Empirical Invariance Principle for Ergodic Torus Automorphisms. Genericity

Olivier Durieu*
olivier.durieu@etu.univ-rouen.fr

Philippe Jouan*
philippe.jouan@univ-rouen.fr

March 18, 2008

Abstract

We consider the dynamical system given by an algebraic ergodic automorphism T on a torus. We study a Central Limit Theorem for the empirical process associated to the stationary process $(f \circ T^i)_{i\in\mathbb{N}}$, where f is a given $\mathbb{R}$-valued function. We give a sufficient condition on f for this Central Limit Theorem to hold.

In a second part, we prove that the distribution function of a Morse function is continuously differentiable if the dimension of the manifold is at least 3 and Hölder continuous if the dimension is 1 or 2. As a consequence, the Morse functions satisfy the empirical invariance principle, which is therefore generically verified.

Keywords: Empirical process; Partially hyperbolic dynamical system; Functional central limit theorem; Multiple mixing; Distribution function; Morse function; Genericity.

AMS Subject Classification: 60G30; 60F17; 28D05; 37D30; 57R45.

Introduction

We are interested in the study of the convergence of an empirical process defined on a dynamical system. This kind of work already appears in the book of Billingsley [2] where a limit theorem for an empirical process defined by the transformation $x \longmapsto 2x \mod 1$ on $[0, 1]$ is established. In 2004, the case of expanding maps of the interval was studied by Collet, Martinez and Schmitt [5], using properties of the transfer operator. The reader can also see (for more examples) the paper by Dedecker and Prieur [6]. Here, we deal with the empirical invariance principle (Central Limit Theorem for an empirical process) in the case of a linear transformation of the torus.

Let us begin by some definitions. Let $(X, \mathcal{A}, \mu)$ be a probability space and $T : X \longrightarrow X$ a measurable μ-invariant transformation of X (i.e. $\mu(T^{-1}A) = \mu(A)$, $\forall A \in \mathcal{A}$). If f is a

*Laboratoire de Mathématiques Raphaël Salem UMR 6085 CNRS-Université de Rouen

measurable function from X to $[0,1]$, the distribution function associated to f is $F : [0,1] \longrightarrow [0,1]$ defined by

$$F(t) = \mu\{f \le t\}, \quad t \in [0,1].$$

The empirical distribution function of order n associated to f is the function on X defined for each t by,

$$F_n(t) = \frac{1}{n} \sum_{i=0}^{n-1} \mathbb{1}_{[0,t]} \left(f \circ T^i \right).$$

We define the empirical process $\{Y_n(t),\ t \in [0,1]\}_{n\in\mathbb{N}}$ by

$$Y_n(t) = \sqrt{n}\,(F_n(t) - F(t)) = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \left(\mathbb{1}_{[0,t]} \left(f \circ T^i \right) - F(t) \right).$$

We say that the empirical invariance principle holds if $\{Y_n(t),\ t \in [0,1]\}_{n\in\mathbb{N}}$ converges in distribution to a Gaussian process in the space $D([0,1])$ of cadlag functions, provided with the Skorohod topology.

Herein X is a torus and T an ergodic algebraic automorphism. Ergodicity means that for all $A \in \mathcal{A}$, $T^{-1}(A) = A$ implies $\mu(A) = 0$ or 1. For this kind of transformation there are two cases: T can be hyperbolic or quasi-hyperbolic (see next Section).

Actually for hyperbolic torus automorphism, an empirical invariance principle can be deduced from known results about fonctionals of absolutely regular (or β-mixing) processes. One can show that the empirical invariance principle holds for Hölder continuous function f having a Lipschitz continuous distribution function. This is a consequence of Borovkova, Burton and Dehling [3], Theorem 5. To apply this theorem, the strategy is to encode the dynamical system into a stationary process $(Z_i)_{i\in\mathbb{Z}}$ having a β-mixing property. Then the stationary process $(f \circ T^i)_{i\in\mathbb{Z}}$ can be viewed as a functional of $(Z_i)_{i\in\mathbb{Z}}$ and one can show that it is a 1-approximation with coefficients decreasing to zero exponentially fast. By definition, a stationary process $(X_k)_{k\in\mathbb{Z}}$ is a 1-approximating functional of $(Z_k)_{k\in\mathbb{Z}}$ if there exist nonnegative constants $(a_k)_{k\ge0}$ with $a_k \to 0$ as $k \to \infty$ and

$$\mathbb{E}\left| X_0 - \mathbb{E}(X_0 | Z_{-k}, \ldots, Z_k) \right| \le a_k$$

holds for any $k \ge 0$.
This is done by using a well adapted Markov partition of the torus. Indeed, this kind of partition generates a β-mixing process (we say that the partition is weak Bernoulli) with exponential rate (see Bowen [4]) and the 1-approximation is due to the regularity of the Markov partition and the Hölder condition on f. This can be done, in general, for the class of Anosov diffeomorphisms.

In the quasi-hyperbolic case, according to Lind [12], no regular partition of the torus is weak Bernoulli. So the Borovkova, Burton and Dehling theorem ([3]) cannot apply. The question is the following: does the empirical invariance principle hold in this case?

In Section 1, we first prove that the answer is positive under a rather technical condition on f, see Theorem 1. The proof uses a multiple mixing property of ergodic torus automorphisms and works as well in the hyperbolic case.

In fact, the technical condition on the function f holds as soon as f and its distribution function are Hölder continuous, see Theorem 2.

In Section 2, we show that the distribution functions of Morse functions defined on a compact Riemannian manifold are at least Hölder continuous, in fact $\mathcal{C}^1$ as soon as the dimension of the manifold is at least 3, see Theorem 5.

This result is used in Section 1 in order to prove that the set of functions for which the empirical invariance principle holds contains an open and dense subset of $\mathcal{C}^r(\mathbb{T}^d)$ for $r \geq 2$, see Theorem 3.

1 Empirical Invariance Principle

1.1 Main Results

Let $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$ be the d-dimensional torus and let $T : \mathbb{T}^d \longrightarrow \mathbb{T}^d$ be an algebraic ergodic automorphism. The transformation T can be represented by a square matrix with integer coefficients and determinant equal to ± 1. Here, ergodicity is equivalent to the fact that no eigenvalue of T is a root of unity. In that case, it is known that the modulus of at least one eigenvalue is strictly bigger than 1 (and the modulus of another one is strictly smaller than 1). The automorphism T is said to be hyperbolic if no eigenvalue has modulus 1 and quasi-hyperbolic (or partially hyperbolic) otherwise. The following statements and proofs work in both cases.

We denote by μ the Lebesgue measure on $\mathbb{T}^d$ (product measure). The aim is to get a limit theorem for the empirical process associated to a function f defined on the torus. For such a theorem to hold this function should of course verify some regularity conditions.

For any subset A of the torus, the notation ∂A stands for the boundary of A and we are interested by the "ε-boundary" of A:

$$\partial_\varepsilon A := \{x/\mathrm{d}(x, \partial A) \leq \varepsilon\}$$

where d is the Euclidian metric on the torus.

Our main result is the following.

Theorem 1. *Let $f : \mathbb{T}^d \longrightarrow \mathbb{R}$ be a bounded measurable function satisfying the condition*

($*$) *there exist $C, \xi > 0$ such that for all $t \in \mathbb{R}$ and for all $\varepsilon > 0$,*

$$\mu(\partial_\varepsilon \{f \leq t\}) \leq C\varepsilon^\xi.$$

Moreover, its distribution function F is assumed to be continuous.

Then $(f \circ T^i)_{i\in\mathbb{N}}$ verifies the empirical invariance principle:
let $[a, b]$ be a compact interval of $\mathbb{R}$ such that $f(\mathbb{T}^d) \subset [a, b]$,

$$\{Y_n(t) = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \left(\mathbb{1}_{[a,t]}(f \circ T^i) - F(t)\right),\ t \in [a, b]\}_{n\in\mathbb{N}}$$

converges in distribution to a Gaussian process Y in $D([a, b])$. Further, Y is μ-almost surely continuous.

This theorem is proved in Section 1.2.

Next Theorem shows that even in the hyperbolic case, our result is slightly more general than the one deduced from the theorem of Borovkova et al.

Theorem 2. *If $f : \mathbb{T}^d \longrightarrow \mathbb{R}$ is a Hölder continuous function having a Hölder continuous distribution function, then $(f \circ T^i)_{i\in\mathbb{N}}$ satisfies the empirical invariance principle.*

Proof.

Let $f : \mathbb{T}^d \longrightarrow \mathbb{R}$ be a function satisfying the assumptions of Theorem 2 and F its distribution function. We will show that Condition $(*)$ holds.

