Tightness in Schauder decomposable Banach spaces*

Ch. Suquet

Laboratoire de Statistique et Probabilités, Bât. M2

Université des Sciences et Technologies de Lille

F-59655 Villeneuve d'Ascq Cedex France

Abstract

We characterize the tightness of a set of probability measures in a large class of Banach spaces including those having a Schauder basis. We give various applications to sequences of stochastic process viewed as random elements in the spaces $\mathrm{L}^p(0,1)$, $\mathrm{L}^p(\mathbb{R})$ or in some Hölder or Besov spaces.

AMS classifications: 60B10, 60F17, 60G30.
Key words: empirical process, Haar basis, multiresolution analysis, Schauder decomposition, strong mixing, tightness, wavelets.

*Preprint

1 Introduction

Relative compactness in the space of probability measures is a key tool in the study of weak convergence. A family $\mathcal{F}$ of probability measures on the general metric space S is said to be tight if for each positive ε, there is a compact set K such that $P(K) > 1 - \varepsilon$ for all P in $\mathcal{F}$. According to Prohorov's theorem, tightness is always a sufficient condition for relative compactness and is also necessary if S is separable and complete.

The Skorohod space $S = D(0,1)$ is the usual framework of many limit theorems for stochastic processes. This is so because it supports processes that contain jumps and weak convergence in $D(0,1)$ provides results about some useful functionnals of paths like those involving the suprema. Nevertheless this space presents some drawbacks. First, tightness in $D(0,1)$ is sometimes difficult to check. Second, under the pointwise addition of functions, the space $D(0,1)$ is not a topological group and hence not a topological vector space.

In many cases, it seems very convenient to treat a stochastic process as a random element in a functional Banach space. The best known case is certainly the $C(0,1)$ one (see Billingsley [4]). As for the Hilbert space case, sufficient conditions for tigthness are given by Prohorov [18], Parthasarathy [17] and Gihman Skorohod [8]. The recents developments in the theory of wavelets and their applications in probability and statistics show the interest of using more sophisticated functions spaces like the Hölder, Sobolev or Besov spaces.

In this paper we will present an unified approach to tightness problems in a large class of Banach spaces including $C(0,1)$, the L^p, Hölder, Sobolev and Besov spaces. Our starting point is the Hilbertian case:

Theorem 1 (Suquet [21]) *Let H be a separable Hilbert space and $(e_i,\ i \in \mathbb{N})$ an orthonormal basis of H. Define for each h in H and each positive integer N:*

$$s_N^2(h) = \sum_{i<N} \langle h, e_i \rangle^2, \qquad r_N^2(h) = \sum_{i \geq N} \langle h, e_i \rangle^2.$$

The family $\mathcal{F}$ of probability measures on H is tight if and only if:

(*i*) $\forall N \geq 1 \quad \lim_{t \to +\infty} \sup_{P \in \mathcal{F}} P(\{h \in H\ :\ s_N^2(h) > t\}) = 0,$

(*ii*) $\forall t > 0 \quad \lim_{N \to +\infty} \sup_{P \in \mathcal{F}} P(\{h \in H\ :\ r_N^2(h) > t\}) = 0.$

Our aim is to generalize this theorem. Let us denote by V_j the finite dimensional subspace of H generated by $\{e_0, \ldots, e_{j-1}\}$ and by E_j the orthogonal projection on V_j. With these notations, the conditions (i) and (ii) can be recast as:

(*i*) $E_j\mathcal{F} = \{P \circ E_j^{-1}, P \in \mathcal{F}\}$ is tight for each $j \geq 1$,

(*ii*) $\forall \varepsilon > 0,\ \lim_{j \to +\infty} \sup_{P \in \mathcal{F}} P(\{h \in H\ :\ \|h - E_j h\| \geq \varepsilon\}) = 0.$

We would like to generalize this result in two directions: dropping the finite dimensionality of the V_j (as in the case of multiresolution analysis) and dropping the orthogonality of the E_j (and hence the hilbertian character of the space). Of course, we have to keep some control on the norms of the projectors E_j. This leads us naturally to deal with Schauder decomposable Banach spaces. The next section will present the functional needed analysis package . The third section exposes the proof of our main result: the extension of the theorem 1 for these spaces.

