathcal{L}$ acting on C^1 is $\leqslant \alpha^{-1} < 1$. Let us now control the outer spectrum of $\mathcal{L}$.

Lemma 3.5. *The operator $\mathcal{L}$ acting on C^1 has a simple eigenvalue at 1, and no other eigenvalue of modulus $\geqslant 1$.*

With this lemma, we can conclude the proof of Theorem 3.1 by applying Theorem 2.4 (the spectral version of Nagaev-Guivarc'h argument). Indeed, we have checked all its assumptions: the coding is proved in Lemma 3.2, the spectral description follows from Lemma 3.5, and the smoothness of $t \mapsto \mathcal{L}_t$ is trivial since this family is analytic as we explained above.

It remains to prove Lemma 3.5.

PROOF OF LEMMA 3.5. The iterates of $\mathcal{L}$ have a uniformly bounded norm, by Corollary 3.4. This shows that $\mathcal{L}$ has no eigenvalue of modulus > 1. Moreover, if there are eigenvalues of modulus 1, they are semisimple (i.e., there are no Jordan blocks). To conclude, one should control these (finitely many) eigenvalues and the corresponding eigenfunctions.

For ρ of modulus 1, denote by Π_ρ the corresponding spectral projection (it vanishes if ρ is not an eigenvalue of $\mathcal{L}$). We have

$$\frac{1}{n}\sum_{k=0}^{n-1} \mathcal{L}^k 1 = \frac{1}{n}\sum_{\rho}\sum_{k=0}^{n-1} \rho^k \Pi_\rho 1 + O(1/n)$$

(where the sum over ρ involves only finitely many terms, the eigenvalues of modulus 1 of $\mathcal{L}$). This converges in C^1 to $\Pi_1 1$ (since for $\rho \neq 1$ the sequence $\sum_{k=0}^{n-1} \rho^k$ is bounded). By bounded distortion, the function $\mathcal{L}^k 1$ takes its values in an interval $[C^{-1}, C]$ for some $C > 0$. It follows that the function $\varphi = \Pi_1 1$ also takes its values in $[C^{-1}, C]$. In particular, Π_1 is not zero, i.e., 1 is an eigenvalue, and φ is a corresponding eigenfunction.

It remains to see that any other eigenfunction is in fact proportional to φ. Consider an eigenfunction ψ, for an eigenvalue ρ of modulus 1. Consider x such that $M = |\psi(x)/\varphi(x)|$ is maximal (since $\varphi \geqslant C^{-1}$, this is well defined). Then

$$\begin{aligned} M\varphi(x) = |\psi(x)| = |\mathcal{L}^n \psi(x)| = \left|\sum h'_{\underline{i}}(x)\psi(h_{\underline{i}}x)\right| \\ \leqslant \sum h'_{\underline{i}}(x)|\psi|(h_{\underline{i}}x) \leqslant \sum h'_{\underline{i}}(x) M\varphi(h_{\underline{i}}x) = M\mathcal{L}^n\varphi(x) = M\varphi(x). \end{aligned}$$

It follows that all these inequalities are equalities. In particular, all the complex numbers $\psi(h_{\underline{i}}x)$ have the same phase, and moreover $|\psi|(h_{\underline{i}}x) = M\varphi(h_{\underline{i}}x)$. This shows that ψ/φ is constant on $T^{-n}(x)$. As this set becomes more and more dense when n tends to infinity, this shows that ψ/φ is constant, as desired. □

Let us give some comments about Theorem 3.1 and its proof:

(1) In the course of the proof, we have constructed an eigenfunction φ of the transfer operator $\mathcal{L}$. One can check that φ is the density of an invariant probability measure m (i.e., it satisfies $m(T^{-1}B) = m(B)$), called the SRB measure, or physical measure, of the system. Moreover, this measure m is ergodic (i.e., if $T^{-1}B = B$, then $m(B) \in \{0, 1\}$): this follows from the fact that the eigenvalue 1 of $\mathcal{L}$ is simple.

(2) One can identify the quantities A and σ^2 in the statement of the theorem: one has $A = \int f \,\mathrm{d}m$, and writing $\tilde{f} = f - \int f \,\mathrm{d}m$,

$$\sigma^2 = \int \tilde{f}^2 \,\mathrm{d}m + 2\sum_{k=1}^{\infty} \int \tilde{f}\cdot\tilde{f}\circ T^k \,\mathrm{d}m. \tag{3.3}$$

Moreover, this series is converging, since the correlations $\int \tilde{f}\cdot\tilde{f}\circ T^k \,\mathrm{d}m$ tend to zero exponentially fast. This again follows from the spectral description of $\mathcal{L}$ (here, what matters is that, apart from the eigenvalue 1, the spectrum of $\mathcal{L}$ is contained in a disk of radius < 1).

(3) The previous items show that the natural measure for the system is not Lebesgue measure, but m. Hence, a central limit theorem on (I, m) would be more natural than on (I, Leb). It turns out that these statements are equivalent, see Theorem 4.9 below.

We have written the proof for $\mathcal{B}$ the space of C^1 functions, hence it only applies if f is C^1. However, essentially the same proof works for Hölder functions (the operator $\mathcal{L}$ is still quasi-compact on the space of Hölder functions). More importantly, the proof also applies to much more general dynamics if they retain uniform expansion. Let us give a (non-exhaustive) list of examples that can be treated using such methods. In most cases, the difficulty is to find a space on which the unperturbed transfer operator $\mathcal{L}$ has a spectral gap.

- Subshifts of finite type with a Gibbs measure [**GH88**, **PP90**].
- Anosov or Axiom A systems. They can be reduced to subshifts of finite type using coding, but one may also work directly with spaces of distributions [**BT07**, **GL08**] (these distributions should be smooth in the unstable direction and dual of smooth in the stable direction).
- Piecewise expanding maps in dimension 1, using spaces of bounded variation functions or variations around this idea [**HK82**, **Kel85**].
- Piecewise expanding maps in higher dimension, if the expansion dominates the local complexity in the sense of [**Buz97**]. See for instance [**Sau00**].
- Lattices of weakly coupled expanding maps [**BGK07**].
- Non-uniformly hyperbolic maps for which the time to wait before seeing the hyperbolicity has exponentially small tails (including for instance billiards, or unimodal maps of the interval of Collet-Eckmann type, or some Hénon maps), see [**You98**].
- Billiard maps [**DZ11**].
- Time-one maps of contact Anosov flows [**Tsu10**, **Tsu12**]. Since there is no hyperbolicity in the direction of the flow, the mixing in this direction should come from a different mechanism related to the non-integrability coming from the contact structure.

To conclude this section, let us describe a strengthening of Theorem 3.1, by estimating the speed of convergence in the central limit theorem:

Theorem 3.6. *Under the assumptions of Theorem* 3.1*, assume moreover that* $\sigma^2 > 0$*. Then there exists* $C > 0$ *such that, for any* $n > 0$*, for any real interval* J*,*

$$\left|\mu\{x \,:\, (S_n f(x) - nA)/\sqrt{n} \in J\} - \mathbb{P}(\mathcal{N}(0, \sigma^2) \in J)\right| \leqslant C/\sqrt{n}.$$

One can show that, for the interval maps under consideration, the condition $\sigma^2 > 0$ is equivalent to the fact that f is not a coboundary, i.e., it can not be written as $f = g - g \circ T + c$, where c is a scalar and g is a C^1 function.

This theorem is not at all specific to interval maps, it is a general consequence of the spectral method, more precisely of a version of the Nagaev-Guivarc'h theorem that mimics the statement (and the proof) of the classical Berry-Esseen theorem in the independent case:

Theorem 3.7 (Nagaev-Guivarc'h method, Berry-Esseen version). *Under the assumptions of Theorem* 2.4*, assume moreover that* $t \mapsto \mathcal{L}_t$ *is* C^3 *and* $\sigma^2 > 0$*. There exists* $C > 0$ *such that, for any* $n > 0$*, for any real interval* J*,*

$$\left|\mathbb{P}((S_n - nA)/\sqrt{n} \in J) - \mathbb{P}(\mathcal{N}(0, \sigma^2) \in J)\right| \leqslant C/\sqrt{n}.$$

The C^3 assumption corresponds to the fact that, in probability theory, the Berry-Esseen theorem is true for random variables in L^3. This theorem readily implies Theorem 3.6 since $t \mapsto \mathcal{L}_t$ is analytic in this case.

Proof of Theorem 3.7. Let X be any real random variable, and Y be Gaussian with variance $\sigma^2 > 0$. The Berry-Esseen lemma (see for instance [**Fel66**, Lemma XVI.3.2]) ensures that, for any real interval J, for any $T > 0$,

$$|\mathbb{P}(X \in J) - \mathbb{P}(Y \in J)| \leqslant C \int_0^T \frac{|\mathbb{E}(e^{\mathbf{i}tX}) - e^{-\sigma^2 t^2/2}|}{t}\, \mathrm{d}t + C/T.$$

We will use this inequality with $T = \varepsilon\sqrt{n}$ (for some suitably small $\varepsilon \leqslant \delta$) and $X = (S_n - nA)/\sqrt{n}$. To conclude, it suffices to show that

$$\int_0^{\varepsilon\sqrt{n}} \frac{|\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/\sqrt{n}}) - e^{-\sigma^2 t^2/2}|}{t}\, \mathrm{d}t \leqslant \frac{C}{\sqrt{n}}. \tag{3.4}$$

By assumption, we have $\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/\sqrt{n}}) = e^{-\mathbf{i}t\sqrt{n}A}\langle \nu, \mathcal{L}_{t/\sqrt{n}}^n \xi\rangle$.

We have to deal in a special way with the interval $t \in [0, 1/n]$, since the factor $1/t$ in (3.4) is not integrable at 0. In this case, we use a crude estimate:

$$\begin{aligned}\left|\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/\sqrt{n}}) - e^{-\mathbf{i}t\sqrt{n}A}\right| &= \left|\langle \nu, \mathcal{L}_{t/\sqrt{n}}^n \xi\rangle - 1\right| = \left|\langle \nu, \mathcal{L}_{t/\sqrt{n}}^n \xi\rangle - \langle \nu, \mathcal{L}_0^n \xi\rangle\right| \\ &= \left|\left\langle \nu, \sum_{k=0}^{n-1} \mathcal{L}_{t/\sqrt{n}}^k (\mathcal{L}_{t/\sqrt{n}} - \mathcal{L}_0)\mathcal{L}_0^{n-k-1}\xi \right\rangle\right|.\end{aligned}$$

The iterates of the operator $\mathcal{L}_0$ have uniformly bounded norm, by assumption. The same holds for $\mathcal{L}_{t/\sqrt{n}}$ since its dominating eigenvalue is bounded in modulus by 1. Finally, $\|\mathcal{L}_{t/\sqrt{n}} - \mathcal{L}_0\| \leqslant Ct/\sqrt{n}$. It follows that the above quantity is bounded by $n \cdot Ct/\sqrt{n}$. Integrating over t, we obtain

$$\begin{aligned}&\int_0^{1/n} \frac{|\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/\sqrt{n}}) - e^{-\sigma^2 t^2/2}|}{t}\, \mathrm{d}t \\ &\quad\leqslant \int_0^{1/n} \frac{|\mathbb{E}(e^{\mathbf{i}t(S_n - nA)/\sqrt{n}}) - e^{-\mathbf{i}t\sqrt{n}A}| + |e^{-\mathbf{i}t\sqrt{n}A} - 1| + |1 - e^{-\sigma^2 t^2/2}|}{t}\, \mathrm{d}t \\ &\quad\leqslant \int_0^{1/n} \frac{n \cdot Ct/\sqrt{n} + t\sqrt{n}A + \sigma^2 t^2/2}{t}\, \mathrm{d}t \leqslant \frac{C}{\sqrt{n}}.\end{aligned}$$

Now, we deal with the remaining interval $t \in [1/n, \varepsilon\sqrt{n}]$. The spectral decomposition of $\mathcal{L}_t$ gives a decomposition $\mathbb{E}(e^{\mathbf{i}tS_n}) = \lambda(t)^n c(t) + d_n(t)$, where c is a C^1 function with $c(0) = 1$, and d_n is exponentially small (see the proof of Theorem 2.4). Moreover, λ is C^3 at 0 (since this is the case of $\mathcal{L}_t$, and the eigenvalue is as smooth as the operator by Proposition 2.3), and it has an asymptotic expansion $\lambda(t) = e^{\mathbf{i}At - \sigma^2 t^2/2 + t^3 h(t)}$ where $h(t) = O(1)$. We will use the shorthand $\tilde{\lambda}(t) = e^{-\mathbf{i}At}\lambda(t) = e^{-\sigma^2 t^2/2 + t^3 h(t)}$, designed so that $\mathbb{E}(e^{\mathbf{i}t(S_n - nA)}) = \tilde{\lambda}(t)^n c(t) + e^{-\mathbf{i}tnA} d_n(t)$.