For $\varepsilon > 0$ and $t \in \mathbb{R}$, the set $\partial_\varepsilon\{f \le t\}$ is the union of the two following disjoint sets:

$$\partial_\varepsilon\{f \le t\}^+ := \partial_\varepsilon\{f \le t\} \cap \{f > t\},$$

$$\partial_\varepsilon\{f \le t\}^- := \partial_\varepsilon\{f \le t\} \cap \{f \le t\}.$$

First of all, the continuity of f implies that $\partial\{f \le t\} \subset \{f = t\}$.

Now, for $x \in \partial_\varepsilon\{f \le t\}^+$, $f(x) \le t + K\varepsilon^\xi$, where K and ξ denote the Hölder constants of f. We have

$$\begin{aligned} \mu\left(\partial_\varepsilon\{f \le t\}^+\right) &\le \mu\left(f \le t + K\varepsilon^\xi\right) - \mu\left(f \le t\right) \\ &= F\left(t + K\varepsilon^\xi\right) - F\left(t\right) \\ &\le L(K\varepsilon^\xi)^\zeta \end{aligned}$$

where L and ζ are the Hölder constants of F.

We obtain the same result for $\partial_\varepsilon\{f \le t\}^-$ and thus

$$\mu(\partial_\varepsilon\{f \le t\}) \le 2LK^\zeta\varepsilon^{\xi\zeta}.$$

□

Remark. *The fact that f is Hölder continuous is not enough to get Condition $(*)$, even if F is continuous. A counterexample is given in Section 1.3.*

Now Theorem 2 together with the upcoming result of Section 2 (Theorem 5) leads to the following genericity result.

Theorem 3. *The set of functions f for which $(f \circ T^i)_{i\in\mathbb{N}}$ satisfies the empirical invariance principle contains an open and dense subset of $\mathcal{C}^r(\mathbb{T}^d)$ for $r \ge 2$.*

1.2 Proof of Theorem 1

We will deduce Theorem 1 from the following:

Theorem 4. *Let $f : \mathbb{T}^d \longrightarrow [0,1]$ be a uniformly distributed function such that Condition $(*)$ is satisfied.*

Then $\{Y_n(t) = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \left(\mathbb{1}_{[0,t]}(f \circ T^i) - t\right),\ t \in [0,1]\}_{n\in\mathbb{N}}$ converges in distribution to a μ-almost surely continuous Gaussian process in $D([0,1])$.

Theorem 4 implies Theorem 1.

Let $f : \mathbb{T}^d \longrightarrow \mathbb{R}$ be a bounded function verifying Condition $(*)$ with a continuous distribution function F. We can assume $f(\mathbb{T}^d) \subset [0,1]$ without loss of generality.

Let us defined as usual F^{-1} by $F^{-1}(t) = \sup\{s \in [0,1]/F(s) \leq t\}$, $t \in [0,1]$. Then $F \circ F^{-1} = \mathrm{id}$, thanks to the continuity of F. Let us consider $g := F \circ f$. We have

$$\{g \leq t\} = \{F \circ f \leq t\} = \{f \leq F^{-1}(t)\}.$$

thanks again to the continuity of F. From this equality we deduce that g is uniformly distributed and satisfies Condition $(*)$ (with the same constants as f).

Let us denote by Y_n and Z_n the empirical processes respectively defined by f and g. For each $t \in [0,1]$, the equality $Y_n(t) = Z_n \circ F(t)$ holds μ-almost surely, because

$$\mu \left\{ \mathbb{1}_{[0,t]}(f) = \mathbb{1}_{[0,F(t)]}(g) \right\} = 1.$$

Moreover the equality $\mathbb{1}_{[0,t]}(f) = \mathbb{1}_{[0,F(t)]}(g)$ holds as soon as F^{-1} does not have a jump at t. Therefore the process equality $Y_n = Z_n \circ F$ holds μ-almost surely.

By Theorem 4, Z_n converges in distribution to an almost surely continuous Gaussian process Z. We recall the following result (see Theorem 5.1 in Billingsley [2]):

Let $\psi : D([0,1]) \longrightarrow D([0,1])$ and D_ψ the set of discontinuities of ψ. If a sequence of random variables h_n converges in distribution to a variable h such that $\mu(h \in D_\psi) = 0$, then $\psi(h_n)$ converges in distribution to $\psi(h)$.

Let us consider the mapping $\psi : h \longmapsto h \circ F$ from $D([0,1])$ to $D([0,1])$. It is continuous on $C([0,1])$ for the induced topology which is the topology of the uniform convergence (see Billingsley [2] p.112). The theorem of Billingsley applied to ψ gives the result for Y_n and Theorem 1 is proved.

$\square$

Proof of Theorem 4.

We consider a uniformly distributed function $f : \mathbb{T}^d \longrightarrow [0,1]$ (i.e. $F \equiv Id_{[0,1]}$).

To get the invariance principle, the proof consists in two steps. First we get the finite-dimensional Central Limit Theorem and then we show the tightness of the process.

The following proposition was proved by Le Borgne in [9].

Proposition 1. *Let $T : \mathbb{T}^d \longrightarrow \mathbb{T}^d$ be an ergodic torus automorphism. If a measurable set A verifies:*

$$\exists C, \xi > 0,\ \forall \varepsilon > 0,\ \mu\left(\partial_\varepsilon A\right) \leq C\varepsilon^\xi,$$

then the Central Limit Theorem holds for the function $\mathbb{1}_A - \mu(A)$.

For each $t \in [0,1]$, we denote by φ_t the function

$$\varphi_t = \mathbb{1}_{\{f \leq t\}} - t.$$

Applying Proposition 1, we get:

Proposition 2. *If a function f satisfies Condition $(*)$, then for all $t \in [0,1]$, φ_t satisfies the Central Limit Theorem. i.e.*

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi_t \circ T^i \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0,\sigma).$$

It follows easily that the same is true for the finite dimensional case.

Proposition 3. *If a function f satisfies Condition $(*)$, then for all $k \in \mathbb{N}$ and $(t_1, ..., t_k) \in [0,1]^k$, $(Y_n(t_1), ..., Y_n(t_k))$ converges in distribution to a k-dimensional Gaussian vector.*

Now it remains to show that the process $\{Y_n(t), t \in [0,1]\}_{n\in\mathbb{N}}$ is tight. Following Billingsley [2], it is sufficient to prove:

$\forall \varepsilon > 0$, $\forall \eta > 0$, $\exists \zeta \in]0,1[$, $\exists n_0 \in \mathbb{N}$ such that $\forall n \geq n_0$,

$$\mu\left(\sup_{|t-s|\leq\zeta} |Y_n(t) - Y_n(s)| \geq \varepsilon\right) \leq \eta. \tag{1}$$

The first step is to establish the following lemma.

Lemma 1. *There exist $C > 0$ and $\delta \in (0,1)$ such that for all $s < t \in [0,1]$, if $\varphi = \varphi_t - \varphi_s$,*

$$\mathbb{E}\left(\left(\sum_{i=1}^{n} \varphi \circ T^i\right)^4\right) \leq C\left(n^2 (t-s)^{\frac{2}{1+\delta}} + n (t-s)^{\frac{1}{1+\delta}}\right).$$

To prove Lemma 1, we will use a multiple mixing property of ergodic torus automorphisms. A version of the following proposition (in the case of diagonal flows and for Hölder observables) can be found in the paper by Le Borgne [10] (see also Le Borgne and Pène [11]). Here, it is adapted to the case of ergodic torus automorphisms and for $\mathcal{C}^1$ functions. We do not give a proof because it follows exactly the one of [11].

Proposition 4. *$\exists C > 0$, $\exists \alpha > 0$, $\forall m$, $p \in \mathbb{N}^*$, $\forall \phi_1, ..., \phi_m, \psi_1, ..., \psi_p \in \mathcal{C}^1(\mathbb{T}^d)$, $\forall k_1 \leq ... \leq k_m \leq 0 \leq l_1 \leq ... \leq l_p$, $\forall n \in \mathbb{N}$,*

$$\left|\operatorname{cov}\left(\prod_{j=1}^{m} \phi_j \circ T^{k_j}, \prod_{j=1}^{p} \psi_j \circ T^{l_j+n}\right)\right|$$
$$\leq C\left(\sum_{j=1}^{m} \prod_{i\neq j} \|\phi_i\|_\infty \|\phi_j\|_{\mathcal{C}^1} |k_j|^r\right)\left(\sum_{j=1}^{p} \prod_{i\neq j} \|\psi_i\|_\infty \|\psi_j\|_{\mathcal{C}^1}\right) \mathrm{e}^{-\alpha n}$$

where r is the size of the bigger Jordan's block of T restricted to its neutral subspace.