In the following section, we use our main theorem and a multiresolution analysis to obtain a sufficient condition for tightness in $\mathrm{L}^p(\mathbb{R})$, $1 < p < +\infty$. We study also the relative compactness in $\mathrm{L}^p(0,1)$ of the random step functions involved in Donsker's theorem (without any assumption on the dependence structure of the underlying random variables). We rederive and clarify a previous condition for tightness in $L^2(0,1)$ due to Jacob, Oliveira and Suquet ([14], [15]). Next we are interested in the convergence of the empirical process based on strong mixing uniform variables $(X_i)_{i\geq 1}$ on $[0,1]$. In the $D(0,1)$ setting, the best result up to now is due to Yoshihara [23] who proved the weak convergence of the empirical process to a gaussian process under the condition $\alpha_n = O(n^{-3-\varepsilon})$ (the α_n being the strong mixing coefficients of the sequence $(X_i)_{i\geq 1}$). Recently, Oliveira and Suquet proved the same convergence in $L^2(0,1)$ under the weaker assumption $\sum \alpha_n < +\infty$. Of course the convergence in $L^2(0,1)$ is weaker than in $D(0,1)$. Here we obtain the convergence in $\mathrm{L}^p(0,1)$ $(2 \leq p < 6)$ under a weaker condition (depending on p) than Yoshihara's one.

In the fifth section we are concerned with the spaces H^0_α of Hölderian functions on $[0,1]$ (*i.e.* $f(0) = 0$, $|f(t) - f(s)| \leq C|t-s|^\alpha$ and $|f(t) - f(s)| = o(|t-s|^\alpha)$). We obtain a tightness criterion and a sufficient condition very similar to the $C(0,1)$ case. Some examples are discussed.

The last section presents an easy application to some sequences spaces and their isomorphic Besov functional spaces. This application could be useful in the study of weak convergence for stochastic processes known by their wavelets coefficients.

2 The functional analysis background

We refer to Singer [20] for the Schauder decompositions and to Meyer [13] for wavelets and multiresolution analysis. Let $(\mathcal{X},\ \|\ \|)$ be a Banach space. A system $\{x_n,\ n \in \mathbb{N}\}$ of elements of $\mathcal{X}$ is called a Schauder basis if for every element $x \in \mathcal{X}$, there is a unique series:

$$x = \sum_{n=0}^{+\infty} a_n x_n, \qquad a_n = a_n(x) \in \mathbb{R}, \tag{1}$$

which converges to x in the norm of $\mathcal{X}$. We define the associated coordinate projections v_n by $v_n(x) = a_n x_n$. These projections are continuous and there is a constant C depending only on the basis $\{x_n,\ n \in \mathbb{N}\}$ such that:

$$\forall N \in \mathbb{N},\ \forall x \in \mathcal{X}, \quad \Big\|\sum_{n=0}^{N} v_n(x)\Big\| \le C\|x\|. \tag{2}$$

The basis is said unconditional if the series (1) is unconditionally convergent, that is, for every permutation $\sigma = \{\sigma(n),\ n \in \mathbb{N}\}$ of the indexes, the series $\sum_{n=0}^{+\infty} a_{\sigma(n)} x_{\sigma(n)}$ converges to x in the norm of $\mathcal{X}$. In the multiresolution analysis setting defined below, we are dealing with bases of wavelets indexed by $\mathbb{N} \times \mathbb{Z}$, $\mathbb{Z} \times \mathbb{Z}$, $\mathbb{Z} \times \mathbb{Z}^d$. Fortunately for many functional spaces, the decomposition on these bases is an unconditionally convergent series so that neither the order of summation nor the groupings of terms do matter. Here the groupings of terms are usually made according to the level of resolution. At one given level, say 2^{-j}, we have a countable family of functions $(\psi_{j,k},\ k \in \mathbb{Z})$ of the basis which is dense in a closed subspace of the involved functional space. In the important case of the Faber-Schauder basis of the space $C(0,1)$, which do not define a multiresolution analysis, we have the groupings $\{\Delta_{j,k},\ 0 \le k < 2^j\}$ which span closed finite dimensional subspaces of $C(0,1)$. Banach spaces having a Schauder decomposition are the natural framework unifying all these situations.

Definition 1 *An infinite sequence $(G_j,\ j \in \mathbb{N})$ of closed linear subspaces of a Banach space $\mathcal{X}$ such that $G_j \neq \{0\}$ $(j \in \mathbb{N})$ is called a Schauder decomposition of $\mathcal{X}$ if for every $x \in \mathcal{X}$ there exists an unique sequence $(y_n,\ n \in \mathbb{N})$ with $y_j \in G_j$ $(j \in \mathbb{N})$ such that:*

$$x = \sum_{j=0}^{+\infty} y_j$$

and if the coordinate projections defined by $v_n(x) = y_n$, are continuous on $\mathcal{X}$.