The contribution of d_n to the integral (3.4) is exponentially small (this is why we had to restrict to $t \geqslant 1/n$). The remaining quantity to be estimated is

$$\begin{aligned}&\int_{t=1/n}^{\varepsilon\sqrt{n}} \frac{\left|\tilde{\lambda}(t/\sqrt{n})^n c(t/\sqrt{n}) - e^{-\sigma^2 t^2/2}\right|}{t}\, \mathrm{d}t \\ &\quad\leqslant \int_{t=1/n}^{\varepsilon\sqrt{n}} |\tilde{\lambda}(t/\sqrt{n})^n| \frac{|c(t/\sqrt{n}) - 1|}{t}\, \mathrm{d}t + \int_{t=1/n}^{\varepsilon\sqrt{n}} \frac{\left|\tilde{\lambda}(t/\sqrt{n})^n - e^{-\sigma^2 t^2/2}\right|}{t}\, \mathrm{d}t.\end{aligned}$$

In the first term, we have $|c(t/\sqrt{n}) - 1| \leqslant Ct/\sqrt{n}$ since c is C^1 with $c(0) = 1$. Moreover, $\tilde{\lambda}(u) \leqslant e^{-\sigma^2 u^2/4}$ if u is small enough, thanks to its expansion $\tilde{\lambda}(u) = e^{-\sigma^2 u^2/2 + u^3 h(u)}$ with $h(u) = O(1)$. Hence, if ε is small enough, this term is $\leqslant C\int_{\mathbb{R}} e^{-\sigma^2 t^2/4}\,\mathrm{d}t/\sqrt{n} = O(1/\sqrt{n})$.

For the second term, using the inequality $|e^x - 1| \leqslant |x|e^{|x|}$, we have

$$\begin{aligned}\left|\tilde{\lambda}(t/\sqrt{n})^n - e^{-\sigma^2 t^2/2}\right| &= e^{-\sigma^2 t^2/2}\left|e^{t^3 h(t/\sqrt{n})/\sqrt{n}} - 1\right| \\ &\leqslant e^{-\sigma^2 t^2/2} e^{t^2 |h(t/\sqrt{n})| \cdot t/\sqrt{n}} \cdot t^3 |h(t/\sqrt{n})|/\sqrt{n} \\ &\leqslant e^{-\sigma^2 t^2/2} e^{t^2 \|h\|_\infty \varepsilon} \cdot t^3 \|h\|_\infty/\sqrt{n}.\end{aligned}$$

If ε is small enough so that $\|h\|_\infty \varepsilon < \sigma^2/4$, this is bounded by $Ct^3 e^{-\sigma^2 t^2/4}/\sqrt{n}$. Dividing by t and integrating, we get that the contribution of this term is also $O(1/\sqrt{n})$, as desired. □

REMARK 3.8. In this proof, we used the spectral decomposition of $\mathcal{L}_t^n$ as $\lambda(t)^n \Pi_t + Q_t^n$, which gives $\mathbb{E}(e^{\mathbf{i}tS_n}) = \lambda(t)^n c(t) + d_n(t)$ where $c(t) = \langle \nu, \Pi_t \xi \rangle$, and $d_n(t) = \langle \nu, Q_t^n \xi \rangle$ is exponentially small as $\|Q_t^n\|$ is itself exponentially small. We claim that, in fact, $|d_n(t)| \leqslant C|t| r^n$ for some $r < 1$, with an additional factor t which makes it possible to avoid the special treatment of the interval $[0, 1/n]$ in the above proof.

To show this claim, let us first note that, for any $\eta \in \mathcal{B}$,

$$1 = \mathbb{E}(e^{\mathbf{i}0S_n}) = \langle \nu, \mathcal{L}_0^n \xi \rangle = \langle \nu, \Pi_0 \xi \rangle + \langle \nu, Q_0^n \xi \rangle.$$

Letting n tend to infinity, we get $\langle \nu, \Pi_0 \xi \rangle = 1$. With the above equation, this gives for any n that $\langle \nu, Q_0^n \xi \rangle = 0$. Finally,

$$d_n(t) = \langle \nu, Q_t^n \xi \rangle = \langle \nu, (Q_t^n - Q_0^n)\xi \rangle = \sum_{k=0}^{n-1} \langle \nu, Q_t^k (Q_t - Q_0) Q_0^{n-k-1} \xi \rangle.$$

As $\|Q_t^k\|$ and $\|Q_0^{n-k-1}\|$ are both exponentially small and $\|Q_t - Q_0\| = O(t)$, this gives the desired conclusion.

4. Stable limit distributions for intermittent maps

In the previous section, we have described what happens to uniformly expanding maps: there is so much chaos that the sequence $f, f \circ T, \dots$ almost behaves like an independent sequence, and satisfies a central limit theorem. However, if T is less chaotic, one might expect a different behavior. As we can guess from the formula (3.3) for σ^2, the critical parameter is the speed of decay of correlations: if f has zero average for a given invariant measure, how fast does $\int f \cdot f \circ T^n$ tend to 0? If this sequence is summable, the formula (3.3) makes sense and one expects a central limit theorem (martingale methods make this intuition precise: if the correlations are summable and some technical conditions are satisfied, then a central limit theorem holds, see for instance [**Liv96**]). However, if the correlations are not summable, one would expect a different behavior. This paragraph is devoted to such an example, so-called *intermittent maps*.

An intermittent map is an expanding map of the interval, with uniform expansion except for a fixed point x_0 where $T'(x_0) = 1$. This implies that a point close to x_0 takes a long time to drift away from x_0, so that the dynamics does not look chaotic for a long time and one expects mixing to be rather slow. The precise behavior depends on the fine asymptotics of T close to x_0: if we have $T(x_0+h) = x_0 + h + ch^{1+\gamma}(1+o(1))$, then the exponent γ dictates almost everything.

While the results we will explain hold in a much wider setting (see [**Zwe98**]), we will for simplicity only consider an explicit family of maps, that were introduced by Liverani, Saussol and Vaienti in [**LSV99**] as a modification of the classical Pomeau-Manneville maps [**PM80**]. Given $\gamma > 0$, the map $T_\gamma : [0,1] \to [0,1]$ is defined by

$$T_\gamma(x) = \begin{cases} x(1+2^\gamma x^\gamma) & \text{if } x \leqslant 1/2, \\ 2x-1 & \text{if } x > 1/2. \end{cases}$$

It has two branches, which are both onto, and is uniformly expanding away from 0. For $\gamma < 1$, it admits an absolutely continuous invariant probability measure m (while, for $\gamma \geqslant 1$, the corresponding measure is infinite), whose density is Lipschitz on any compact subset of $(0,1]$, and grows like c/x^γ for small x. It is also known that the correlations for m decay exactly like $c/n^{1/\gamma-1}$ (see for instance [**You99**, **Sar02**, **Gou04b**]). Hence, we expect a central limit theorem for $\gamma < 1/2$, and a different behavior for $\gamma \geqslant 1/2$. This is indeed the case:

THEOREM 4.1. *Let $f : I \to \mathbb{R}$ be a C^1 function. Write $A = \int f \,\mathrm{d}m$.*

(1) *If $\gamma < 1/2$, there exists $\sigma^2 \geqslant 0$ such that $(S_n f - nA)/\sqrt{n}$ converges in distribution to a Gaussian random variable $\mathcal{N}(0,\sigma^2)$.*
(2) *If $\gamma \in (1/2, 1)$ and f is generic in the sense that $f(0) \neq \int f \,\mathrm{d}m$, there exists a stable law W of index $1/\gamma$ such that $(S_n f - nA)/n^\gamma$ tends in distribution to W.*
(3) *If $\gamma \in (1/2,1)$ and $f(0) = \int f \,\mathrm{d}m$, there exists $\sigma^2 \geqslant 0$ such that $(S_n f - nA)/\sqrt{n}$ converges in distribution to a Gaussian random variable $\mathcal{N}(0,\sigma^2)$.*

The convergence in distribution in the theorem holds in both probability spaces (I, m) and (I, Leb).

We will concentrate on the stable law case: there are numerous proofs for the central limit theorem, while the only available proof for the stable law case relies on the Nagaev-Guivarc'h spectral method. Contrary to the previous section, we will not give a complete proof of the result, rather a detailed sketch of the argument. Note that the behavior is also known for $\gamma = 1/2$: there is always convergence to a Gaussian distribution, but with a normalization $1/\sqrt{n \log n}$ if $f(0) \neq \int f \,\mathrm{d}m$, and $1/\sqrt{n}$ otherwise (see [**Gou04a**]). The proofs are very similar, so we will not say more about this case.

REMARK 4.2. Refining the method of proof below, one can also obtain a speed of convergence towards the limit distribution, as in Theorem 3.6. See [**Gou05**] for the case $\gamma < 1/2$ (convergence to a Gaussian distribution), and [**GM14**] for $\gamma > 1/2$ (convergence to stable laws).

4.1. Stable laws. Stable laws are the probability distributions on $\mathbb{R}$ answering the following question: if $X_1, X_2, \dots$ is a sequence of i.i.d. random variables, with partial sums S_n, what are the possible limits of $(S_n - A_n)/B_n$ (where A_n and B_n are suitable normalizing sequences).

DEFINITION 4.3. A probability distribution on $\mathbb{R}$ (which is not a Dirac measure) is stable if it arises as such a limit.

For instance, a Gaussian distribution is stable (just take X_i to be in L^2), but there are other examples:

EXAMPLE 4.4. Assume that the X_i are i.i.d., bounded from below and satisfy $\mathbb{P}(X_i \geqslant z) \sim z^{-\beta}$ when $z \to \infty$, for some $\beta \in (1, 2)$. Then, for $t > 0$, a simple computation shows that we have $\mathbb{E}(e^{\mathbf{i}tX_i}) = 1 + \mathbf{i}tA + ct^\beta(1 + o(1))$ when $t \to 0$, where $A = \mathbb{E}(X_i)$ and c is a complex number. It follows that, for any $t > 0$,

$$\begin{aligned}\mathbb{E}\left(e^{\mathbf{i}t\frac{S_n - nA}{n^{1/\beta}}}\right) &= \mathbb{E}(e^{\mathbf{i}tX_i/n^{1/\beta}})^n e^{-n\mathbf{i}tA/n^{1/\beta}} \\ &= \exp\bigl(\mathbf{i}tA/n^{1/\beta} + ct^\beta/n(1 + o(1))\bigr)^n e^{-n\mathbf{i}tA/n^{1/\beta}} \to e^{ct^\beta}.\end{aligned}$$

Since our random variable are real valued, their characteristic functions satisfy $\mathbb{E}(e^{-\mathbf{i}tY}) = \overline{\mathbb{E}(e^{\mathbf{i}tY})}$. Hence, convergence of the characteristic functions for $t \geqslant 0$ implies the same convergence for all $t \in \mathbb{R}$. This shows that the characteristic function of $(S_n - nA)/n^{1/\beta}$ converges pointwise to a continuous function, and therefore that this sequence converges in distribution. The limit W is a stable law by definition, with an explicit characteristic function (in particular, it is not Gaussian). One checks that it satisfies $\mathbb{P}(W > z) \sim c' z^{-\beta}$ and $\mathbb{P}(W < -z) = o(z^{-\beta})$ when $z \to \infty$, for some $c' > 0$.

This example is significant: all non-Gaussian stable laws have the same kind of behavior, characterized by an exponent $\beta \in (0, 2)$ called the *index* of the stable law. Their tails are heavy, of the order $z^{-\beta}$. Their characteristic function are explicit, of the form $\mathbb{E}(e^{\mathbf{i}tW}) = \exp(\mathbf{i}tA + ct^\beta)$ for $t > 0$ (for some $A \in \mathbb{R}$ and $c \in \mathbb{C}$). The complex number c is related to a skewness parameter, in $[0, 1]$, parameterizing the balance between the tails at $+\infty$ and $-\infty$ – the above example is called totally asymmetric since the tails at $-\infty$ are negligible with respect to the tails at $+\infty$, corresponding to a skewness parameter 1. The precise classification, due to Lévy, can for instance be found in [**Fel66**, Chapter XVII]. It will not be important for us, since we will only encounter the totally asymmetric stable laws that we have described in the above example.