In Lemma 1, φ is a discontinuous function. We will use the fact that the boundary of the set $\{s < f \leq t\}$ is enough regular (Condition $(*)$) to approximate φ by $\mathcal{C}^1$ functions.
In the sequel, C always denotes a constant, but its value may change.

Lemma 2. *Under Condition* $(*)$, *there exists* $C > 0$, *for all* $\beta > 0$, *for all* $s, t \in [0, 1]$, *there exists a sequence* $(\varphi_k)_{k\in\mathbb{N}}$ *of* $\mathcal{C}^1$ *functions such that*

$$\|\varphi_k\|_{\mathcal{C}^1} \le C\mathrm{e}^{\beta k} \quad \text{and} \quad \|\varphi - \varphi_k\|_{L^p} \le C\mathrm{e}^{-\frac{\gamma k}{p}}, \quad \forall p \ge 1$$

where $\gamma = \frac{\xi\beta}{d+1}$ *and* $\varphi = \varphi_t - \varphi_s$.

Proof.

The considered balls are defined with respect to the Euclidian norm of $\mathbb{R}^d$ which is inducted on the torus by identifying $\mathbb{T}^d$ to $[0, 1]^d$. Let $\rho : \mathbb{T}^d \longrightarrow [0, +\infty)$ be a $\mathcal{C}^1$ function such that $\mathbb{E}(\rho) = 1$ and ρ equals 0 outside $B(0, \frac{1}{2})$.

Write

$$\rho_k(x) = \begin{cases} \mathrm{e}^{\frac{\beta k d}{d+1}}\rho(\mathrm{e}^{\frac{\beta k}{d+1}}x) & \text{if } x \in B(0, \frac{1}{2}\mathrm{e}^{-\frac{\beta k}{d+1}}) \\ 0 & \text{else} \end{cases} .$$

Then ρ_k is $\mathcal{C}^1$, $\rho_k^{-1}((0, +\infty)) \subset B(0, \frac{1}{2}\mathrm{e}^{-\frac{\beta k}{d+1}})$, $\mathbb{E}(\rho_k) = 1$ and for $i = 1, \dots, d$,

$$\left\|\frac{\partial \rho_k}{\partial x_i}\right\|_\infty \le \mathrm{e}^{\frac{\beta k d}{d+1}}\mathrm{e}^{\frac{\beta k}{d+1}}\left\|\frac{\partial \rho}{\partial x_i}\right\|_\infty \le C\mathrm{e}^{\beta k}.$$

Write $\varphi_k = \varphi * \rho_k$. Then

$$\|\varphi_k\|_{\mathcal{C}^1} = \|\varphi_k\|_\infty + \max_{i=1}^{d}\left\|\varphi * \frac{\partial \rho_k}{\partial x_i}\right\|_\infty \le 1 + \max_{i=1}^{d}\left\|\frac{\partial \rho_k}{\partial x_i}\right\|_\infty \le C\mathrm{e}^{\beta k}$$

and

$$\begin{aligned} \|\varphi_k - \varphi\|_p^p = \mathbb{E}\left(|\varphi_k - \varphi|^p\right) &\le \mathbb{E}|\varphi_k - \varphi| \\ &\le \|\varphi_k - \varphi\|_\infty \, \mu\left(\partial_{\frac{1}{2}\mathrm{e}^{-\frac{\beta k}{d+1}}}\{f \in (s, t]\}\right) \\ &\le C\mathrm{e}^{-\frac{\beta k}{d+1}\xi} \\ &= C\mathrm{e}^{\gamma k}. \end{aligned}$$

□

The function φ is always bounded by 1. From this, we deduce the following remark which is useful in the proof of Lemma 1.

Remark. *For all* $p \ge 1$, $\mathbb{E}(|\varphi|^p) \le 2|t - s|$.

Proof of Lemma 1.

Developing the term $\mathbb{E}\left(\left(\sum_{i=1}^n \varphi \circ T^i\right)^4\right)$, we will have to study terms like

$$\mathbb{E}(\varphi\, \varphi \circ T^i\, \varphi \circ T^{i+j}\, \varphi \circ T^{i+j+k})$$

with $i + j + k \le n$. We fix three integers i, j, k like this. Note that,

$$\mathbb{E}(\varphi\, \varphi \circ T^i\, \varphi \circ T^{i+j}\, \varphi \circ T^{i+j+k}) = \mathrm{cov}(\varphi\, , (\varphi\, \varphi \circ T^j\, \varphi \circ T^{j+k}) \circ T^i).$$

We wish to apply Proposition 4. To do that, we consider the sequence $(\varphi_l)_{l\in\mathbb{N}}$ defined by Lemma 2. We have

$$\begin{aligned}
\mathrm{cov}(\varphi\,,(\varphi\,\varphi\circ T^j\,\varphi\circ T^{j+k})\circ T^i) &= \mathbb{E}(\varphi_i\,(\varphi_i\,\varphi_i\circ T^j\,\varphi_i\circ T^{j+k})\circ T^i) &(2)\\
&+\mathbb{E}((\varphi-\varphi_i)\,(\varphi\,\varphi\circ T^j\,\varphi\circ T^{j+k})\circ T^i) &(3)\\
&+\mathbb{E}(\varphi_i\,((\varphi-\varphi_i)\,\varphi\circ T^j\,\varphi\circ T^{j+k})\circ T^i) &(4)\\
&+\mathbb{E}(\varphi_i\,(\varphi_i\,(\varphi-\varphi_i)\circ T^j\,\varphi\circ T^{j+k})\circ T^i) &(5)\\
&+\mathbb{E}(\varphi_i\,(\varphi_i\,\varphi_i\circ T^j\,(\varphi-\varphi_i)\circ T^{j+k})\circ T^i). &(6)
\end{aligned}$$

We will distinguish two cases.

First case : $|t-s|\geq \mathrm{e}^{-\gamma i}$.

Notice that γ is the constant equal to $\frac{\xi\beta}{d+1}$ that appears in Lemma 2. For the seek of clarity it will be fixed farther.

Let $p,q\geq 1$ such that $\frac{1}{p}+\frac{3}{q}=1$. By Hölder inequality and Lemma 2,

$$\begin{aligned}
|(3)| &\leq \|\varphi-\varphi_i\|_p\|\varphi\|_q^3\\
&\leq C\mathrm{e}^{-\frac{\gamma i}{p}}\|\varphi\|_q^3\\
&\leq C\mathrm{e}^{-\frac{\gamma i}{p}}2^{\frac{3}{q}}|t-s|^{\frac{3}{q}}.
\end{aligned}$$

On the other hand,

$$|(6)|\leq \|\varphi-\varphi_i\|_p\|\varphi_i\|_q^3$$

and

$$\|\varphi_i\|_q\leq \|\varphi-\varphi_i\|_q+\|\varphi\|_q\leq C\mathrm{e}^{-\frac{\gamma i}{q}}+\|\varphi\|_q.$$

Thus by assumption,

$$\begin{aligned}
|(6)| &\leq C\mathrm{e}^{-\frac{\gamma i}{p}}\left[C\mathrm{e}^{-\frac{\gamma i}{q}}+\|\varphi\|_q\right]^3\\
&\leq C\mathrm{e}^{-\frac{\gamma i}{p}}\left[C|t-s|^{\frac{1}{q}}+2^{\frac{1}{q}}|t-s|^{\frac{1}{q}}\right]^3\\
&\leq C\mathrm{e}^{-\frac{\gamma i}{p}}|t-s|^{\frac{3}{q}}.
\end{aligned}$$

In the same way,

$$|(4)|\leq \|\varphi-\varphi_i\|_p\|\varphi_i\|_q\|\varphi\|_q^2\leq C\mathrm{e}^{-\frac{\gamma i}{p}}|t-s|^{\frac{3}{q}}.$$

The same thing is true for (5). We get

$$|(3)+(4)+(5)+(6)|\leq C\mathrm{e}^{-\frac{\gamma i}{p}}|t-s|^{\frac{3}{q}}. \tag{7}$$

Now (2) $=\mathrm{cov}(\varphi_i\,,(\varphi_i\,\varphi_i\circ T^j\,\varphi_i\circ T^{j+k})\circ T^i)$, because $\mathbb{E}(\varphi_i)=0$. So we can apply the multiple mixing inequality. There exist $C,\alpha>0$ which depend only on T, such that

$$\begin{aligned}
|(2)| &\leq C\|\varphi_i\|_{\mathcal{C}^1}.\|\varphi_i\|_\infty^2\|\varphi_i\|_{\mathcal{C}^1}\mathrm{e}^{-\alpha i}\\
&\leq C\mathrm{e}^{2\beta i}\mathrm{e}^{-\alpha i}
\end{aligned}$$

because $\|\varphi_i\|_\infty \leq 1$. As $|t-s| \geq \mathrm{e}^{-\gamma i}$,

$$|(2)| \leq C\mathrm{e}^{-\eta i}|t-s|^{\frac{3}{q}} \tag{8}$$

where $\eta = \alpha - 2\beta - \frac{3\gamma}{q}$.