In other words, $(G_j,\ j \in \mathbb{N})$ is a decomposition of $\mathcal{X}$ if and only if, $\mathcal{X}$ is the direct topological sum of the subspaces G_j. It should be noticed here that some Banach space do not possess a Schauder decomposition, for instance the space ℓ^∞ (see Singer [20]) and that a Schauder decomposable Banach space need not be separable.

Let us denote $V_j = \bigoplus_{i \le j} G_i$ and $E_j = \sum_{i \le j} v_i$ the continuous projections of $\mathcal{X}$ onto V_j. It follows from proposition 15.3 p. 488 in Singer [20] that:

$$C = \sup_{j \in \mathbb{N}} \|E_j\| < +\infty. \tag{3}$$

By the orthogonality relations between the coordinate projections (i.e.: $v_i \circ v_j = \delta_{ij} v_i = \delta_{ij} v_j$) we have:

$$E_{j'} \circ E_j = E_j \circ E_{j'} = E_{j \wedge j'}, \qquad j, j' \in \mathbb{N}. \tag{4}$$

This implies:

$$\|x - E_i x\| \le (1 + C)\|x - E_j x\|, \qquad x \in \mathcal{X},\ i > j. \tag{5}$$

In separable Banach spaces having a Schauder decomposition, we have a very simple criterion for relative compactness. This criterion is a generalization of the Hilbert space case (see for instance Akhiezer and Glazman [1]). We did not find this result in its general form in the literature, so we give a detailed proof.

Theorem 2 *Let $\mathcal{X}$ be a separable Banach space having a Schauder decomposition. A subset K is relatively compact in $\mathcal{X}$ if and only if:*

(*i*) *For each $j \in \mathbb{N}$, $E_j K$ is relatively compact in V_j,*

(*ii*) $\sup_{x \in K} \|x - E_j x\| \to 0$ *as* $j \to +\infty$.

Proof :
Sufficiency of (i) *and* (ii): Let $(z_n,\ n \in \mathbb{N})$ be a sequence in K. We have to check that $(z_n,\ n \in \mathbb{N})$ contains a convergent subsequence $(z_n,\ n \in I)$ where I is some infinite subset of $\mathbb{N}$. Using repeatedly (i), we can construct a sequence (J_j) of infinite subsets of $\mathbb{N}$ such that: $\mathbb{N} \supset J_0 \supset J_1 \supset \cdots$ and for each j, $(E_j z_n,\ n \in J_j)$ converges in V_j. Moreover we can require that $\min(J_j) < \min(J_{j+1})$, $j \in \mathbb{N}$. Let us define then $I = \{\min(J_j),\ j \in \mathbb{N}\}$. By construction we have $\{n \in I,\ n \ge \min(J_j)\} \subset J_j$ for each $j \in \mathbb{N}$ and hence the sequence $(E_j z_n,\ n \in I)$ converges in V_j towards some y_j. Now we show that $(z_n,\ n \in I)$ is a Cauchy sequence in $\mathcal{X}$. For fixed positive ε there is by (ii) an integer j such that $\|x - E_j x\| < \varepsilon$ for all x in K. By the construction of I, there is an integer n_0 such that for all $n > n_0$ in I, $\|E_j z_n - y_j\| < \varepsilon$. So for every n and p larger than n_0 in I, $\|z_n - z_p\| < 4\varepsilon$. ■

Necessity of (i) *and* (ii): The necessity of (i) follows obviously from the continuity of the projections E_j. To prove (ii), we can assume without lose generality that K is closed and hence compact. Thus the continuous function $x \mapsto \|x - E_j x\|$ takes its maximum over K for some $z_j \in K$. Put:

$$y_j = z_j - E_j z_j, \qquad \|y_j\| = \sup_{x \in K} \|x - E_j x\|.$$

It suffices then to prove the convergence to zero of $(y_j,\ j \in \mathbb{N})$. First we observe that $(y_j,\ j \in \mathbb{N})$ is relatively compact in $\mathcal{X}$. Indeed, taking subsequences it suffices to check that if $(z_j,\ j \in J)$ converges to z, $(E_j z_j,\ j \in J)$ is a convergent sequence. Writing:

$$\|E_j z_j - z\| \le \|E_j\| \cdot \|z_j - z\| + \|E_j z - z\|,$$

this follows from (3) and the definition 1. Moreover we have $\lim E_j z_j = \lim z_j$, so that the only possible limit for a subsequence of $(y_j\ j \in \mathbb{N})$ is zero. Hence y_j converges to 0, which ends the proof. ■

Let us now have a more detailed look at the examples of Schauder decompositions referred above. The first instance is provided by the Schauder bases $(x_n,\, n \in \mathbb{N})$, taking $G_j = \operatorname{span}[x_j]$. For some bases it is more convenient to have the G_j as span of a finite number of vectors of the basis. This is the case of the Haar and Faber-Schauder bases we are now recalling the definition.