Let us explain heuristically why stable laws show up for intermittent maps. Consider a function f with $\int f \,\mathrm{d}m = 0$ (so that the growth of $S_n f$ is typically sublinear) and $f(0) > 0$. If x is very close to 0, so are $T(x)$, and $T(T(x))$, and so on. In particular, the Birkhoff sums $S_n f(x)$ will grow linearly, like $nf(0)$, until $T^n x$ is far away from 0. The quantities one adds to the Birkhoff sums until one regains independence behave like $\varphi \cdot f(0)$, where φ is the time to drift away from 0. If the fixed point 0 of T is very neutral (i.e., if γ is large), then φ has heavy tails, and the Birkhoff sums really behave like the addition of random variables with heavy tails, just as in Example 4.4.

The rigorous proof of Theorem 4.1 follows this intuition. It is done in two main steps, that we will explain with more details in the following paragraphs:

(1) to regain some independence, we replace long excursions close to 0 (in which there are strong correlations) by a single step: in dynamical terms, this is an *inducing* process. We will show the convergence to a stable law for the induced map.
(2) Then, we need to go back from the induced map to the original map, using general arguments. Since there is no completely general exposition of this argument in the literature, we will give all details here.

4.2. Stable limits for the induced map. Let $Y = (1/2, 1]$. The first return time to Y of $x \in Y$ is

$$\varphi_Y(x) = \inf\{n \geqslant 1 \,:\, T^n x \in Y\}.$$

The combinatorics of T ensure that all points come back to Y. We define the *induced map* (or *first return map*) $T_Y : Y \to Y$ by $T_Y(x) = T^{\varphi_Y(x)}(x)$. Note that T_Y is uniformly expanding, i.e., $T_Y' \geqslant \alpha > 1$, since $T' = 2$ on $(1/2, 1]$.

Let $x_0 = 1$, and $x_{n+1} = T^{-1}x_n \cap [0, 1/2]$. The sequence x_n converges to 0. One checks that

$$x_n \sim c/n^{1/\gamma}, \quad x_n - x_{n+1} \sim c'/n^{1+1/\gamma}.$$

The points in $[x_{n+1}, x_n]$ are precisely those that take n iterates to reach Y. Let $y_n = T^{-1}(x_n) \cap Y$. Then the interval $Y_n = [y_n, y_{n-1}]$ comes back to Y precisely in n steps, i.e., $\varphi_Y = n$ on Y_n. Moreover, T_Y has full branches, i.e., $T_Y(Y_n) = Y$ for all $n \geqslant 1$.

Take a C^1 function f for which we want to prove a limit theorem, with $\int f \,\mathrm{d}m = 0$. Define on Y the induced function

$$f_Y(x) = \sum_{k=0}^{\varphi_Y(x)-1} f(T^k x).$$

The Birkhoff sums of f_Y for the map T_Y (that we will denote by $S_n^Y f_Y$) form a subsequence of the Birkhoff sums of f for the map T, corresponding precisely to those times where $T^n x$ comes back to Y. To prove a limit theorem for $S_n f$, the idea will be to first prove a limit theorem for $S_n^Y f_Y$ using the Nagaev-Guivarc'h spectral method (this step is due to Aaronson and Denker [**AD01**]) and then deduce a result for $S_n f$ (using general arguments of [**MT04**] and [**Zwe07**]). The original argument for Theorem 4.1, given in [**Gou04a**], relies only on the spectral method and is more complicated (but it allows extensions that are not available by more elementary methods, for instance to the local limit theorem [**Gou05**]).

PROPOSITION 4.5. *Let $\gamma \in (1/2, 1)$. Consider a C^1 function f with $\int f \,\mathrm{d}m = 0$ and $f(0) > 0$. There exists a stable law W of index $1/\gamma$ such that $S_n^Y f_Y / n^\gamma$ converges in distribution (on the probability space $(Y, \mathrm{Leb}_{|Y} / \mathrm{Leb}(Y))$) towards W.*

This proposition is proved using the spectral method. We first need to understand the distribution of f_Y. On the interval Y_n (which has a measure $\sim C/n^{1+1/\gamma}$), this function is equal to $nf(0) + o(n)$: indeed, $f_Y(x)$ is the sum $\sum_{k=0}^{n-1} f(T^k x)$, and among the $T^k x$ most are very close to 0. Writing $\mathrm{Leb}_Y = \mathrm{Leb}_{|Y} / \mathrm{Leb}(Y)$, we obtain:

$$\mathrm{Leb}_Y\{x \,:\, f_Y(x) \geqslant z\} \sim \sum_{n=z/f(0)}^{\infty} \mathrm{Leb}_Y(Y_n) \sim \sum_{n=z/f(0)}^{\infty} C/n^{1+1/\gamma} \sim C'/z^{1/\gamma}. \tag{4.1}$$

This shows that the function f_Y has heavy tails, of the order $1/z^{1/\gamma}$, just as in Example 4.4. We expect that the map T_Y, being uniformly expanding, will give enough independence to ensure that $S_n^Y f_Y$ behaves like a sum of independent random variables. In this way, the convergence to a stable law would follow from Example 4.4.

We define a transfer operator $\mathcal{L}$ for the map T_Y as in Section 3, in (3.1). Our map T_Y has infinitely many branches, but it is uniformly expanding so that most arguments of Section 3 still work: one checks that $\mathcal{L}$, acting on the space $\mathcal{B}$ of C^1

functions on Y, is a quasi-compact operator, and that it has a unique eigenvalue of modulus 1, at 1. Moreover, this eigenvalue is simple. The only difficulty is to check the distortion lemma 3.3, even for the branches of T_Y, since there are infinitely many of them. This is an elementary computation, see [**You99**] (one can also give a proof using the fact that the branches of T have a negative Schwarzian derivative).

Let us define perturbed operators $\mathcal{L}_t$, for small t, by $\mathcal{L}_t u = \mathcal{L}(e^{\mathbf{i}tf_Y}u)$. By definition, it satisfies as in Section 3 (see Lemma 3.2) the identity

$$\mathbb{E}(e^{\mathbf{i}tS_n^Y f_Y}) = \int \mathcal{L}_t^n 1 \,\mathrm{dLeb}_Y .$$

Since the function f_Y is unbounded, $e^{\mathbf{i}tf_Y}u$ is in general not C^1 even if u is. However, the operator $\mathcal{L}$ has an additional averaging effect (and it gives a small weight to regions where the derivative of f_Y is large). One can therefore check that $\mathcal{L}_t u$ is C^1 if u is C^1, i.e., $\mathcal{L}_t$ maps $\mathcal{B}$ to itself. Even more, the same computation gives that

$$\|\mathcal{L}_t - \mathcal{L}\|_{\mathcal{B}\to\mathcal{B}} \leqslant C|t|. \tag{4.2}$$

We are almost in a situation to apply Theorem 2.4. There is one assumption of this theorem which is not satisfied, of course (otherwise, $S_n^Y f_Y$ would satisfy a central limit theorem, while we expect it to converge, suitably renormalized, to a stable law): the family $t \mapsto \mathcal{L}_t$ is continuous by (4.2), but it is not C^2. Copying the first steps of the proof of Theorem 2.4, we obtain the following. Denote by $\lambda(t)$ the dominating eigenvalue of $\mathcal{L}_t$, then

$$\mathbb{E}(e^{\mathbf{i}tS_n f}) = c(t)\lambda(t)^n + d_n(t), \tag{4.3}$$

where $c(t)$ tends to 1 when $t \to 0$, and $\|d_n\|_{L^\infty([-\delta,\delta])} \to 0$. However, we do not know the expansion of $\lambda(t)$ for small t. If we could prove that this expansion is similar to the expansion of $\mathbb{E}(e^{\mathbf{i}tf_Y})$, then we could follow the argument in Example 4.4 and get the desired limit theorem (this is the essence of the Nagaev-Guivarc'h spectral method: follow the proofs of the independent case).

It remains to study $\lambda(t)$. We do it by hand, instead of relying on an abstract argument such as smoothness in Theorem 2.4. Denote by $\tilde\xi_t = \Pi_t(1)$ an eigenfunction of $\mathcal{L}_t$ for its dominating eigenvalue $\lambda(t)$ (where Π_t denotes the corresponding spectral projection). By Proposition 2.3, $t \mapsto \Pi_t$ is as smooth as $t \mapsto \mathcal{L}_t$. With (4.2), we get $\|\tilde\xi_t - \tilde\xi_0\|_{C^1} = O(|t|)$. The function $\tilde\xi_0$ is the density of the invariant measure, it is positive and therefore satisfies $\int \tilde\xi_0 > 0$. We deduce that $\int \tilde\xi_t \neq 0$ for small t. Hence, we can define a normalized eigenfunction $\xi_t = \tilde\xi_t / \int \tilde\xi_t \,\mathrm{dLeb}_Y$. It still satisfies

$$\|\xi_t - \xi_0\|_{C^1} = O(|t|).$$

Integrating the equation $\lambda(t)\xi_t = \mathcal{L}_t \xi_t$, we get:

$$\begin{aligned}\lambda(t) = \int \mathcal{L}_t\xi_t \,\mathrm{dLeb}_Y &= \int \mathcal{L}_t(\xi_t - \xi_0)\,\mathrm{dLeb}_Y + \int \mathcal{L}_t\xi_0 \,\mathrm{dLeb}_Y \\ &= \int (\mathcal{L}_t - \mathcal{L}_0)(\xi_t - \xi_0)\,\mathrm{dLeb}_Y + \int \mathcal{L}_t\xi_0 \,\mathrm{dLeb}_Y,\end{aligned}$$

where we have inserted $\mathcal{L}_0(\xi_t - \xi_0)$ in the second line since its integral vanishes by our normalization choice for ξ_t. In the first term, $\|\mathcal{L}_t - \mathcal{L}_0\| = O(|t|)$ and $\|\xi_t - \xi_0\| = O(|t|)$, hence this term is $O(t^2)$. In the second term,

$$\int \mathcal{L}_t\xi_0 \,\mathrm{dLeb}_Y = \int \mathcal{L}(e^{\mathbf{i}tf_Y}\xi_0)\,\mathrm{dLeb}_Y = \int e^{\mathbf{i}tf_Y}\xi_0 \,\mathrm{dLeb}_Y = \int e^{\mathbf{i}tf_Y}\,\mathrm{d}m_Y,$$

where m_Y is the measure with density ξ_0 (it is the normalized restriction of the absolutely continuous invariant measure to Y). We have proved that

$$\lambda(t) = \mathbb{E}_{m_Y}(e^{\mathbf{i}tf_Y}) + O(t^2). \tag{4.4}$$

The function f_Y has heavy tails, as we explain in (4.1) (there, we did the computation for Lebesgue measure, but the density ξ_0 of m_Y is continuous and the heavy tails all come from small neighborhoods of $1/2$, so the same estimate follows for m_Y). Hence, the asymptotic expansion in Example 4.4 gives, for $t > 0$,

$$\mathbb{E}_{m_Y}(e^{\mathbf{i}tf_Y}) = 1 + \mathbf{i}t\mathbb{E}_{m_Y}(f_Y) + ct^{1/\gamma}(1 + o(1)).$$

Since f_Y is the function induced by f on Y, we have $\mathbb{E}_{m_Y}(f_Y) = \frac{1}{m(Y)}\mathbb{E}_m(f) = 0$. Since $1/\gamma < 2$, the term $O(t^2)$ in (4.4) is negligible with respect to $t^{1/\gamma}$. We get, for $t > 0$,

$$\lambda(t) = 1 + ct^{1/\gamma}(1 + o(1)).$$

We can then use (4.3) and repeat the computation in Example 4.4 to deduce that $S_n^Y f_Y/n^\gamma$ converges to a stable law of index $1/\gamma$. This proves Proposition 4.5. □

4.3. Inducing limit theorems. In this paragraph, we explain how (under suitable assumptions) limit theorems for an induced map imply limit theorems for the original map. This statement is inspired by [**MT04**] and [**CG07**].

Definition 4.6. A continuous function $L : \mathbb{R}_+^* \to \mathbb{R}_+^*$ is slowly varying if, for any $\lambda > 0$, $L(\lambda x)/L(x) \to 1$ when $x \to \infty$. A function f is regularly varying with index d if it can be written as $x^d L(x)$ where L is slowly varying. A sequence a_n is regularly varying with index d if there exists a function f, regularly varying with index d, such that $a_n = f(n)$.

Definition 4.7. A family of real random variables $(X_i)_{i\in I}$ is tight if there is no possible loss of mass at infinity, i.e., for all $\varepsilon > 0$, there exists $M > 0$ such that $\mathbb{P}(|X_i| \geqslant M) < \varepsilon$ for all $i \in I$.

For instance, a sequence of random variables which converges in distribution is tight.