β being arbitrary, we can choose $\beta > 0$ such that $\frac{\xi+2(d+1)}{d+1}\beta < \frac{\alpha}{2}$. Then, for all $q \geq 3$, $\eta > \frac{\alpha}{2} > 0$. Now, we can fix $\delta \in (0,1)$ and choose $p = \frac{1+\delta}{\delta}$ (so, $q = 3+3\delta$) in order to have $\frac{\gamma}{p} < \eta$. This is possible because $p \to \infty$ when $\delta \to 0$. We get $\mathrm{e}^{-\eta} \leq \mathrm{e}^{-\frac{\gamma}{p}}$ and by (7) and (8),

$$\begin{aligned} |\mathbb{E}(\varphi\,(\varphi\,\varphi\circ T^j\,\varphi\circ T^{j+k})\circ T^i)| &\leq C|t-s|^{\frac{3}{q}}\mathrm{e}^{-\frac{\gamma i}{p}} \\ &= C|t-s|^{\frac{1}{1+\delta}}\mathrm{e}^{-\frac{\gamma\delta i}{1+\delta}}. \end{aligned}$$

Second case : $|t-s| < \mathrm{e}^{-\gamma i}$.
By Hölder inequality with p and q previously fixed,

$$\begin{aligned} |\mathbb{E}(\varphi\,(\varphi\,\varphi\circ T^j\,\varphi\circ T^{j+k})\circ T^i)| &\leq \|\varphi\|_p\|\varphi\|_q^3 \\ &\leq 2^{\frac{1}{p}}|t-s|^{\frac{1}{p}}2^{\frac{3}{q}}|t-s|^{\frac{3}{q}} \\ &\leq C\mathrm{e}^{-\frac{\gamma i}{p}}|t-s|^{\frac{3}{q}} \\ &= C|t-s|^{\frac{1}{1+\delta}}\mathrm{e}^{-\frac{\gamma\delta i}{1+\delta}}. \end{aligned}$$

In each case, we have

$$|\mathbb{E}(\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j}\,\varphi\circ T^{i+j+k})| \leq C|t-s|^{\frac{1}{1+\delta}}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}i}. \tag{9}$$

In the same way, playing with k instead of i, we get

$$\begin{aligned} |\mathbb{E}(\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j}\,\varphi\circ T^{i+j+k})| &= |\mathrm{cov}((\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j})\,,\varphi\circ T^{i+j}\circ T^k)| \\ &\leq C|t-s|^{\frac{1}{1+\delta}}(1+(i+j)^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}k} \end{aligned} \tag{10}$$

where the term $(i+j)^r$ is related to the neutral subspace of T and appears in application of Proposition 4.

Now, with j, there is a remainder term:

$$\begin{aligned} &|\mathbb{E}(\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j}\,\varphi\circ T^{i+j+k})| \\ &= |\mathrm{cov}((\varphi\,\varphi\circ T^i), (\varphi\circ T^i\,\varphi\circ T^{i+k})\circ T^j) + \mathbb{E}(\varphi\,\varphi\circ T^i)\mathbb{E}(\varphi\,\varphi\circ T^k)| \\ &\leq C|t-s|^{\frac{1}{1+\delta}}(1+i^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}j} + |\mathbb{E}(\varphi\,\varphi\circ T^i)\mathbb{E}(\varphi\,\varphi\circ T^k)| \end{aligned}$$

Again the term i^r appears in application of Proposition 4.

So, we study

$$\begin{aligned} \mathbb{E}(\varphi\,\varphi\circ T^i) &= \mathbb{E}(\varphi_i\,\varphi_i\circ T^i) \\ &\quad + \mathbb{E}((\varphi-\varphi_i)\,\varphi\circ T^i) \\ &\quad + \mathbb{E}(\varphi_i\,(\varphi-\varphi_i)\circ T^i). \end{aligned}$$

By Proposition 4 (or by the exponential mixing inequality (see [12])),

$$|\mathbb{E}(\varphi_i\,\varphi_i\circ T^i)| \le C\|\varphi_i\|_{\mathcal{C}^1}^2 \mathrm{e}^{-\alpha i} \le C\mathrm{e}^{2\beta i-\alpha i}.$$

As above, considering two cases and keeping the same notations, we have

$$|\mathbb{E}(\varphi\,\varphi\circ T^i)| \le C|t-s|^{\frac{1}{1+\delta}}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}i}$$

and the same inequality holds for k. Thus we get

$$\begin{aligned} |\mathbb{E}(\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j}\,\varphi\circ T^{i+j+k})| \ &\le\ C|t-s|^{\frac{1}{1+\delta}}(1+(i+k)^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}j} \\ &\quad + C|t-s|^{\frac{2}{1+\delta}}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}(i+k)}. \end{aligned} \tag{11}$$

Now we can obtain the desired majoration.

$$\begin{aligned} \mathbb{E}\left(\left(\sum_{i=1}^n \varphi\circ T^i\right)^4\right) &\le 4!n \sum_{\{i,j,k: i+j+k\le n\}} |\mathbb{E}(\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j}\,\varphi\circ T^{i+j+k})| \\ &\le 4!n\left[\sum_{i=1}^n\sum_{j,k\le i} |\mathbb{E}(\varphi\,(\varphi\,\varphi\circ T^j\,\varphi\circ T^{j+k})\circ T^i)|\right. \\ &\quad + \sum_{j=1}^n\sum_{i,k\le j} |\mathbb{E}((\varphi\,\varphi\circ T^i)(\varphi\circ T^i\,\varphi\circ T^{i+k})\circ T^j)| \\ &\quad \left. + \sum_{k=1}^n\sum_{i,j\le k} |\mathbb{E}((\varphi\,\varphi\circ T^i\,\varphi\circ T^{i+j})\,\varphi\circ T^{i+j}\circ T^k)|\right] \\ &= 4!n\,[\quad I_1 \quad + \quad I_2 \quad + \quad I_3 \quad]. \end{aligned}$$

By (9),

$$\begin{aligned} I_1 &\le C\left[\sum_{i=1}^n\sum_{j,k\le i}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}i}\right]|t-s|^{\frac{1}{1+\delta}} \\ &\le C\left[\sum_{i=1}^n i^2\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}i}\right]|t-s|^{\frac{1}{1+\delta}} \\ &\le C|t-s|^{\frac{1}{1+\delta}} \end{aligned}$$

because the series is convergente.

By (10),

$$\begin{aligned} I_3 &\le C\left[\sum_{k=1}^n\sum_{i,j\le k}(1+(i+j)^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}k}\right]|t-s|^{\frac{1}{1+\delta}} \\ &\le C\left[\sum_{k=1}^n k^2(1+(2k)^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}k}\right]|t-s|^{\frac{1}{1+\delta}} \\ &\le C|t-s|^{\frac{1}{1+\delta}}. \end{aligned}$$

Finally, by (11),

$$\begin{aligned}
I_2 &\le C\left[\sum_{j=1}^{n}\sum_{i,k\le j}(1+i^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}j}\right]|t-s|^{\frac{1}{1+\delta}} + C\left[\sum_{j=1}^{n}\sum_{i,k\le j}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}(i+k)}\right]|t-s|^{\frac{2}{1+\delta}} \\
&\le C\left[\sum_{j=1}^{n}j^2(1+j^r)\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}j}\right]|t-s|^{\frac{1}{1+\delta}} + C\left[n\left(\sum_{i=1}^{n}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}i}\right)\left(\sum_{k=1}^{n}\mathrm{e}^{-\frac{\gamma\delta}{1+\delta}k}\right)\right]|t-s|^{\frac{2}{1+\delta}} \\
&\le C|t-s|^{\frac{1}{1+\delta}} + Cn|t-s|^{\frac{2}{1+\delta}}
\end{aligned}$$

because, again, the series are convergente.