The Haar basis $(e_n,\, n \in \mathbb{N})$ is an unconditional basis for the spaces $\mathrm{L}^p(0,1)$, $(1 < p < +\infty)$. Put $\psi(t) = \mathbb{1}_{[0,1/2[}(t) - \mathbb{1}_{[1/2,1[}(t)$. The Haar basis is defined by $e_0(t) = 1$ and $e_n(t) = e_{j,k}(t) = 2^{j/2}\psi\big(2^j t - k\big)$ where $n = 2^j + k$ with $0 \leq k < 2^j$. Here the Schauder decomposition of $\mathrm{L}^p(0,1)$ we are interested in is given by $G_0 = \operatorname{span}[e_0]$, $G_j = \operatorname{span}[e_{j,k},\, 0 \leq k < 2^j]$. Let us recall that the projection of f onto $V_j = \bigoplus_{0 \leq i \leq j} G_i$ is its approximation by a step function equal to the mean value of f over each intervall $\big[k2^{-j},\, (k+1)2^{-j}\big[$, $0 \leq k < 2^j$.

The Faber-Schauder basis $(\Delta_n,\, n \in \mathbb{N} \cup \{-1\})$ is a Schauder basis for the space $C(0,1)$ of continuous functions on $[0,1]$. This space have no unconditional basis. Put $\Delta(t) = 2t\mathbb{1}_{[0,1/2[}(t) + 2(1-t)\mathbb{1}_{[1/2,1[}(t)$. Then $\Delta_{-1}(t) = 1$, $\Delta_0(t) = t$, $\Delta_1(t) = \Delta_{0,0}(t) = \Delta(t)$, $\Delta_n(t) = \Delta_{j,k}(t) = \Delta(2^j t - k)$ where $n = 2^j + k$ with $0 \leq k < 2^j$. Here we define the G_j as in the Haar basis case (and $G_{-1} = \operatorname{span}[\Delta_{-1}]$). The projection of a continuous function f onto V_j is simply its approximation by linear interpolation with knots $\big(k2^{-j},\, f(k2^{-j})\big)$. The Faber Schauder basis is also a Schauder basis in the Hölderian spaces H^0_α (see section 5 below).

Finally, we recall some useful facts about wavelets and multiresolution analysis (in a reduced version adapted to our purpose, the general definitions can be found in Meyer [13] or Daubechies [7]). In what follows, for $g \in L^2(\mathbb{R})$, we write $g_{j,k}$ for the function $g_{j,k}(t) = 2^{j/2}g\big(2^j t - k\big)$, $j,\, k \in \mathbb{Z}$. By multiresolution analysis with scaling function φ, we mean a ladder of closed subspaces $(V_j$, $j \in \mathbb{Z})$ of $L^2(\mathbb{R})$ such that:

a) $\displaystyle\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$, $\displaystyle\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$,

b) $V_j \subset V_{j+1}$,

c) $V_j = \overline{\operatorname{span}}\,[\varphi_{j,k},\, k \in \mathbb{Z}]$,

d) $(\varphi_{0,k},\, k \in \mathbb{Z})$ is an orthonormal basis of V_0.

The multiresolution analysis is called r-regular ($r \in \mathbb{N}$) if φ is of C^r class and for each integer m there is a constant a_m such that:

$$\big|\varphi^{(\alpha)}\big|(t) \leq a_m\big(1 + |t|\big)^{-m}, \quad t \in \mathbb{R},\ \alpha \leq r. \tag{6}$$

Define W_j as the orthogonal complement of V_j in V_{j+1}, then we have for each $j \in \mathbb{Z}$ the decomposition:

$$L^2(\mathbb{R}) = V_j \oplus \bigoplus_{i \geq j} W_i. \tag{7}$$

One can construct a function ψ such that $(\psi_{0,k},\ k \in \mathbb{Z})$ is an orthonormal basis of W_0 and $(\psi_{j,k},\ j,\ k \in \mathbb{Z})$ is an orthonormal wavelets basis of $L^2(\mathbb{R})$. If the multiresolution analysis is r-regular, ψ verify also the property (6). We write E_j (resp. D_j) for the orthogonal projection from $L^2(\mathbb{R})$ onto V_j (resp. W_j) and its associated integral kernel:

$$E_j f = \int_{\mathbb{R}} E_j(\,.\,,s) f(s)\,ds, \qquad E_j(t,s) = \sum_{k\in\mathbb{Z}} \varphi_{j,k}(t)\bar{\varphi}_{j,k}(s), \tag{8}$$

$$D_j f = \int_{\mathbb{R}} D_j(\,.\,,s) f(s)\,ds, \qquad D_j(t,s) = \sum_{k\in\mathbb{Z}} \psi_{j,k}(t)\bar{\psi}_{j,k}(s). \tag{9}$$