Theorem 4.8. *Let $T : X \to X$ be an ergodic probability preserving map, let $\alpha(n)$ and B_n be two sequences of integers which are regularly varying with positive indexes, let $A_n \in \mathbb{R}$, and let $Y \subset X$ be a subset with positive measure. We will denote by $m_Y := m_{|Y}/m(Y)$ the induced probability measure.*

Let $\varphi : Y \to \mathbb{N}^*$ be the first return time to Y for T, and $T_Y = T^\varphi : Y \to Y$ the induced map. Consider a measurable function $f : X \to \mathbb{R}$ and define $f_Y : Y \to \mathbb{R}$ by $f_Y = \sum_{k=0}^{\varphi-1} f \circ T^k$. Let us define the sequence of Birkhoff sums $S_n^Y f_Y = \sum_{k=0}^{n-1} f_Y \circ T_Y$. Assume that $(S_n^Y f_Y - A_n)/B_n$ converges in distribution (with respect to m_Y) to a random variable W. Additionally, assume either that

$$\frac{S_n^Y \varphi - n/m(Y)}{\alpha(n)} \text{ tends in probability to } 0 \text{ and } \max_{0\leqslant k\leqslant \alpha(n)} |S_k^Y f_Y|/B_n \text{ is tight} \tag{4.5}$$

or

$$\frac{S_n^Y \varphi - n/m(Y)}{\alpha(n)} \text{ is tight and } \max_{0\leqslant k\leqslant \alpha(n)} |S_k^Y f_Y|/B_n \text{ tends in probability to } 0. \tag{4.6}$$

Then $(S_n f - A_{\lfloor nm(Y)\rfloor})/B_{\lfloor nm(Y)\rfloor}$ converges in distribution (with respect to m) to W.

The intuition behind this theorem is the following. For $x \in Y$, one can write $S_N f(x) = S^Y_{n(x,N)} f_Y(x) + R_N(x)$, where $n(x,N)$ is the number of visits of x to Y until time N, and $R_N(x)$ is a remainder term corresponding to the last excursion. By Birkhoff theorem, $n(x,N)$ is close to $Nm(Y)$. Hence, we may write

$$\frac{S_N f(x) - A_{Nm(Y)}}{B_{Nm(Y)}} = \frac{S^Y_{Nm(Y)} f_Y(x) - A_{Nm(Y)}}{B_{Nm(Y)}} \\ + \frac{S^Y_{n(x,N)-Nm(Y)} f_Y(T_Y^{Nm(Y)} x)}{B_{Nm(Y)}} + \frac{R_N(x)}{B_{Nm(Y)}}.$$

The term on the right of the first line converges in distribution (with respect to m_Y) to W, by assumption. To conclude, we should show that the remaining two terms are small. For the last one, this follows from a general argument. For the middle one, one should control the growth of the Birkhoff sums of f_Y, on a time scale given by the fluctuations of $S^Y_n \varphi$ around its average $n/m(Y)$. This is exactly the role of the assumptions (4.5) or (4.6), showing that such a good control holds for the time scale $\alpha(n)$.

To complete the argument, one should finally understand how the above convergence in distribution with respect to m_Y yields a corresponding convergence with respect to m.

The precise technical implementation of the above idea to prove Theorem 4.8 is given in Subsection 4.6, relying on intermediate results of independent interest that are described in Subsections 4.4 and 4.5. Before this, we explain how Theorem 4.1 follows from the above theorem.

Proof of Theorem 4.1. We concentrate on Case (2), i.e., the case of stable laws, and indicate briefly at the end the modifications for the other cases.

Let $\gamma \in (1/2, 1)$, consider a C^1 function f with $\int f \, dm = 0$ and $f(0) > 0$. By Proposition 4.5, the sequence $S^Y_n f_Y$ converges in distribution to a stable law (for the measure Leb_Y or for the measure m_Y, this is equivalent according to Theorem 4.9 below). Denote by φ the first return time to Y. The function $\varphi - 1/m(Y)$ is the induced function on Y of the function g equal to $1 - 1/m(Y)$ on Y, and 1 outside of Y. Hence, Proposition 4.5 shows that $(S^Y_n \varphi - n/m(Y))/n^\gamma$ converges in distribution to a stable law. (Strictly speaking, g is not C^1, but it is C^1 on $[0, 1/2]$ and $[1/2, 1]$, which is sufficient for the proposition.)

We will apply Theorem 4.8 (more precisely (4.6)) with $\alpha(n) = B_n = n^\gamma$. We have just shown that $(S^Y_n \varphi - n/m(Y))/n^\gamma$ is tight. To conclude, we should show that $\max_{0 \leqslant k \leqslant N} |S^Y_k f_Y|/N$ tends in probability to 0 (where $N = n^\gamma$). It suffices to show that $S^Y_k f_Y / k$ tends almost surely to 0. This follows from Birkhoff's ergodic theorem since f_Y is integrable. This concludes the proof of the theorem in the stable law case.

To handle the other cases, one should show that the induced function f_Y satisfies a central limit theorem for T_Y, and then induce back to T. When $\gamma < 1/2$, or $\gamma > 1/2$ and $f(0) = 0 = \int f \, dm$, a simple computation shows that the induced function is square integrable. Hence, one can for instance apply Theorem 2.4 to obtain a central limit theorem for f_Y, or rely on the martingale arguments of [**Liv96**]. To go back to the original system, one may again apply Theorem 4.8. □

4.4. Limit theorems do not depend on the reference measure. The following theorem has been proved by Eagleson [**Eag76**] and popularized by Zweimüller (in much more general contexts, [**Zwe07**]). It is instrumental in the proof of Theorem 4.8.

THEOREM 4.9. *Let $T : X \to X$ be a transformation preserving an ergodic probability measure m. Let $f : X \to \mathbb{R}$ be measurable, let $A_n \in \mathbb{R}$, let B_n tend to ∞ and let m' be an absolutely continuous probability measure. Then $(S_n f - A_n)/B_n$ converges in distribution to a random variable W with respect to m if and only if it satisfies the same convergence with respect to m'.*

PROOF. For the proof, let us write $M(n, g, \varphi) = \int g((S_n f - A_n)/B_n)\varphi \, \mathrm{d}m$, where g is a bounded Lipschitz function and φ is an integrable function.

We claim that

$$M(n, g, \varphi) - M(n, g, \varphi \circ T) \to 0 \text{ when } n \to \infty. \tag{4.7}$$

Let us first prove this assuming that φ is bounded. Since g is Lipschitz and bounded, it satisfies $|g(a) - g(b)| \leqslant C \min(1, |a - b|)$. Therefore,

$$\begin{aligned}
|M(n, g, \varphi)-&M(n, g, \varphi \circ T)| \\
&= \left| \int \left(g\left[\frac{S_n f(x) - A_n}{B_n}\right] - g\left[\frac{S_n f(Tx) - A_n}{B_n}\right]\right) \varphi(Tx) \, \mathrm{d}m(x) \right| \\
&\leqslant C \int \min(1, |S_n f(x) - S_n f(Tx)|/B_n) \, \mathrm{d}m \\
&= C \int \min(1, |f(x) - f(T^n x)|/B_n) \, \mathrm{d}m \\
&\leqslant C \int (\min(1, |f|/B_n) + \min(1, |f| \circ T^n/B_n)) \, \mathrm{d}m \\
&= 2C \int \min(1, |f|/B_n) \, \mathrm{d}m.
\end{aligned}$$

This quantity converges to 0 when $n \to \infty$, since $B_n \to \infty$. This proves (4.7) for bounded φ.

In general, we have

$$|M(n, g, \psi)| \leqslant \|g\|_\infty \|\psi\|_{L^1(m)}. \tag{4.8}$$

Hence, the general case of (4.7) follows by writing $\varphi = \varphi_1 + \varphi_2$ with φ_1 bounded and $\|\varphi_2\|_{L^1(m)} \leqslant \varepsilon$: we obtain $\limsup|M(n, g, \varphi) - M(n, g, \varphi \circ T)| \leqslant 2\varepsilon\|g\|_\infty$.

Assume now that $(S_n f - A_n)/B_n$ converges in distribution with respect to m' towards W. Write $\mathrm{d}m' = \varphi \, \mathrm{d}m$ with φ integrable (and of integral 1). Let g be a bounded Lipschitz function. Then $M(n, g, \varphi) \to \mathbb{E}(g(W))$, hence $M(n, g, \varphi \circ T^k) \to \mathbb{E}(g(W))$ for any fixed k, by (4.7). Therefore, $M(n, g, S_k\varphi/k) \to \mathbb{E}(g(W))$. Let $\varepsilon > 0$, choose k large enough so that $\|S_k\varphi/k - 1\|_{L^1(m)} \leqslant \varepsilon$. Then

$$\begin{aligned}
&\limsup|M(n, g, 1) - \mathbb{E}(g(W))| \\
&\quad\leqslant \limsup|M(n, g, 1) - M(n, g, S_k\varphi/k)| + \limsup|M(n, g, S_k\varphi/k) - \mathbb{E}(g(W))|.
\end{aligned}$$

The first term is at most ε by (4.8), while the second one is 0. Hence, $M(n, g, 1)$ converges to $\mathbb{E}(g(W))$. This proves the convergence of $(S_n f - A_n)/B_n$ to W with respect to m.

Conversely, if $(S_n f - A_n)/B_n$ converges to W with respect to m, the convergence with respect to m' follows in the same way:

$$\limsup |M(n,g,\varphi) - \mathbb{E}(g(W))| \leqslant \limsup |M(n,g,\varphi) - M(n,g,S_k\varphi/k)| \\ + \limsup |M(n,g,S_k\varphi/k) - M(n,g,1)| + \limsup |M(n,g,1) - \mathbb{E}(g(W))|.$$

The third term tends to 0 by assumption, the first one tends to 0 by (4.7), and the second one is at most $C\|S_k\varphi/k - 1\|_{L^1(m)}$, which can be made arbitrarily small by choosing k large enough. □

4.5. Limit theorems do not depend on random indices. In this paragraph, we show that random time changes, if they do not deviate too much from linearity, do not change the validity of limit theorems. Again, this statement is instrumental in the proof of Theorem 4.8.

THEOREM 4.10. *Let $T : X \to X$ be a probability preserving map, and let $\alpha(n)$ and B_n be two sequences of integers which are regularly varying with positive indexes. Let also $A_n \in \mathbb{R}$. Let $f : X \to \mathbb{R}$ measurable be such that $(S_n f - A_n)/B_n$ converges in distribution to a random variable W. Let also $t_1, t_2, \dots$ be a sequence of integer valued measurable functions on X, and let $c > 0$. Assume that either*

$$\frac{t_n - cn}{\alpha(n)} \text{ tends in probability to } 0 \text{ and } \max_{0 \leqslant k \leqslant \alpha(n)} |S_k f|/B_n \text{ is tight} \tag{4.9}$$

or

$$\frac{t_n - cn}{\alpha(n)} \text{ is tight and } \max_{0 \leqslant k \leqslant \alpha(n)} |S_k f|/B_n \text{ tends in probability to } 0. \tag{4.10}$$

Then the sequence $(S_{t_n} f - A_{\lfloor cn \rfloor})/B_{\lfloor cn \rfloor}$ converges in distribution to W.

As in Theorem 4.8, the assumptions (4.9) or (4.10) say that t_n deviates from its average behavior cn roughly on an order $\alpha(n)$, which is sufficiently small so that the Birkhoff sums of f do not vary too much on this time scale.

PROOF. We will show that, under (4.9) or (4.10), there exists a sequence $\beta(n)$ of integers such that

$$|t_n - cn|/\beta(n) \text{ and } \max_{0 \leqslant k \leqslant 2\beta(n)} |S_k f|/B_n \text{ both tend in probability to } 0. \tag{4.11}$$

Let us show how it implies the theorem. It is sufficient to prove that

$$m\left\{x : \left|\frac{S_{t_n(x)} f - S_{\lfloor cn \rfloor} f}{B_{\lfloor cn \rfloor}}\right| \geqslant \varepsilon\right\} \to 0.$$

Abusing notations, we will omit the integer parts. The measure of the set in the last equation is bounded by $m\{|t_n - cn| \geqslant \beta(n)\} + m\{\exists i \in [\gamma(n), \beta(n)],\ |S_{cn+i} f - S_{cn} f| \geqslant \varepsilon B_{cn}\}$, where $\gamma(n) = -\min(cn, \beta(n))$. The measure of the first set tends to 0 by (4.11). If x belongs to the second set, then either $|S_{cn} f - S_{cn+\gamma(n)} f| \geqslant \varepsilon B_{cn}/2$ or $|S_{cn+i} f - S_{cn+\gamma(n)} f| \geqslant \varepsilon B_{cn}/2$. In both cases, $\max_{0 \leqslant k \leqslant 2\beta(n)} |S_k f|(T^{cn+\gamma(n)} x) \geqslant \varepsilon B_{cn}/2$. Since B_n is regularly varying, there exists C such that $B_{cn}/2 \geqslant C B_n$. Hence, the measure of the second set is bounded by $m\{\max_{0 \leqslant k \leqslant 2\beta(n)} |S_k f| \geqslant C\varepsilon B_n\}$, which also tends to 0 by (4.11).