In conclusion, there exists $\delta > 0$ such that for all $s, t \in [0,1]$,

$$\mathbb{E}\left(\left(\sum_{i=1}^{n}\varphi\circ T^i\right)^4\right) \le C\left(n^2|t-s|^{\frac{2}{1+\delta}} + n|t-s|^{\frac{1}{1+\delta}}\right).$$

□

Lemma 1 is the key inequality for the sequel. Now, the method leading to the tightness of the process $(Y_n)_{n\in\mathbb{N}}$ is the classical chaining one. The reader can see, for more details, the paper by Dehling and Philipp [7]. For convenience, we expose here the main steps of the method.

To check Condition (1), we study $|Y_n(t) - Y_n(s)|$ for $s \le t \le s+\zeta$ where ζ is to be determined. The idea is to introduce a subdivision of step h of the interval, where h is also to be determined. Next Lemma is proved in Billingsley [2].

Lemma 3. *$\forall h \in [0,1]$ and $\forall s \le t \le s+h$,*

$$|Y_n(t) - Y_n(s)| \le |Y_n(s+h) - Y_n(s)| + h\sqrt{n}.$$

We deduce

Lemma 4. *$\forall s, h \in [0,1)$ and $\forall m \in \mathbb{N}$ such that $s + mh \le 1$,*

$$\sup_{s\le t\le s+mh}|Y_n(t) - Y_n(s)| \le 3\max_{i\le m}|Y_n(s+ih) - Y_n(s)| + h\sqrt{n}.$$

The proof is left to the reader.

Now, assume that $h < \frac{\varepsilon}{\sqrt{n}}$. Then Lemma 4 gives

$$\begin{aligned}
\mu\left(\sup_{s\le t\le s+mh}|Y_n(t) - Y_n(s)| \ge 4\varepsilon\right) &\le \mu\left(3\max_{i\le m}|Y_n(s+ih) - Y_n(s)| + h\sqrt{n} \ge 4\varepsilon\right) \\
&\le \mu\left(\max_{i\le m}|Y_n(s+ih) - Y_n(s)| \ge \varepsilon\right). \qquad (12)
\end{aligned}$$

Lemma 1 gives: $\forall s, t \in [0,1]$,

$$\mathbb{E}\left((Y_n(t) - Y_n(s))^4\right) \le C\left(|t-s|^{\frac{2}{1+\delta}} + \frac{1}{n}|t-s|^{\frac{1}{1+\delta}}\right).$$

Applying this with $s+ih$ and $s+(i+k)h$, we get

$$\mathbb{E}\left((Y_n(s+(i+k)h) - Y_n(s+ih))^4 \right) \leq C\left((kh)^{\frac{2}{1+\delta}} + \frac{1}{n}(kh)^{\frac{1}{1+\delta}} \right).$$

We can choose $h \geq \frac{\varepsilon}{n^{1+\delta}}$ (for n large enough). Thus, we get $h^{\frac{1}{1+\delta}} \geq \frac{\varepsilon^{\frac{1}{1+\delta}}}{n} \geq \frac{\varepsilon}{n}$ ($\varepsilon < 1$) and

$$\mathbb{E}\left((Y_n(s+(i+k)h) - Y_n(s+ih))^4 \right) \leq C\left((kh)^{\frac{2}{1+\delta}} + \frac{(kh)^{\frac{2}{1+\delta}}}{\varepsilon} \right) \leq \frac{2C}{\varepsilon}(kh)^{\frac{2}{1+\delta}}.$$

According to Billingsley [2], Theorem 12.2,

$$\mu\left(\max_{i\leq m} |Y_n(s+ih) - Y_n(s)| \geq \varepsilon \right) \leq \frac{C(mh)^{\frac{2}{1+\delta}}}{\varepsilon^5}. \tag{13}$$

From (12) and (13), we derive

$$\mu\left(\sup_{s\leq t\leq s+mh} |Y_n(t) - Y_n(s)| \geq 4\varepsilon \right) \quad \leq \quad \frac{C(mh)^{\frac{2}{1+\delta}}}{\varepsilon^5}. \tag{14}$$

In conclusion, let ε, η be fixed, δ defined by Lemma 1, and ζ belonging to $[0,1]$ such that $\zeta < \left(\frac{\eta\varepsilon^5}{C}\right)^{\frac{1+\delta}{2}}$. Pick n_0 large enough to verify

$$\frac{\zeta n_0^{1+\delta}}{\varepsilon} - \frac{\zeta\sqrt{n_0}}{\varepsilon} \geq 1.$$

Then for all $n \geq n_0$, there exist $m \in \mathbb{N}$ and $h \in (0,1)$ such that

$$\frac{\varepsilon}{n^{1+\delta}} \leq h < \frac{\varepsilon}{\sqrt{n}} \qquad \text{and} \qquad \zeta = mh.$$

Finally, (14) implies

$$\mu\left(\sup_{s\leq t\leq s+\zeta} |Y_n(t) - Y_n(s)| \geq 4\varepsilon \right) \leq \frac{C}{\varepsilon^5}\zeta^{\frac{2}{1+\delta}} \leq \eta.$$

The tightness condition (1) is verified and the result is proved.

□

1.3 Counterexample

We have shown in Theorem 2 that Condition $(*)$ holds as soon as the function f and its distribution function F are Hölder continuous. In this section, we prove that the fact that f is Hölder continuous function is not enough to get Condition $(*)$, even if F is continuous.

We define :

$\forall k \geq 1$,

$$s_k = 1 - \frac{2}{2^k} \quad \text{et} \quad S_k = 1 - \frac{1}{2^k}$$

and the intervals $I_k = [s_k, S_k]$, $k \geq 1$.

Further, for each $k \geq 2$, we write

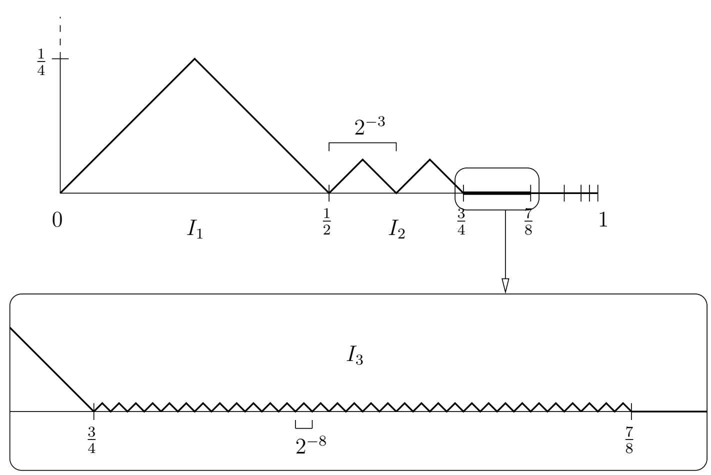
$$s^l_{(k)} = s_k + \frac{l}{2^{k^2}} = \frac{2^{k^2} - 2^{k^2-k+1} + l}{2^{k^2}}, \quad 0 \leq l \leq 2^{k^2-k}.$$

So, we have, $s_k = s^0_{(k)}$ and $S_k = s^{k^2-k}_{(k)}$.

Let $f : [0,1] \longrightarrow [0,1]$ be defined as follows:

$$\begin{cases} \forall x \in [0, \frac{1}{4}], \quad f(x) = x; \\ \forall x \in [\frac{1}{4}, \frac{1}{2}], \quad f(x) = \frac{1}{2} - x; \\ \forall k \geq 2, \qquad \forall 0 \leq l \leq 2^{k^2-k}, \qquad f(s^l_{(k)}) = \begin{cases} 0 & \text{if } l \text{ is even} \\ \frac{1}{2^{k^2}} & \text{if } l \text{ is odd} \end{cases} \\ \qquad \text{and } f \text{ affine function on each interval } [s^l_{(k)}, s^{l+1}_{(k)}]. \end{cases}$$

Finally, we obtain a continuous function with $f(1) = 0$. Further, f is continuous on the one dimensional torus.



Hence, f is clearly Lipschitz continuous with a Lipschitz constant equal to 1. Remark that its distribution function is continuous (the preimages by f of each point are at most countable).

We will show that f does not satisfy Condition $(*)$.