To conclude the proof, it is therefore sufficient to construct $\beta(n)$ satisfying (4.11).

LEMMA 4.11. *Let Y_n be a sequence of real random variables tending in probability to 0. There exists a non-decreasing sequence $A(n) \to \infty$ such that $A(n)Y_n$ still tends in probability to 0.*

PROOF. For $k > 0$, let $N(k)$ be such that, for $n \geqslant N(k)$, $\mathbb{P}(|Y_n| > 1/k^2) \leqslant 1/k$. We can also assume that $N(k+1) > N(k)$. Define A by $A(n) = k$ when $N(k) \leqslant n < N(k+1)$, this sequence tends to infinity. Consider $k \in \mathbb{N}$, and $n \geqslant N(k)$. Let $\ell \geqslant k$ be such that $N(\ell) \leqslant n < N(\ell+1)$. Then

$$\mathbb{P}(A(n)|Y_n| > 1/k) \leqslant \mathbb{P}(A(n)|Y_n| > 1/\ell) = \mathbb{P}(|Y_n| > 1/\ell^2) \leqslant 1/\ell = 1/A(n).$$

Hence, $\mathbb{P}(A(n)|Y_n| > 1/k)$ tends to 0 for any k. □

LEMMA 4.12. *Let B_n be a regularly varying sequence with positive index, and let Y_n be a sequence of real random variables such that Y_n/B_n converges in probability to 0. Then there exists a non-decreasing sequence $\varphi(n) = o(n)$ such that $Y_n/B_{\varphi(n)}$ still converges in probability to 0. We can also ensure $\varphi(n+1) \leqslant 2\varphi(n)$ for any n, and $\varphi(n) \to \infty$.*

PROOF. Applying the previous lemma to Y_n/B_n, we obtain a non-decreasing sequence $A(n)$ tending to infinity such that $A(n)Y_n/B_n$ converges in probability to 0. Replacing $A(n)$ with $\min(A(n), \log n)$ if necessary, we can assume $A(n) = O(\log n)$. Write $B_n = n^d L(n)$ where L is slowly varying. Let $\varphi(n)$ be the integer part of $n/A(n)^{1/(2d)}$, it satisfies the equation $\varphi(n+1) \leqslant 2\varphi(n)$ since A is non-decreasing, tends to infinity since $A(n) = O(\log n)$, and

$$\frac{Y_n}{B_{\varphi(n)}} = \frac{A(n)Y_n}{B_n} \cdot \frac{B_n}{A(n)B_{\varphi(n)}}.$$

The first factor tends to 0 in probability, while the second one is equivalent to

$$\frac{n^d L(n)}{A(n)(n^d/A(n)^{1/2})L(n/A(n)^{1/(2d)})}.$$

By Potter's bounds [**BGT87**, Theorem 1.5.6], for any $\varepsilon > 0$, there exists $C > 0$ such that $L(n)/L(n/A(n)^{1/(2d)}) \leqslant CA(n)^{\varepsilon}$. Taking $\varepsilon = 1/4$, we obtain that the last equation is bounded by $C/A(n)^{1/4}$, and therefore tends to 0. □

We can now prove (4.11).

Assume first (4.9). Applying the previous lemma to $Y_n = t_n - cn$, we obtain a non-decreasing sequence $\varphi(n) = o(n)$ such that $(t_n - cn)/\alpha(\varphi(n)) \to 0$. Let $\beta(n) = \alpha(\varphi(n))/2$, then

$$\max_{0 \leqslant k \leqslant 2\beta(n)} |S_k f|/B_n = \frac{B_{\varphi(n)}}{B_n} \max_{0 \leqslant k \leqslant \alpha(\varphi(n))} |S_k f|/B_{\varphi(n)}. \tag{4.12}$$

The factor $B_{\varphi(n)}/B_n$ tends to 0 since $\varphi(n) = o(n)$ and B_n is regularly varying with positive index. The second factor is tight by assumption. Hence, (4.12) tends in probability to 0, as desired.

Assume now (4.10). Applying the previous lemma to $Y_n = \max_{0 \leqslant k \leqslant \alpha(n)} |S_k f|$, we obtain a non-decreasing sequence $\varphi(n) = o(n)$ such that

$$\max_{0 \leqslant k \leqslant \alpha(n)} |S_k f|/B_{\varphi(n)} \text{ tends in probability to } 0. \tag{4.13}$$

Let $\psi(n)$ be the smallest integer p such that $\varphi(p) \geqslant n/2$. Then $\varphi(\psi(n)) \geqslant n/2$. Moreover, $\varphi(\psi(n)-1) < n/2$, hence $\varphi(\psi(n)) < n$ by the inequality $\varphi(k+1) \leqslant 2\varphi(k)$.

Therefore, $B_n \geqslant C^{-1} B_{\varphi(\psi(n))}$ since the sequence B_n, being regularly varying with positive index, is increasing up to a constant multiplicative factor.

Let $\beta(n) = \alpha(\psi(n))/2$, we get

$$\max_{0\leqslant k\leqslant 2\beta(n)} |S_k f|/B_n = \max_{0\leqslant k\leqslant \alpha(\psi(n))} |S_k f|/B_n \leqslant C \max_{0\leqslant k\leqslant \alpha(\psi(n))} |S_k f|/B_{\varphi(\psi(n))}.$$

This converges to 0 in probability by (4.13).

Since $\varphi(\psi(n)) \geqslant n/2$ and $\varphi(k) = o(k)$, we have $n = o(\psi(n))$. Since α is regularly varying with positive index, this yields $\alpha(n) = o(\beta(n))$. In particular, the tightness of $(t_n - cn)/\alpha(n)$ implies the convergence to 0 of $(t_n - cn)/\beta(n)$. We have proved (4.11) as desired. □

4.6. Proof of Theorem 4.8. In this paragraph, we prove Theorem 4.8. Going to the natural extension, we can without loss of generality assume that T is invertible. Abusing notations, we will write $B_{nm(Y)}$ instead of $B_{\lfloor nm(Y)\rfloor}$. We will prove that $(S_n f - A_{nm(Y)})/B_{nm(Y)}$ converges to W in distribution with respect to m_Y: this will imply the desired result by Theorem 4.9, since m_Y is absolutely continuous with respect to m.

For $x \in Y$ and $N \in \mathbb{N}$, let $n(x, N) = \mathrm{Card}\{1 \leqslant i < N : T^i x \in Y\}$ denote the number of visits of x to Y. By construction, it satisfies

$$n(x, N) \geqslant k \iff S_k^Y \varphi(x) < N. \tag{4.14}$$

Define also a function H on X by $H(x) = \sum_{k=1}^{\psi(x)} f(T^{-k}x)$, where $\psi(x) = \inf\{n \geqslant 1 : T^{-n}x \in Y\}$. By construction, for $x \in Y$,

$$S_N f(x) = S_{n(x,N)}^Y f_Y(x) + H(T^N x).$$

Moreover, $H \circ T^N / B_{Nm(Y)}$ converges to 0 in distribution on X (since the measure is invariant and B_n tends to infinity), and therefore on Y. To prove the theorem, it is therefore sufficient to show that

$$\frac{S_{n(x,N)}^Y f_Y - A_{Nm(Y)}}{B_{Nm(Y)}} \to W.$$

This will follow from Theorem 4.10 if we can check its assumptions (4.9) or (4.10) for $t_N(x) = n(x, N)$ and $c = m(Y)$. The assumptions concerning $S_k^Y f_Y$ are already contained in (4.5) or (4.6) respectively, we only have to check the assumptions about t_N.

Birkhoff's theorem ensures that $n(x, N) = Nm(Y) + o(N)$ for almost every x. Therefore, along any subsequence N_k for which $\alpha(N_k) \geqslant \delta N_k$ with $\delta > 0$, we get that $|n(x, N_k) - N_k m(Y)|/\alpha(N_k)$ converges in probability to 0, and there is nothing left to prove. Thus, it is sufficient to consider only values of N along which $\alpha(N)/N \to 0$.

For any $a > 0$, we have by (4.14)

$$\begin{aligned} m_Y \left\{ \frac{n(x, N) - Nm(Y)}{\alpha(N)} \geqslant a \right\} &= m_Y \left\{ S_{Nm(Y)+\alpha(N)a}^Y \varphi < N \right\} \\ &= m_Y \left\{ \frac{S_{Nm(Y)+\alpha(N)a}^Y \varphi - (Nm(Y) + \alpha(N)a)/m(Y)}{\alpha(Nm(Y) + \alpha(N)a)} \right. \\ &\qquad\qquad \left. < -\frac{\alpha(N)}{\alpha(Nm(Y) + \alpha(N)a)} \frac{a}{m(Y)} \right\}. \end{aligned}$$

Since we are considering values of N for which $\alpha(N) = o(N)$, we have $Nm(Y) + \alpha(N)a \leqslant 2Nm(Y)$ if N is large enough. Since α is regularly varying with positive index, this yields $\frac{\alpha(N)}{\alpha(Nm(Y)+\alpha(N)a)} \geqslant C > 0$. Hence,

$$m_Y\left\{\frac{n(x,N) - Nm(Y)}{\alpha(N)} \geqslant a\right\} \leqslant m_Y\left\{\frac{S^Y_{p(N)}\varphi - p(N)/m(Y)}{\alpha(p(N))} < -Ca\right\} \tag{4.15}$$

for some integer $p(N)$ which tends to infinity with N.

Let us now study $m_Y\left\{\frac{n(x,N)-Nm(Y)}{\alpha(N)} < -a\right\}$. Using again $\alpha(N) = o(N)$, we get $Nm(Y) - \alpha(N)a \geqslant Nm(Y)/2 > 0$ for large enough N. Hence,

$$\begin{aligned} m_Y\left\{\frac{n(x,N) - Nm(Y)}{\alpha(N)} < -a\right\} &= m_Y\left\{S^Y_{Nm(Y)-\alpha(N)a}\varphi \geqslant N\right\} \\ &= m_Y\left\{\frac{S^Y_{Nm(Y)-\alpha(N)a}\varphi - (Nm(Y) - \alpha(N)a)/m(Y)}{\alpha(Nm(Y) - \alpha(N)a)}\right. \\ &\qquad\qquad \left.\geqslant \frac{\alpha(N)}{\alpha(Nm(Y) - \alpha(N)a)}\frac{a}{m(Y)}\right\}. \end{aligned}$$

Since $\frac{\alpha(N)}{\alpha(Nm(Y)-\alpha(N)a)} \geqslant C > 0$, we obtain

$$m_Y\left\{\frac{n(x,N) - Nm(Y)}{\alpha(N)} < -a\right\} \leqslant m_Y\left\{\frac{S^Y_{q(N)}\varphi - q(N)/m(Y)}{\alpha(q(N))} \geqslant Ca\right\}, \tag{4.16}$$

for some $q(N)$ tending to infinity with N.

The equations (4.15) and (4.16) together show that the tightness (resp. the convergence in probability to 0) of $(S^Y_n\varphi - n/m(Y))/\alpha(n)$ implies the tightness (resp. the convergence in probability to 0) of $(n(x,N) - Nm(Y))/\alpha(N)$. We can therefore apply Theorem 4.10, to conclude the proof. □

5. The almost-sure invariance principle

In this section, we describe another application of the Nagaev-Guivarc'h method to prove a limit theorem for dynamical systems, the almost sure invariance principle. This limit theorem ensures that the trajectories of a process can be coupled with a Brownian motion so that, almost surely, the difference between the processes is negligible with respect to their size. The precise definition is the following (we formulate it for vector-valued observables since the Nagaev-Guivarc'h method applies directly to this case, while a lot of martingale-based arguments are restricted to real-valued observables).

Definition 5.1. Let $\rho \in (0, 1/2]$ and let Σ^2 be a (possibly degenerate) $d \times d$ symmetric nonnegative matrix. A random process $(X_0, X_1, \dots)$ taking its values in $\mathbb{R}^d$ satisfies the almost sure invariance principle with error rate $o(n^\rho)$ and limiting covariance Σ^2 if there exist a probability space Ω and two processes $(X_0^*, X_1^*, \dots)$ and $(B_0, B_1, \dots)$ on Ω such that:

(1) The processes $(X_0, X_1, \dots)$ and $(X_0^*, X_1^*, \dots)$ have the same distribution.
(2) The random variables $B_0, B_1, \dots$ are i.i.d. and distributed like $\mathcal{N}(0, \Sigma^2)$.