Let $\varepsilon_k = \frac{1}{2^{k^2}}$, $k \geq 2$. Then, by definition of f, for all $k \geq 2$,

$$I_k \subset \partial_{\varepsilon_k}\{f = 0\}.$$

Thus,

$$\mu\left(\partial_{\varepsilon_k}\{f = 0\}\right) \geq \frac{1}{2^k} = 2^k \varepsilon_k^{\frac{2}{k}}.$$

Now, for all $C > 0$ and $\xi > 0$, there exists an integer $k_0 \geq 2$ such that $2^{k_0} > C$ and $\frac{2}{k_0} < \xi$. So, there exist $\varepsilon = \varepsilon_{k_0}$ and $t = 0$ such that

$$\mu\left(\partial_\varepsilon\{f \leq t\}\right) > C\varepsilon^\xi.$$

To get a counterexample in dimension $d > 1$, it is enough to take

$$\begin{array}{rccl} g: & [0,1]^d & \longrightarrow & [0,1] \\ & (x_1, \ldots, x_d) & \longmapsto & \frac{1}{d}(f(x_1) + \cdots + f(x_d)). \end{array}$$

2 Regularity of the distribution functions of Morse functions

In Theorem 2 the empirical invariance principle is stated under the assumption that the distribution function of f is Hölder continuous. For this reason it is of interest to determine a class as large as possible of functions that present some characteristics allowing the study of the regularity properties of their distribution functions.

The class of Morse functions seems to be a good candidate. Let us recall that a $\mathcal{C}^r$ function f, with $r \geq 2$, is a Morse function if

1. its critical points, or singularities, that is the points where the differential vanishes, are isolated;

2. at each critical point the Hessian, which is well defined (see Hirsch [8] or Milnor [13]), is a non degenerate quadratic form.

According to the well known Morse Lemma, it is possible to find local coordinates in a neighbourhood of a critical point such that f can be written

$$f(x_1, \ldots, x_d) = x_1^2 + x_2^2 + \cdots + x_p^2 - x_{p+1}^2 - x_{p+2}^2 - \cdots - x_d^2$$

with $d = \dim M$ since the critical point is not degenerate.

On the one hand these characteristics are sufficient to study the behaviour of the distribution function of f. On the other hand the set of Morse functions is open and dense in the set of $\mathcal{C}^r$ functions, with $r \geq 2$ or $r = +\infty$, for the $\mathcal{C}^r$ convergence topology.

In the statement of next Theorem we consider a compact Riemannian manifold instead of a torus because this generalization does not change anything in the proof.

The measure on the compact Riemannian manifold M is of course the natural Lebesguian one defined by its metric (it exists even if M is not orientable, see Berger-Gostiaux [1]). We will assume that the volume of M is equal to 1 but this does not matter. Notice that the torus, viewed as $\mathbb{R}^d/\mathbb{Z}^d$, is a compact Riemannian manifold, the measure of which is the Lebesgue measure on $[0,1]^d$.

2.1 Main result

Theorem 5. *Let M be a compact, d-dimensional Riemannian manifold. The natural measure is denoted by μ and we assume $\mu(M) = 1$.*

Let f be a Morse function on M and F its distribution function: $F(a) = \mu\{f \le a\}$.

1. $\dim M = 1$. *The distribution function F is $\frac{1}{2}$-Hölder continuous and is $\mathcal{C}^1$ outside the singular values of f.*

2. $\dim M = 2$. *If f has some hyperbolic singularities, then F is $\frac{1}{2}$-Hölder continuous. Otherwise, F is piecewise $\mathcal{C}^1$ hence Lipschitz. In any case it is $\mathcal{C}^1$ outside the singular values of f.*

3. $\dim M \ge 3$. *The distribution function F is $\mathcal{C}^1$.*

This theorem will be proved in next Section. The following corollary is straightforward:

Corollary 1. *The set of $\mathcal{C}^r$ functions whose distribution functions are $\mathcal{C}^1$ (resp. $\frac{1}{2}$-Hölder continuous) contains an open and dense subset of $\mathcal{C}^r(M;\mathbb{R})$ for $\dim M \ge 3$ (resp. $\dim M = 1, 2$).*

The dimension 2 is particular. On the sphere S^2 it is easy to find functions without hyperbolic singularities, hence functions whose all critical points are extrema: consider for example the "height" function. According to Theorem 5 their distribution functions are piecewise $\mathcal{C}^1$ and therefore Lipschitz. Moreover the set of these functions is open (but not dense!) in $\mathcal{C}^r(M)$ because the non degenerate singularities are stable.

On the opposite any Morse function on the 2-dimensional torus has at least two hyperbolic singularities. This intuitive fact can be proved by considering the index of the gradient of f, see for instance Hirsch [8]. Therefore the distribution function of a Morse function on the torus $\mathbb{T}^2$ is never Lipschitz and we can state:

The set of $\mathcal{C}^r$ functions on the 2-dimensional torus $\mathbb{T}^2$ whose distribution functions are Lipschitz is contained in a closed subset with empty interior of $\mathcal{C}^r(\mathbb{T}^2;\mathbb{R})$ for $r \ge 2$.

2.2 Proof of Theorem 5

Let us begin by some notations. A point of M is usually denoted by x. With a clear abuse of notation the local coordinates for the same point will be $x = (x_1, x_2, \ldots, x_d)$ or even $(x, y) = (x_1, \ldots, x_p, y_1, \ldots, y_q)$ when we will identify $\mathbb{R}^d$ with the product $\mathbb{R}^p \times \mathbb{R}^q$.

Around the hyperbolic singularities we will consider the domain $D_{p,q}(r)$ of $\mathbb{R}^p \times \mathbb{R}^q$ defined as the intersection of

$$\{\|y\| \le r, \quad \|x\|^2 \le \|y\|^2 + r^2\} \cup \{\|x\| \le r, \quad \|y\|^2 \le \|x\|^2 + r^2\}$$

and

$$\{\|x\| \, \|y\| \le r^2\}.$$

There are two reasons to define these rather complicated sets. The first one is that $D_{p,q}(r)$ has to be symmetric with respect to x and y. Secondly we need the boundary of $D_{p,q}(r)$ to be parallel to the gradient of f near the set $\{\|x\| = \|y\|\}$. In these local coordinates f is given by

$$f(x, y) = x_1^2 + x_2^2 + \cdots + x_p^2 - y_1^2 - y_2^2 - \cdots - y_q^2 + a = \|x\|^2 - \|y\|^2 + a.$$

and one can verify that grad f is parallel to the hypersurface $\{\|x\|\,\|y\| = r^2\}$.

To finish the following change of coordinates is used several times:

$$\begin{array}{rcl} \mathbb{R}_+^* \times S^{d-1} & \longmapsto & \mathbb{R}^d \\ (s,z) & \longmapsto & sz \end{array}$$

where S^{d-1} stands for the $(d-1)$-dimensional sphere. Let us denote by ϕ this diffeomorphism, by dx the canonical volume of $\mathbb{R}^d$, and by σ the induced volume on S^{d-1}. Then we have

$$\phi^* dx = s^{d-1} ds \wedge \sigma$$

or, in other words, we replace dx by $s^{d-1}ds \wedge \sigma$ in the integrals (see Berger-Gostiaux [1]). In the proof of Lemma 6 we will do two changes of variable of this kind and will write σ_{d-1} instead of σ to avoid confusion.

In order to prove Theorem 5 we will first consider the regular values of f and then the singular ones. If a is a singular value we can assume without loss of generality that the set $\{f = a\}$ contains an unique singularity x_0, according to the fact that the singularities are isolated. We will distinguish the case where x_0 is an extremum, and the one where x_0 is a hyperbolic singularity.

First case: a is a regular value of f.

We proceed as in [13]. Let $r > 0$ such that the set $\{a - 2r \le f \le a + 2r\}$ does not contain any singularity. Let X be a $\mathcal{C}^1$ vector field on M that verifies

$$\forall x \in \{a - r \le f \le a + r\} \qquad X(x) = \frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|^2}$$

and vanishes outside the set $\{a - 2r \le f \le a + 2r\}$. The vector field X is complete, and we denote its flow by Φ_t. For all $x \in \{a - r \le f \le a + r\}$ we have

$$\frac{d}{dt} f(\Phi_t(x)) = < \operatorname{grad} f, \frac{\operatorname{grad} f}{\|\operatorname{grad} f\|^2} > |_{\Phi_t(x)} = 1,$$

and therefore $f(\Phi_t(x)) = f(x) + t$, as long as $\Phi_t(x) \in \{a - r \le f \le a + r\}$.

Consequently we have

$$\{f \le a + t\} = \Phi_t\{f \le a\}$$

for $|t| \le r$. Then

$$F(a+t) = \mu(\{f \le a + t\}) = \mu(\Phi_t\{f \le a\}) = \int_{\{f \le a\}} |J\Phi_t| \, d\mu.$$

where $J\Phi_t(x)$ stands for the Jacobian of Φ_t evaluated between orthonormal basis at the points x and $\Phi_t(x)$.