(3) Almost surely in Ω,

$$\left|\sum_{\ell=0}^{n-1} X_\ell^* - \sum_{\ell=0}^{n-1} B_\ell\right| = o(n^\rho).$$

This property implies the central limit theorem, but it is much more precise (for instance, it readily gives the law of the iterated logarithm, describing the almost sure growth rate of the partial sums $\sum_{\ell=0}^{n-1} X_\ell$). Such a result is well known for sums of i.i.d. random variables, but it is delicate: the optimal error rate ($O(n^{1/p})$ for random variables in L^p) has only been proved for real-valued random variables in 1975 for $p \geqslant 4$ [**KMT75**] (the case $p < 4$ is easier, and can be handled with different methods, for instance using Skorokhod embedding). For $\mathbb{R}^d$-valued random variables, the result has been proved even more recently [**Ein89**, **Zaĭ98**], and is really difficult.

The almost sure invariance principle can be proved for some dynamical systems using martingale or approximation arguments, if there is a well behaved underlying filtration (see for instance [**HK82**, **DP84**, **MN09**]). In this section, we describe the results of [**Gou10a**] relying on the Nagaev-Guivarc'h spectral method, that give better error bounds and apply to a whole range of dynamical systems (for instance those described at the end of Section 3). To simplify notations, we write tx instead of $\langle t, x\rangle$ for $t \in \mathbb{R}^d$ and $x \in \mathbb{R}^d$. The main theorem is the following.

THEOREM 5.2. *Let $X_1, X_2, \dots$ be a centered stationary sequence of $\mathbb{R}^d$-valued random variables, with partial sums denoted by S_n. Assume that there exist a complex Banach space $\mathcal{B}$, a family of operators $\mathcal{L}_t$ acting on $\mathcal{B}$ (for $t \in \mathbb{R}^d$ with $|t| \leqslant \delta$) and $\xi \in \mathcal{B}$, $\nu \in \mathcal{B}^*$ such that:*

(1) *strong coding: for all $n \in \mathbb{N}$, for all $t_0, \dots, t_{n-1} \in \mathbb{R}^d$ with $|t_i| \leqslant \delta$,*

$$\mathbb{E}(e^{\mathbf{i}\sum_{\ell=0}^{n-1} t_\ell X_\ell}) = \langle \nu, \mathcal{L}_{t_{n-1}} \mathcal{L}_{t_{n-2}} \cdots \mathcal{L}_{t_1} \mathcal{L}_{t_0} \xi\rangle.$$

(2) *spectral description: $r_{\mathrm{ess}}(\mathcal{L}_0) < 1$, and $\mathcal{L}_0$ has a single eigenvalue of modulus $\geqslant 1$, located at 1. It is an isolated eigenvalue, of multiplicity one.*

(3) *weak regularity: there exists $C > 0$ such that for all $|t| \leqslant \delta$ and all $n \in \mathbb{N}$, we have $\|\mathcal{L}_t^n\|_{\mathcal{B}\to\mathcal{B}} \leqslant C$.*

(4) *there exists $p > 2$ such that $\|X_i\|_{L^p} \leqslant C$.*

Then there exists a matrix Σ^2 such that $S_n/\sqrt{n}$ converges in distribution to $\mathcal{N}(0, \Sigma^2)$. Moreover, the process $(X_\ell)_{\ell\in\mathbb{N}}$ satisfies the almost sure invariance principle with error rate $o(n^\rho)$ for all ρ with

$$\rho > \frac{p}{4p-4} = \frac{1}{4} + \frac{1}{4p-4}.$$

When $p = \infty$, the condition on the error rate becomes $\rho > 1/4$. This bound is rather good (in particular dimension-independent), although it is much weaker than the result for i.i.d. random variables (one should be able to take any $\rho > 0$). In really non-independent situations, all the methods seem to be stuck at this 1/4-boundary, excepted the very recent paper [**BLW14**].

This theorem should be compared to the usual Nagaev-Guivarc'h result giving the central limit theorem, Theorem 2.4.

- The coding assumption (1) in Theorem 5.2 is stronger (instead of looking at only one Fourier parameter, one should be able to change it with time),

but this is minor since operators that satisfy the usual coding property of Theorem 2.4 usually also satisfy the strong coding property. For instance, one readily checks that the proof of Lemma 3.2 also gives the strong coding property for dynamical systems.
- The spectral assumption (2) on $\mathcal{L}_0$ is the same.
- The regularity assumption (3) is considerably weaker in Theorem 5.2: since it does not require any continuity, the operator $\mathcal{L}_t$ for $t \neq 0$ may not have a dominating eigenvalue $\lambda(t)$. In particular, the usual strategy of the Nagaev-Guivarc'h theory (reduce everything to the study of the dominating eigenvalue) can not work!
- The price to pay for the very weak condition (3) is an additional condition (4) of L^p boundedness for some $p > 2$. This condition is not needed in Theorem 2.4 (and, indeed, significant applications are to merely L^2 functions). However, it is natural for the almost sure invariance principle since it is necessary even for i.i.d. random variables.

REMARK 5.3. If $t \mapsto \mathcal{L}_t$ is continuous, then the weak regularity condition (3) is satisfied. Indeed, in this case, Proposition 2.3 shows that one can write for small t a decomposition $\mathcal{L}_t^n = \lambda(t)^n \Pi_t + Q_t^n$ where Π_t is the eigenprojection corresponding to the dominating eigenvalue $\lambda(t)$ of $\mathcal{L}_t$, and $Q_t = (I - \Pi_t)\mathcal{L}_t$ corresponds to the rest of the spectrum of $\mathcal{L}_t$. It satisfies $\|Q_t^n\| \leqslant Cr^n$ for some $r < 1$. To prove the condition (3), it is therefore sufficient to show that $|\lambda(t)| \leqslant 1$ for small t.

We have

$$\begin{aligned}\mathbb{E}(e^{\mathbf{i}t\sum_{\ell=0}^{n-1} X_\ell}) &= \langle \nu, \mathcal{L}_t^n \xi \rangle = \lambda(t)^n \langle \nu, \Pi_t \xi \rangle + \langle \nu, Q_t^n \xi \rangle \\ &= \lambda(t)^n \langle \nu, \Pi_t \xi \rangle + O(r^n).\end{aligned} \tag{5.1}$$

When $t \to 0$, by continuity, the quantity $\langle \nu, \Pi_t \xi \rangle$ converges to $\langle \nu, \Pi_0 \xi \rangle = 1$. In particular, if t is small enough, $\langle \nu, \Pi_t \xi \rangle \neq 0$. As the right hand side of (5.1) should remain bounded by 1 in modulus, this gives $|\lambda(t)| \leqslant 1$ as desired.

The L^p condition (4) hints at the fact that the proof of the theorem is not spectral or dynamical, but rather probabilistic in spirit. (In this sense, although the statement looks very similar to Theorem 2.4, the above theorem is not a genuine application of the spectral method.) Indeed, Theorem 5.2 will follow from a corresponding statement relying only on a technical decorrelation condition that we now describe. This condition ensures that the characteristic function of the process we consider is close enough to that of an independent process. The condition, denoted by (H), is the following: there exist $\delta > 0$ and $C, c > 0$ such that, for all $n, m > 0$, for all $b_1 < b_2 < \cdots < b_{n+m+1}$, for all $k > 0$ and for all $t_1, \dots, t_{n+m} \in \mathbb{R}^d$ with $|t_j| \leqslant \delta$, we have

$$\begin{aligned}\Bigg| &\mathbb{E}\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j \left(\sum_{\ell=b_j}^{b_{j+1}-1} X_\ell\right) + \mathbf{i}\sum_{j=n+1}^{n+m} t_j \left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} X_\ell\right)}\right) \\ &\qquad - \mathbb{E}\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j \left(\sum_{\ell=b_j}^{b_{j+1}-1} X_\ell\right)}\right) \cdot \mathbb{E}\left(e^{\mathbf{i}\sum_{j=n+1}^{n+m} t_j \left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} X_\ell\right)}\right)\Bigg| \\ &\leqslant C(1 + \max|b_{j+1} - b_j|)^{C(n+m)} e^{-ck}.\end{aligned} \tag{H}$$

This assumption means that, if we group the random variables in $n + m$ blocks, then a gap of size k between two blocks yields characteristic functions that are exponentially close (in terms of k) to independent characteristic functions, with an error that is, for each block, polynomial in terms of the size of the block. This control is only required for Fourier parameters close to 0.

This condition is of course true for independent random variables. Its main interest is that it is also satisfied under the assumptions of Theorem 5.2:

LEMMA 5.4. *Under the assumptions of Theorem* 5.2, *the property* (H) *is satisfied.*

PROOF. Write Π_0 for the eigenprojection associated to the dominating eigenvalue 1 of $\mathcal{L}_0$. Then

$$\begin{aligned}\mathbb{E}&\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j\left(\sum_{\ell=b_j}^{b_{j+1}-1} X_\ell\right)+\mathbf{i}\sum_{j=n+1}^{n+m} t_j\left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} X_\ell\right)}\right)\\
&= \left\langle \nu, \mathcal{L}_{t_{n+m}}^{b_{n+m+1}-b_{n+m}} \cdots \mathcal{L}_{t_{n+1}}^{b_{n+2}-b_{n+1}} \mathcal{L}_0^k \mathcal{L}_{t_n}^{b_{n+1}-b_n} \cdots \mathcal{L}_{t_1}^{b_2-b_1} \mathcal{L}_0^{b_1} \xi \right\rangle\\
&= \left\langle \nu, \mathcal{L}_{t_{n+m}}^{b_{n+m+1}-b_{n+m}} \cdots \mathcal{L}_{t_{n+1}}^{b_{n+2}-b_{n+1}} (\mathcal{L}_0^k - \Pi_0) \mathcal{L}_{t_n}^{b_{n+1}-b_n} \cdots \mathcal{L}_{t_1}^{b_2-b_1} \mathcal{L}_0^{b_1} \xi \right\rangle\\
&\quad + \left\langle \nu, \mathcal{L}_{t_{n+m}}^{b_{n+m+1}-b_{n+m}} \cdots \mathcal{L}_{t_{n+1}}^{b_{n+2}-b_{n+1}} \Pi_0 \mathcal{L}_{t_n}^{b_{n+1}-b_n} \cdots \mathcal{L}_{t_1}^{b_2-b_1} \mathcal{L}_0^{b_1} \xi \right\rangle.\end{aligned}$$

The operators $\mathcal{L}_{t_i}$ all satisfy $\|\mathcal{L}_{t_i}^j\| \leqslant C$ by the weak regularity assumption. As $\|\mathcal{L}_0^k - \Pi_0\| \leqslant Cr^k$, we deduce that the term on the penultimate line is bounded by $C^{n+m}r^k$, which is compatible with (H). On the last line, the projection Π_0 decouples both parts, hence one proves that this line is equal to

$$\mathbb{E}\left(e^{\mathbf{i}\sum_{j=1}^{n} t_j\left(\sum_{\ell=b_j}^{b_{j+1}-1} X_\ell\right)}\right)\cdot \mathbb{E}\left(e^{\mathbf{i}\sum_{j=n+1}^{n+m} t_j\left(\sum_{\ell=b_j+k}^{b_{j+1}+k-1} X_\ell\right)}\right) + O(C^m r^{b_{n+1}+k}).$$

This proves (H). ☐

The main probabilistic result is the following.

THEOREM 5.5. *Let* $(X_0, X_1, \dots)$ *be a centered stationary sequence of* $\mathbb{R}^d$*-valued random variables, in* L^p *for some* $p > 2$, *satisfying* (H). *Then*

1. *The covariance matrix* $\operatorname{cov}(\sum_{\ell=0}^{n-1} X_\ell)/n$ *converges to a limiting matrix* Σ^2.
2. *The sequence* $\sum_{\ell=0}^{n-1} X_\ell/\sqrt{n}$ *converges in distribution to* $\mathcal{N}(0, \Sigma^2)$.
3. *The process* $(X_0, X_1, \dots)$ *satisfies the almost sure invariance principle with limiting covariance* Σ^2 *and error rate* $o(n^\rho)$ *for any* $\rho > p/(4p-4)$.

Theorem 5.2 readily follows from this theorem together with Lemma 5.4. Note that Theorem 5.5 admits a non-stationary version, implying a non-stationary version of Theorem 5.2 (see [**Gou10a**] for details). For simplicity, we stick to the stationary version in these notes.

5.1. Tools for the proof of Theorem 5.5. The proof of Theorem 5.5 only uses probabilistic tools (note that there is no dynamical system involved in the statement of the result, and no spectral assumption). It relies in a crucial way on the notion of *coupling*.