Now f is $\mathcal{C}^2$, the vector field X and its flow Φ_t are $\mathcal{C}^1$. Thus $J\Phi_t$ is continuously derivable with respect to t and does not vanish. As a conclusion the distribution function F is $\mathcal{C}^1$ on $]a - r, a + r[$ and

$$F'(a+t) = \int_{\{f \le a\}} \frac{\partial}{\partial t} |J\Phi_t| \, d\mu.$$

Second case: $f(x_0) = a$ is a local extremum.

We can assume without loss of generality that x_0 is a minimum.

Let (U, φ) be a local card centered at x_0 such that $\varphi(U) = B(0, 2r)$, where $B(0, 2r)$ is the the ball of radius $2r > 0$ in $\mathbb{R}^d$, and f can be written

$$f(x) = x_1^2 + x_2^2 + \cdots + x_d^2 + a = \|x\|^2 + a.$$

We can assume that x_0 is the unique singularity of the set $\{a - 2r \le f \le a + 2r\}$. Let us consider the manifold with boundary $N = M \setminus U$. If the set $\{a - 2r \le f \le a + 2r\} \cap N$ is not empty we can build a vector field X on N as in the previous case which verifies

$$\forall x \in \{a - r \le f \le a + r\} \bigcap N \qquad X(x) = \frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|^2}$$

and is zero outside of $\{a - 2r \le f \le a + 2r\}$. The previous proof shows that

$$F_1(a + t) = \mu(\{f \le a + t\} \cap N)$$

is $\mathcal{C}^1$ for $t \in]-r, r[$. Now let us consider the function

$$F_2(a + t) = \mu(\{f \le a + t\} \cap U) = \mu(\varphi^{-1}(B(0, \sqrt{t}))).$$

We have

$$F_2(a + t) = \int_{B(0,\sqrt{t})} \left|J\varphi^{-1}(x)\right| dx$$

where $J\varphi^{-1}(x)$ stands for the Jacobian of φ^{-1} evaluated between orthonormal basis at the points x and $\varphi^{-1}(x)$. According to Lemma 5, with $\theta = \left|J\varphi^{-1}\right|$, this function is $\mathcal{C}^1$ (resp. piecewise $\mathcal{C}^1$) (resp. $\frac{1}{2}$-Hölder continuous) for $-r < t < r$ if $d \ge 3$ (resp. $d = 2$) (resp. $d = 1$) and so is $F = F_1 + F_2$.

Third case: $f(x_0) = a$ is a hyperbolic singularity.

Let (U, φ) be a local card centered at x_0 such that $\varphi(U) = D_{p,q}(2r)$, and such that in these coordinates f is given by

$$f(x, y) = x_1^2 + x_2^2 + \cdots + x_p^2 - y_1^2 - y_2^2 - \cdots - y_q^2 + a = \|x\|^2 - \|y\|^2 + a.$$

We can assume that x_0 is the unique singularity of the set $\{a - 2r \le f \le a + 2r\}$. Let us consider the manifold with corners $N = M \setminus U$.

The vector field $\operatorname{grad} f$ is tangent to the boundary of N at any point (x, y) of this boundary such that $\|x\| \, \|y\| = 4r^2$. Therefore we can build a vector field X on N that verifies

$$\forall x \in \{a - r \le f \le a + r\} \bigcap N \qquad X(x) = \frac{\operatorname{grad} f(x)}{\|\operatorname{grad} f(x)\|^2}$$

that is zero outside of $\{a - 2r \le f \le a + 2r\}$, and that is tangent to the boundary of N. The function

$$F_1(a + t) = \mu(\{f \le a + t\} \cap N)$$

is as previously $\mathcal{C}^1$ for $t \in]-r, r[$. Now the function

$$F_2(a + t) = \mu(\{f \le a + t\} \cap U)$$

is equal to

$$F_2(a+t) = \int_{D_{p,q}(r)} \left| J\varphi^{-1}(x,y) \right| dxdy.$$

According to Lemma 6 this function is $\mathcal{C}^1$ (resp. $\frac{1}{2}$-Hölder continuous) for $-r < t < r$ if $d \geq 3$ (resp. $d = 2$) and so is $F = F_1 + F_2$.

□

Lemma 5. *Let θ be a $\mathcal{C}^1$ positive map on $B(0,r) \subset \mathbb{R}^d$, the ball of radius $r > 0$, and*

$$V(t) = \int_{B(0,\sqrt{t})} \theta(x)dx.$$

If $d \geq 2$ then V is a $\mathcal{C}^1$ function on $[0, r^2]$ with

$$\begin{array}{ll} V'(0) = 0 & \text{if } d \geq 3 \\ V'(0) = \pi\theta(0) & \text{if } d = 2 \end{array}.$$

If $d = 1$ then V is not derivable at 0 but the function

$$t \longmapsto V(t^2)$$

is $\mathcal{C}^1$ on $[0, r]$.

Proof.

Let us do the change of variable

$$\begin{array}{rcl} \mathbb{R}_+^* \times S^{d-1} & \longmapsto & \mathbb{R}^d \\ (s,z) & \longmapsto & sz \end{array}.$$

We have

$$V(t) = \int_{[0,\sqrt{t}]\times S^{d-1}} \theta(sz)s^{d-1}ds \wedge \sigma = \int_0^{\sqrt{t}} s^{d-1}ds \int_{S^{d-1}} \theta(sz)\sigma = \int_0^{\sqrt{t}} s^{d-1}\omega(s)ds$$

where $\omega(s)$ is the $\mathcal{C}^1$ function defined by

$$\omega(s) = \int_{S^{d-1}} \theta(sz)\sigma.$$

Whenever $d \geq 2$ we set $u = s^2$ and we obtain

$$V(t) = \frac{1}{2}\int_0^t u^{\frac{d-2}{2}}\omega(\sqrt{u})du.$$

As $d \geq 2$, the function $u \longmapsto u^{\frac{d-2}{2}}\omega(\sqrt{u})$ is continuous and V is $\mathcal{C}^1$. Moreover $V'(0) = 0$ if $d \geq 3$ and $V'(0) = \frac{1}{2}\omega(0) = \pi\theta(0)$ if $d = 2$.

In the case $d = 1$, the function

$$V(t) = \int_{-\sqrt{t}}^{\sqrt{t}} \theta(x)dx$$

is not derivable at 0 because θ is positive. But clearly

$$t \longmapsto V(t^2) = \int_{-t}^{t} \theta(x)dx$$

is $\mathcal{C}^1$.

□

Lemma 6. *Let θ be a $\mathcal{C}^1$ positive map on $D_{p,q}(r) \subset \mathbb{R}^p \times \mathbb{R}^q$, and*

$$V(t) = \int_{\{\|x\|^2 - \|y\|^2 \leq t\} \cap D_{p,q}(r)} \theta(x,y)dxdy.$$

If $d = p+q \geq 3$ then V is $\mathcal{C}^1$ on $]-r^2, r^2[$.
If $d = p+q = 2$ then V is not derivable at 0 but the function

$$t \longmapsto V(t^2)$$

is $\mathcal{C}^1$ on $]-r, r[$.

Proof.

Because of the symmetry of the domain $D_{p,q}(r)$ it is enough to compute $V(t)$ for $t \geq 0$ and to verify that its derivative at 0 is symmetric with respect to p and q, in order to ensure that V is $\mathcal{C}^1$, not only piecewise $\mathcal{C}^1$.