Given $Z_1 : \Omega_1 \to E_1$ and $Z_2 : \Omega_2 \to E_2$ two random variables on two (possibly different) probability spaces, a coupling between Z_1 and Z_2 is a way to associate those random variables, generally in order to show that they are close in a suitable sense. Formally, a coupling between Z_1 and Z_2 is given by a probability space

Ω' and two random variables $Z'_1 : \Omega' \to E_1$ and $Z'_2 : \Omega' \to E_2$ such that Z'_i is distributed like Z_i. By considering the distribution of (Z'_1, Z'_2) in $E_1 \times E_2$, one can without loss of generality take $\Omega' = E_1 \times E_2$, and Z'_1 and Z'_2 respectively the first and second projection. If $E_1 = E_2$, one often tries to ensure that Z'_1 and Z'_2 are close, for instance by minimizing $\|Z'_1 - Z'_2\|_{L^p}$ for some p, or $\mathbb{P}(Z'_1 \neq Z'_2)$.

DEFINITION 5.6. Let P, Q be two probability measures on a metric space. Their Prokhorov distance $\pi(P, Q)$ is the smallest $\varepsilon > 0$ such that $P(B) \leqslant \varepsilon + Q(B^\varepsilon)$ for any Borel set B, where B^ε denotes the ε-neighborhood of B.

This distance is symmetric, although this is not completely trivial from the definition. Its interest is that it makes it possible to construct good couplings, thanks to the following Dudley-Strassen theorem [**Bil99**, Theorem 6.9].

THEOREM 5.7. *Let X and Y be two random variables taking their values in a metric space, with respective distributions P_X and P_Y. If $\pi(P_X, P_Y) < c$, there exists a coupling between X and Y such that $\mathbb{P}(d(X, Y) > c) < c$.*

To construct good couplings during the proof of Theorem 5.5, we will only have information about the characteristic functions of the processes under study. Hence, it will be important to estimate the Prokhorov distance just in terms of characteristic functions. This is done in the following lemma. Let $d > 0$ and $N > 0$. We consider $\mathbb{R}^{dN}$ with the norm

$$|(x_1, \dots, x_N)|_N = \sup_{1 \leqslant i \leqslant N} |x_i|,$$

where $|x|$ is the Euclidean norm of $x \in \mathbb{R}^d$.

LEMMA 5.8. *There exists a constant $C(d)$ with the following property. Let F and G be two probability measures on $\mathbb{R}^{dN}$, with characteristic functions φ and γ. For all $T > 0$,*

$$\pi(F, G) \leqslant \sum_{j=1}^{N} F(|x_j| > T) + \left(C(d) T^{d/2}\right)^N \left[\int_{\mathbb{R}^{dN}} |\varphi - \gamma|^2\right]^{1/2}.$$

PROOF. Using an approximation argument, we may assume that F and G have densities f and g. For any Borel set B,

$$\begin{aligned} F(B) - G(B) &\leqslant F(B \cap \max|x_j| \leqslant T) + F(\max|x_j| > T) - G(B \cap \max|x_j| \leqslant T) \\ &\leqslant \int_{|x_1|, \dots, |x_N| \leqslant T} |f - g| + \sum_{j=1}^{N} F(|x_j| > T). \end{aligned}$$

As a consequence, $\pi(F, G)$ is bounded by the last line of this equation. To conclude, it suffices to estimate $\int_{|x_1|, \dots, |x_N| \leqslant T} |f - g|$. We have

$$\int_{|x_1|, \dots, |x_N| \leqslant T} |f - g| \leqslant \|f - g\|_{L^2} \|1_{|x_1|, \dots, |x_N| \leqslant T}\|_{L^2} = \|\varphi - \gamma\|_{L^2} (CT)^{dN/2},$$

since the Fourier transform is an isometry on L^2, up to a multiplicative factor $(2\pi)^{dN/2}$. $\square$

We illustrate the usefulness of the above tools with the following proposition.

PROPOSITION 5.9. *Let $(X_0, X_1, \dots)$ be a centered process, bounded in L^p for some $p > 2$, satisfying* (H). *For every $\eta > 0$, there exists $C > 0$ such that, for all $m, n \in \mathbb{N}$,*

$$\left\| \sum_{\ell=m}^{m+n-1} X_\ell \right\|_{L^{p-\eta}} \leqslant C n^{1/2}.$$

This kind of moment estimates is classical for a large class of weakly dependent processes. The interest of the proposition is that, here, those bounds are proved assuming only the assumption (H) on characteristic functions, which apparently does not give any information on moments.

PROOF. The crucial step is to obtain an L^2 bound. The bound in $L^{p-\eta}$ follows using the same techniques and Rosenthal inequality for sums of independent random variables [**Ros70**].

Write $u_n = \max_{m\in\mathbb{N}} \|\sum_{\ell=m}^{m+n-1} X_\ell\|_{L^2}^2$. We will show that u_{a+b} is essentially bounded by $u_a + u_b$, which gives the result inductively. To prove this, we decompose the sum $\sum_{\ell=m}^{m+a+b-1} X_\ell$ as $Y_1 + Y_2 + Y_3$ where $Y_1 = \sum_{m}^{m+a-1} X_\ell$ and $Y_3 = \sum_{m+a+b^\alpha}^{m+a+b-1} X_\ell$, while Y_2 is the sum over the central interval. If α is small enough, Y_2 is negligible and can safely be ignored. Let Y_1' and Y_3' be two independent copies of Y_1 and Y_3. The assumption (H) ensures that the characteristic function of (Y_1, Y_3) is very close to that of (Y_1', Y_3'), for Fourier parameters close enough to 0. Then, we regularize things to eliminate large Fourier parameters, as follows. Let U be a fixed random variable in L^p, whose characteristic function is supported in $\{|t| \leqslant \delta\}$, and denote by U_1, U_3, U_1', U_3' four independent copies of U. Then the characteristic function of $(Y_1 + U_1, Y_3 + U_3)$ is everywhere close to that of $(Y_1' + U_1', Y_3' + U_3')$. Lemma 5.8 makes it possible to couple those random variables, with a difference that is very small on a very big part of the space. Forgetting about the U variables (that are bounded in L^p independently of m, n), we obtain

$$\|Y_1 + Y_3\|_{L^2} \leqslant \|Y_1 - Y_1'\|_{L^2} + \|Y_3 - Y_3'\|_{L^2} + \|Y_1' + Y_3'\|_{L^2}.$$

The first two terms are small, while the last one is equal to $(\|Y_1'\|_{L^2}^2 + \|Y_3'\|_{L^2}^2)^{1/2}$ by independence, and is therefore bounded by $(u_a + u_b)^{1/2}$. We have obtained the inequality $u_{a+b} \leqslant u_a + u_b + r(a, b)$, where $r(a, b)$ is small enough not to be a serious problem. □

5.2. Sketch of proof of Theorem 5.5. We will now give some ideas of the proof of Theorem 5.5. Let us consider a process $(X_0, X_1, \dots)$ satisfying the assumptions of this theorem. The first step of the proof is to show the convergence of $\mathrm{cov}(\sum_{\ell=0}^{n-1} X_\ell)/n$ to some matrix Σ^2. This is the same kind of argument as in the above moment control (reduction to the independent situation by a coupling argument, and careful control of the error), we will not say more about it. From this point on, we will also assume that Σ^2 is non-degenerate (the degenerate case should be handled differently, it is in fact easier). Our goal is to prove the almost sure invariance principle (from which the central limit follows), with a good control on the error term.

The strategy of the proof is very classical: we will use small blocks of variables that we will throw away, replacing them with gaps, and big blocks that we will couple with independent copies (the gaps giving enough independence, thanks to the assumption (H)). The almost sure invariance principle for independent variables will then be applied, giving the desired result. One should be careful enough so that the small blocks are small enough to be negligible (this will be proved using moment estimates such as Proposition 5.9), and large enough to give enough independence.

This method is usually implemented using blocks of polynomial size. In our case, we obtain better results by using a Cantor-like triadic approach, as follows. First, we write $\mathbb{N}$ as the union of the intervals $[2^n, 2^{n+1})$. In each of these intervals, we put a small (but not too small) block in the middle, then a smaller block in the middle of the two newly created intervals, then an even smaller block in the middle of the four remaining intervals, and so on. The interest of this construction is that, to obtain n well separated big blocks, the classical argument uses small blocks whose lengths add up to n^2, while here their lengths only add up to n, making the final estimates better.

The precise construction depends on two parameters $\beta \in (0,1)$ and $\varepsilon < 1-\beta$. Let $f = f(n) = \lfloor \beta n \rfloor$. We decompose $[2^n, 2^{n+1})$ as a union of $F = 2^f$ intervals $(I_{n,j})_{0 \leqslant j < F}$ which all have the same length (the big blocks), and F small blocks $(J_{n,j})_{0 \leqslant j < F}$ giving enough independence, constructed as explained above. The central small block $J_{n,F/2}$ has length $2^{\lfloor \varepsilon n \rfloor} 2^{f-1}$, the next two small blocks have half size, and so on. The big and small blocks are laid alternatively and in increasing order, as follows:

$$[2^n, 2^{n+1}) = J_{n,0} \cup I_{n,0} \cup J_{n,1} \cup I_{n,1} \cdots \cup J_{n,F-1} \cup I_{n,F-1}.$$

We will write $(n', j') \prec (n, j)$ if the interval $I_{n',j'}$ is to the left of $I_{n,j}$ (this corresponds to the lexicographic order), and denote by $i_{n,j}$ the smallest element of $I_{n,j}$.

Let $X_{n,j} = \sum_{\ell \in I_{n,j}} X_\ell$, for $n \in \mathbb{N}$ and $0 \leqslant j < F(n)$. Denote by $\mathcal{I} = \bigcup_{n,j} I_{n,j}$ the union of the big blocks (on which we will do the coupling), and by $\mathcal{J} = \bigcup_{n,j} J_{n,j}$ the union of the small blocks (that we will neglect). The different steps in the proof of Theorem 5.5 are the following:

(1) There exists a coupling between $(X_{n,j})$ and a sequence of independent random variables $(Y_{n,j})$, where $Y_{n,j}$ is distributed like $X_{n,j}$, such that, almost surely, when $(n,j) \to \infty$,

$$\left| \sum_{(n',j') \prec (n,j)} X_{n',j'} - Y_{n',j'} \right| = o(2^{(\beta+\varepsilon)n/2}).$$

(2) There exists a coupling between $(Y_{n,j})$ and a sequence of independent Gaussian random variables $Z_{n,j}$, with covariance $\operatorname{cov}(Z_{n,j}) = |I_{n,j}| \Sigma^2$, such that, almost surely, when $(n,j) \to \infty$,

$$\left| \sum_{(n',j') \prec (n,j)} Y_{n',j'} - Z_{n',j'} \right| = o(2^{(\beta+\varepsilon)n/2} + 2^{((1-\beta)/2 + \beta/p + \varepsilon)n}).$$

(3) Coupling $X_{n,j}$ with $Z_{n,j}$ thanks to the first two steps, and writing $Z_{n,j}$ as the sum of $|I_{n,j}|$ Gaussian random variables $\mathcal{N}(0, \Sigma^2)$, we obtain a coupling between $(X_\ell)_{\ell \in \mathcal{I}}$ and $(B_\ell)_{\ell \in \mathcal{I}}$ where the B_ℓ are independent and

distributed like $\mathcal{N}(0, \Sigma^2)$. Moreover, when $(n, j) \to \infty$,

$$\left| \sum_{\ell < i_{n,j},\, \ell \in \mathcal{I}} X_\ell - B_\ell \right| = o(2^{(\beta+\varepsilon)n/2} + 2^{((1-\beta)/2+\beta/p+\varepsilon)n}).$$

(4) We easily check that, when $(n, j) \to \infty$,

$$\max_{m < |I_{n,j}|} \left| \sum_{\ell = i_{n,j}}^{i_{n,j}+m} X_\ell \right| = o(2^{((1-\beta)/2+\beta/p+\varepsilon)n}).$$

Moreover, the B_ℓ satisfy an analogous estimate.

(5) The two last steps show that, when k tends to infinity,

$$\left| \sum_{\ell < k,\, \ell \in \mathcal{I}} X_\ell - B_\ell \right| = o(k^{(\beta+\varepsilon)/2} + k^{(1-\beta)/2+\beta/p+\varepsilon}).$$

(6) Finally, we show that the small blocks are negligible: almost surely,

$$\sum_{\ell < k,\, \ell \in \mathcal{J}} X_\ell = o(k^{\beta/2+\varepsilon}),$$

and the B_ℓ satisfy the same estimate.

In this way, we get a coupling such that, almost surely,

$$\left| \sum_{\ell < k} X_\ell - B_\ell \right| = o(k^{\beta/2+\varepsilon} + k^{(1-\beta)/2+\beta/p+\varepsilon}).$$

Finally, we choose β so that the two error terms coincide, i.e., $\beta = p/(2p-2)$. This yields the almost sure invariance principle with error rate $o(n^{p/(4p-4)+\varepsilon})$, for any $\varepsilon > 0$, proving Theorem 5.5.

Steps (3) and (5) above are trivial, we should justify the other ones. They are not very difficult:

- (1) follows from the assumption (H) and from Lemma 5.8.
- For (2), we should couple independent random variables. This is certainly the most difficult step, but luckily for us it has already been considered in the literature (see for instance [**Zaĭ07**, Corollary 3]).
- (4) and (6) rely on the moment estimate given in Proposition 5.9. Technically, (6) is more complicated since Proposition 5.9 deals with consecutive blocks of random variables, while the set $\mathcal{J}$ we have to control is far from being connected.

References

[AD01] Jon Aaronson and Manfred Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn. **1** (2001), no. 2, 193–237, DOI 10.1142/S0219493701000114. MR1840194 (2002h:37014)

[BGK07] Jean-Baptiste Bardet, Sébastien Gouëzel, and Gerhard Keller, *Limit theorems for coupled interval maps*, Stoch. Dyn. **7** (2007), no. 1, 17–36, DOI 10.1142/S0219493707001913. MR2293068 (2008f:37186)

[BGT87] Nicholas H. Bingham, Charles M. Goldie, and Jozef L. Teugels, *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27, Cambridge University Press, Cambridge, 1987. MR898871 (88i:26004)

[Bil99] Patrick Billingsley, *Convergence of probability measures*, 2nd ed., Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, 1999. A Wiley-Interscience Publication. MR1700749 (2000e:60008)

[BLW14] István Berkes, Weidong Liu, and Wei Biao Wu, *Komlós-Major-Tusnády approximation under dependence*, Ann. Probab. **42** (2014), no. 2, 794–817, DOI 10.1214/13-AOP850. MR3178474

[BT07] Viviane Baladi and Masato Tsujii, *Anisotropic Hölder and Sobolev spaces for hyperbolic diffeomorphisms* (English, with English and French summaries), Ann. Inst. Fourier (Grenoble) **57** (2007), no. 1, 127–154. MR2313087 (2008d:37034)

[Buz97] Jérôme Buzzi, *Intrinsic ergodicity of affine maps in* $[0,1]^d$, Monatsh. Math. **124** (1997), no. 2, 97–118, DOI 10.1007/BF01300614. MR1462857 (98g:58098)

[CG07] Jean-René Chazottes and Sébastien Gouëzel, *On almost-sure versions of classical limit theorems for dynamical systems*, Probab. Theory Related Fields **138** (2007), no. 1-2, 195–234, DOI 10.1007/s00440-006-0021-6. MR2288069 (2008a:60066)

[DP84] Manfred Denker and Walter Philipp, *Approximation by Brownian motion for Gibbs measures and flows under a function*, Ergodic Theory Dynam. Systems **4** (1984), no. 4, 541–552, DOI 10.1017/S0143385700002637. MR779712 (86g:28025)

[DZ11] Mark F. Demers and Hong-Kun Zhang, *Spectral analysis of the transfer operator for the Lorentz gas*, J. Mod. Dyn. **5** (2011), no. 4, 665–709, DOI 10.3934/jmd.2011.5.665. MR2903754

[Eag76] G. K. Eagleson, *Some simple conditions for limit theorems to be mixing* (Russian), Teor. Verojatnost. i Primenen. **21** (1976), no. 3, 653–660. MR0428388 (55 #1409)

[Ein89] Uwe Einmahl, *Extensions of results of Komlós, Major, and Tusnády to the multivariate case*, J. Multivariate Anal. **28** (1989), no. 1, 20–68, DOI 10.1016/0047-259X(89)90097-3. MR996984 (90g:60032)

[Fel66] William Feller, *An introduction to probability theory and its applications. Vol. II*, John Wiley & Sons, Inc., New York-London-Sydney, 1966. MR0210154 (35 #1048)

[GH88] Yves Guivarc'h and Jean Hardy, *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d'Anosov* (French, with English summary), Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), no. 1, 73–98. MR937957 (89m:60080)

[GL08] Sébastien Gouëzel and Carlangelo Liverani, *Compact locally maximal hyperbolic sets for smooth maps: fine statistical properties*, J. Differential Geom. **79** (2008), no. 3, 433–477. MR2433929 (2010a:37050)

[GM14] Sébastien Gouëzel and Ian Mebourne, *Moment bounds and concentration inequalities for slowly mixing dynamical systems*, Electron. J. Probab. **19** (2014), no. 93, 30.

[Gou04a] Sébastien Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Probab. Theory Related Fields **128** (2004), no. 1, 82–122, DOI 10.1007/s00440-003-0300-4. MR2027296 (2004k:37016)

[Gou04b] Sébastien Gouëzel, *Sharp polynomial estimates for the decay of correlations*, Israel J. Math. **139** (2004), 29–65, DOI 10.1007/BF02787541. MR2041223 (2004m:37011)

[Gou05] Sébastien Gouëzel, *Berry-Esseen theorem and local limit theorem for non uniformly expanding maps* (English, with English and French summaries), Ann. Inst. H. Poincaré Probab. Statist. **41** (2005), no. 6, 997–1024, DOI 10.1016/j.anihpb.2004.09.002. MR2172207 (2007b:60071)

[Gou10a] Sébastien Gouëzel, *Almost sure invariance principle for dynamical systems by spectral methods*, Ann. Probab. **38** (2010), no. 4, 1639–1671, DOI 10.1214/10-AOP525. MR2663640 (2011h:60073)

[Gou10b] Sébastien Gouëzel, *Characterization of weak convergence of Birkhoff sums for Gibbs-Markov maps*, Israel J. Math. **180** (2010), 1–41, DOI 10.1007/s11856-010-0092-z. MR2735054 (2011k:37010)

[Hen93] Hubert Hennion, *Sur un théorème spectral et son application aux noyaux lipchitziens* (French, with English summary), Proc. Amer. Math. Soc. **118** (1993), no. 2, 627–634, DOI 10.2307/2160348. MR1129880 (93g:60141)

[Her05] Loïc Hervé, *Théorème local pour chaînes de Markov de probabilité de transition quasi-compacte. Applications aux chaînes V-géométriquement ergodiques et aux modèles itératifs* (French, with English and French summaries), Ann. Inst. H. Poincaré Probab.

Statist. **41** (2005), no. 2, 179–196, DOI 10.1016/j.anihpb.2004.04.001. MR2124640 (2005k:60223)

[HH01] Hubert Hennion and Loïc Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Lecture Notes in Mathematics, vol. 1766, Springer-Verlag, Berlin, 2001. MR1862393 (2002h:60146)

[HK82] Franz Hofbauer and Gerhard Keller, *Ergodic properties of invariant measures for piecewise monotonic transformations*, Math. Z. **180** (1982), no. 1, 119–140, DOI 10.1007/BF01215004. MR656227 (83h:28028)

[HP10] Loïc Hervé and Françoise Pène, *The Nagaev-Guivarc'h method via the Keller-Liverani theorem*, Bull. Soc. Math. France **138** (2010), 415–489. MR2729019

[Kat66] Tosio Kato, *Perturbation theory for linear operators*, Die Grundlehren der mathematischen Wissenschaften, Band 132, Springer-Verlag New York, Inc., New York, 1966. MR0203473 (34 #3324)

[Kel85] Gerhard Keller, *Generalized bounded variation and applications to piecewise monotonic transformations*, Z. Wahrsch. Verw. Gebiete **69** (1985), no. 3, 461–478, DOI 10.1007/BF00532744. MR787608 (86i:28024)

[KMT75] János Komlós, Péter Major, and Gábor Tusnády, *An approximation of partial sums of independent* RV*'s and the sample* DF*. I*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **32** (1975), 111–131. MR0375412 (51 #11605b)

[Liv96] Carlangelo Liverani, *Central limit theorem for deterministic systems*, International Conference on Dynamical Systems (Montevideo, 1995), Pitman Res. Notes Math. Ser., vol. 362, Longman, Harlow, 1996, pp. 56–75. MR1460797 (98k:28025)

[LSV99] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti, *A probabilistic approach to intermittency*, Ergodic Theory Dynam. Systems **19** (1999), no. 3, 671–685, DOI 10.1017/S0143385799133856. MR1695915 (2000d:37029)

[MN09] Ian Melbourne and Matthew Nicol, *A vector-valued almost sure invariance principle for hyperbolic dynamical systems*, Ann. Probab. **37** (2009), no. 2, 478–505, DOI 10.1214/08-AOP410. MR2510014 (2010i:37016)

[MT04] Ian Melbourne and Andrei Török, *Statistical limit theorems for suspension flows*, Israel J. Math. **144** (2004), 191–209, DOI 10.1007/BF02916712. MR2121540 (2006c:37005)

[Nag57] S. V. Nagaev, *Some limit theorems for stationary Markov chains* (Russian, with English summary), Teor. Veroyatnost. i Primenen. **2** (1957), 389–416. MR0094846 (20 #1355)

[Nag61] S. V. Nagaev, *More exact limit theorems for homogeneous Markov chains* (Russian, with English summary), Teor. Verojatnost. i Primenen. **6** (1961), 67–86. MR0131291 (24 #A1143)

[PM80] Yves Pomeau and Paul Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Comm. Math. Phys. **74** (1980), no. 2, 189–197. MR576270 (81g:58024)

[PP90] William Parry and Mark Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics* (English, with French summary), Astérisque **187-188** (1990), 268. MR1085356 (92f:58141)

[RE83] J. Rousseau-Egele, *Un théorème de la limite locale pour une classe de transformations dilatantes et monotones par morceaux* (French, with English summary), Ann. Probab. **11** (1983), no. 3, 772–788. MR704569 (84m:60032)

[Ros70] Haskell P. Rosenthal, *On the subspaces of L^p $(p > 2)$ spanned by sequences of independent random variables*, Israel J. Math. **8** (1970), 273–303. MR0271721 (42 #6602)

[Sar02] Omri Sarig, *Subexponential decay of correlations*, Invent. Math. **150** (2002), no. 3, 629–653, DOI 10.1007/s00222-002-0248-5. MR1946554 (2004e:37010)

[Sau00] Benoît Saussol, *Absolutely continuous invariant measures for multidimensional expanding maps*, Israel J. Math. **116** (2000), 223–248, DOI 10.1007/BF02773219. MR1759406 (2001e:37037)

[Tsu10] Masato Tsujii, *Quasi-compactness of transfer operators for contact Anosov flows*, Nonlinearity **23** (2010), no. 7, 1495–1545, DOI 10.1088/0951-7715/23/7/001. MR2652469 (2012e:37040)

[Tsu12] Masato Tsujii, *Contact Anosov flows and the Fourier-Bros-Iagolnitzer transform*, Ergodic Theory Dynam. Systems **32** (2012), no. 6, 2083–2118, DOI 10.1017/S0143385711000605. MR2995886

[You98] Lai-Sang Young, *Statistical properties of dynamical systems with some hyperbolicity*, Ann. of Math. (2) **147** (1998), no. 3, 585–650, DOI 10.2307/120960. MR1637655 (99h:58140)

[You99] Lai-Sang Young, *Recurrence times and rates of mixing*, Israel J. Math. **110** (1999), 153–188, DOI 10.1007/BF02808180. MR1750438 (2001j:37062)

[Zaĭ98] Andrei Yu. Zaitsev, *Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments*, ESAIM Probab. Statist. **2** (1998), 41–108 (electronic), DOI 10.1051/ps:1998103. MR1616527 (99k:60094)

[Zaĭ07] Andrei Yu. Zaĭtsev, *Estimates for the accuracy of the strong approximation in the multidimensional invariance principle* (Russian, with English and Russian summaries), Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **339** (2006), no. Veroyatn. i Stat. 10, 37–53, 176, DOI 10.1007/s10958-007-0319-7; English transl., J. Math. Sci. (N. Y.) **145** (2007), no. 2, 4856–4865. MR2355400 (2008g:60093)

[Zwe98] Roland Zweimüller, *Ergodic structure and invariant densities of non-Markovian interval maps with indifferent fixed points*, Nonlinearity **11** (1998), no. 5, 1263–1276, DOI 10.1088/0951-7715/11/5/005. MR1644385 (99e:58125)

[Zwe07] Roland Zweimüller, *Mixing limit theorems for ergodic transformations*, J. Theoret. Probab. **20** (2007), no. 4, 1059–1071, DOI 10.1007/s10959-007-0085-y. MR2359068 (2008h:60119)

IRMAR, CNRS UMR 6625, Université de Rennes 1, 35042 Rennes, France
E-mail address: sebastien.gouezel@univ-rennes1.fr