For this purpose we first compute the integral of θ on the domain

$$\{\|y\| \leq r, \quad \|y\|^2 \leq \|x\|^2 \leq \|y\|^2 + t\}$$

and we denote this integral by $V_1(t)$. We use the two changes of variable

$$\begin{array}{ccc} \mathbb{R}_+^* \times S^{p-1} & \longmapsto & \mathbb{R}^p \\ (s,z) & \longmapsto & sz \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathbb{R}_+^* \times S^{q-1} & \longmapsto & \mathbb{R}^q \\ (\tau,\zeta) & \longmapsto & \tau\zeta \end{array}.$$

$$\begin{aligned} V_1(t) &= \int_{\{\|y\| \leq r\}} \int_{\{\|y\| \leq \|x\| \leq \sqrt{\|y\|^2 + t}\}} \theta(x,y)dx \\ &= \int_{\{\|y\| \leq r\}} \int_{\|y\|}^{\sqrt{\|y\|^2+t}} s^{p-1}ds \int_{S^{p-1}} \theta(sz,y)\sigma_{p-1}(z) \\ &= \int_0^r \tau^{q-1}d\tau \int_{S^{q-1}} \sigma_{q-1}(\zeta) \int_\tau^{\sqrt{\tau^2+t}} s^{p-1}ds \int_{S^{p-1}} \theta(sz,\tau\zeta)\sigma_{p-1}(z) \\ &= \int_0^r \tau^{q-1}d\tau \int_\tau^{\sqrt{\tau^2+t}} s^{p-1}\omega(s,\tau)ds \end{aligned}$$

where

$$\omega(s,\tau) = \int_{S^{p-1}\times S^{q-1}} \theta(sz,\tau\zeta)\sigma_{p-1}(z) \wedge \sigma_{q-1}(\zeta)$$

is a $\mathcal{C}^1$ map. We have now to distinguish three cases.
First case: $p \geq 2$

We can set $u = s^2$ and obtain

$$V_1(t) = \frac{1}{2}\int_0^r \tau^{q-1}d\tau \int_{\tau^2}^{\tau^2+t} u^{\frac{p-2}{2}}\omega(\sqrt{u},\tau)du.$$

Clearly V_1 is $\mathcal{C}^1$ on $[0, r^2]$ and

$$V_1'(t) = \frac{1}{2}\int_0^r \tau^{q-1}(\tau^2+t)^{\frac{p-2}{2}}\omega(\sqrt{\tau^2+t},\tau)d\tau.$$

In particular at $t = 0$

$$V_1'(0) = \frac{1}{2}\int_0^r \tau^{d-3}\omega(\tau,\tau)d\tau.$$

Second case: $p = 1$, $q \geq 2$

We set

$$g(u,\tau) = \int_\tau^u \omega(s,\tau)ds \quad \text{and} \quad h(t,\tau) = \tau^{q-1}g(\sqrt{\tau^2+t},\tau).$$

For $0 < \tau \leq r$, and $0 \leq t \leq r^2$, we have

$$\begin{aligned} \tfrac{\partial}{\partial t}h(t,\tau) &= \tau^{q-1}\tfrac{\partial}{\partial u}g(\sqrt{\tau^2+t},\tau)\tfrac{1}{2\sqrt{\tau^2+t}} \\ &\leq \tfrac{1}{2}\tau^{q-2}\tfrac{\partial}{\partial u}g(\sqrt{\tau^2+t},\tau) \end{aligned}$$

and

$$\lim_{t\mapsto 0}\frac{\partial}{\partial t}h(t,\tau) = \frac{\partial}{\partial t}h(0,\tau) = \frac{1}{2}\tau^{q-2}\frac{\partial}{\partial u}g(\tau,\tau).$$

Therefore V_1 is $\mathcal{C}^1$ and

$$V_1'(0) = \int_0^r \frac{1}{2}\tau^{q-2}\frac{\partial}{\partial u}g(\tau,\tau)d\tau = \frac{1}{2}\int_0^r \tau^{d-3}\omega(\tau,\tau)d\tau.$$

Third case: $p = q = 1$

We have

$$V_1(t) = \int_0^r d\tau \int_\tau^{\sqrt{\tau^2+t}} \omega(s,\tau)ds.$$

In the particular case where $\omega = 1$, it is clear that V_1 is not derivable at $t = 0$ because the slope at 0 is infinite (the computation is left to the reader). In the general case there exists $m > 0$ such that $\omega \geq m$, and V_1 is no more derivable at $t = 0$.

However let us consider

$$V_1(t^2) = \int_0^r d\tau \int_\tau^{\sqrt{\tau^2+t^2}} \omega(s,\tau)ds.$$

We have

$$\frac{\partial}{\partial t}\int_\tau^{\sqrt{\tau^2+t^2}} \omega(s,\tau)ds = \frac{t}{\sqrt{\tau^2+t^2}}\omega(s,\tau) \leq \omega(s,\tau)$$

and

$$\frac{\partial}{\partial t}\int_\tau^{\sqrt{\tau^2+t^2}} \omega(s,\tau)ds \longmapsto_{t\mapsto} 0.$$

Therefore $t \mapsto V_1(t^2)$ is $\mathcal{C}^1$ and its derivative vanishes at $t = 0$.

It remains to compute the integral of θ on the compact

$$\{\|y\| \leq r, \quad \|y\|^2 \leq \|x\|^2 \leq \|y\|^2 + t\}\bigcap\{\|x\|\,\|y\| \leq r^2\}.$$

Let us denote this integral by $W(t)$. We have

$$W(t) = \int_{\{m(t)\leq\|y\|\leq r\}} dy \int_{\frac{r^2}{\|y\|}}^{\sqrt{\|y\|^2+t}} \theta(x,y)dx$$

where $m(t)$ is the minimum value of $\|y\|$ for which $\frac{r^2}{\|y\|} \leq \sqrt{\|y\|^2+t}$. Notice that

$$m(t) = \sqrt{\frac{\sqrt{t^2+4r^4}-t}{2}}$$

is $\mathcal{C}^1$ and has a positive lower bound as long as $t \leq r^2$. Therefore

$$\psi(t,y) = \int_{\frac{r^2}{\|y\|}}^{\sqrt{\|y\|^2+t}} \theta(x,y)dx$$

is $\mathcal{C}^1$, and, using a now classical change of variable,

$$W(t) = \int_{m(t)}^{r} \tau^{q-1}d\tau \int_{S^{q-1}} \psi(t,\tau\zeta)\sigma_{q-1}(\zeta)$$

is also $\mathcal{C}^1$. Let us show that $W'(0) = 0$. First at all we have $\psi(0, r\zeta) = 0$ because for $t = 0$, $m(t) = r$ and the bounds of the integral are r and r. Hence

$$\begin{aligned} W'(0) = & \int_{m(0)}^{r} \tau^{q-1}\frac{\partial}{\partial t}\left(\int_{S^{q-1}} \psi(t,\tau\zeta)\sigma_{q-1}(\zeta)\right)_{t=0} d\tau \\ & -m'(0)\left(\tau^{q-1}\int_{S^{q-1}} \psi(0,\tau\zeta)\sigma_{q-1}(\zeta)\right)_{\tau=m(0)} = 0. \end{aligned}$$

□

Acknowledgement

We want to thank S. Le Borgne for the helpfull discussions we had about torus automorphisms, Professor D. Volný for suggesting the problem and Professor J.M. Strelcyn for his precious advices.

References

[1] Marcel Berger and Bernard Gostiaux. *Géométrie différentielle: variétés, courbes et surfaces.* Mathématiques. [Mathematics]. Presses Universitaires de France, Paris, second edition, 1992.

[2] Patrick Billingsley. *Convergence of probability measures.* John Wiley & Sons Inc., New York, 1968.

[3] Svetlana Borovkova, Robert Burton, and Herold Dehling. Limit theorems for functionals of mixing processes with applications to U-statistics and dimension estimation. *Trans. Amer. Math. Soc.*, 353(11):4261–4318 (electronic), 2001.

[4] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms.* Springer-Verlag, Berlin, 1975. Lecture Notes in Mathematics, Vol. 470.

[5] P. Collet, S. Martinez, and B. Schmitt. Asymptotic distribution of tests for expanding maps of the interval. *Ergodic Theory Dynam. Systems*, 24(3):707–722, 2004.

[6] Jérôme Dedecker and Clémentine Prieur. An empirical central limit theorem for dependent sequences. *Stochastic Process. Appl.*, 117(1):121–142, 2007.

[7] Herold Dehling and Walter Philipp. Empirical process techniques for dependent data. In *Empirical process techniques for dependent data*, pages 3–113. Birkhäuser Boston, Boston, MA, 2002.

[8] Morris W. Hirsch. *Differential topology.* Springer-Verlag, New York, 1976. Graduate Texts in Mathematics, No. 33.

[9] Stéphane Le Borgne. *Dynamique symbolique et propriétés stochastiques des automorphismes du tore: cas hyperbolique et quasi hyperbolique.* PhD thesis, Université de Rennes 1, 1997.

[10] Stéphane Le Borgne. Exemples de systèmes dynamiques quasi-hyperboliques à décorrélations lentes. *http://perso.univ-rennes1.fr/stephane.leborgne/*, 2006.

[11] Stéphane Le Borgne and Françoise Pène. Vitesse dans le théorème limite central pour certains systèmes dynamiques quasi-hyperboliques. *Bull. Soc. Math. France*, 133(3):395–417, 2005.

[12] D. A. Lind. Dynamical properties of quasihyperbolic toral automorphisms. *Ergodic Theory Dynamical Systems*, 2(1):49–68, 1982.

[13] J. Milnor. *Morse theory.* Based on lecture notes by M. Spivak and R. Wells. Annals of Mathematics Studies, No. 51. Princeton University Press, Princeton, N.J., 1963.