The kernels E_j and D_j verify:

$$E_j(s,t) = 2^j E_0(2^j s, 2^j t), \quad D_j(s,t) = 2^j D_0(2^j s, 2^j t), \quad s,\ t \in \mathbb{R}. \tag{10}$$

Using (6), it is easily verified that E_0 and D_0 are majorized by convolution kernels. More precisely, there exists two rapidly decreasing functions K and L such that:

$$\big|E_0(s,t)\big| \le K(s-t), \quad \big|D_0(s,t)\big| \le L(s-t). \tag{11}$$

As shown in Meyer [13], the usefulness of the wavelets bases associated to a regular multiresolution analysis goes far beyond the $L^2(\mathbb{R})$ space. They provide unconditional bases for many functions spaces as $\mathrm{L}^p(\mathbb{R})$ $(1 < p < +\infty)$, Sobolev, Hölder and Besov spaces. In each case, the function space $\mathcal{X}$ is the topological direct sum of V_0 and the W_i $(i \ge 0)$ (these subspaces being redefined in an adapted way). We have then a Schauder decomposition of $\mathcal{X}$ given by $G_0 = V_0$, $G_j = W_{j-1}$ $(j \ge 1)$.

3 Main result

We give now the characterization of the tightness for separable Schauder decomposable Banach spaces.

Theorem 3 *Let $\mathcal{X}$ be a separable Banach space having a Schauder decomposition:*

$$\mathcal{X} = \bigoplus_{i=0}^{+\infty} G_i, \quad V_j = \bigoplus_{i=0}^{j} G_i, \qquad j = 0,1,2,\ldots$$

and denote by E_j the continuous projection from $\mathcal{X}$ onto V_j. Let $\mathcal{F}$ be a family of probability measures on $\mathcal{X}$ and $E_j\mathcal{F} = \{\mu \circ E_j^{-1},\ \mu \in \mathcal{F}\}$. Then $\mathcal{F}$ is tight if and only if:

(i) $E_j\mathcal{F}$ is tight, $j = 0,1,2,\ldots$

(ii) For each positive ε, $\lim_{j\to+\infty} \sup_{\mu\in\mathcal{F}} \mu\big(x \in \mathcal{X} :\ \|x - E_j x\| > \varepsilon\big) = 0.$

Proof :
Sufficiency of (i) *and* (ii): For fixed positive η, put $\eta_l = 2^{-l}$, $l = 1, 2, \ldots$ and choose a sequence (ε_l) decreasing to 0. By (ii), there is an integer j_l such that:

$$\forall \mu \in \mathcal{F}, \quad \mu\big(x \in \mathcal{X} : \|x - E_{j_l}x\| > \varepsilon_l\big) < \eta_l. \tag{12}$$

By (i), there is a compact subset K_l of $\mathcal{X}$ such that:

$$\forall \mu \in \mathcal{F}, \quad \mu\big(x \in \mathcal{X} : E_{j_l}x \in K_l\big) > 1 - \eta_l. \tag{13}$$

From (12) and (13) we deduce:

$$\forall \mu \in \mathcal{F}, \quad \mu\Big(\bigcap_{l \geq 1}\{x \in \mathcal{X} : E_{j_l}x \in K_l \text{ and } \|x - E_{j_l}x\| \leq \varepsilon_l\}\Big) > 1 - 2\eta. \tag{14}$$

It remains to check the compacity in $\mathcal{X}$ of the intersection in (14). This follows easily from the continuity of the E_j, (4), (5) and the theorem 2. ■

Necessity of (i) *and* (ii): As tightness is preserved by continuous mappings, the necessity of (i) follows from the continuity of the E_j. To prove the necessity of (ii), we need the following lemma.

Lemma 1 *Let $\mathcal{F}$ be a compact family (for the topology of weak convergence) of probability measures on the separable metric space S. Let $(F_l,\ l \in \mathbb{N})$ be a sequence of closed subsets of S decreasing to $\emptyset$. Define the functions u_l $(l \in \mathbb{N})$ by: $u_l :\quad P \longmapsto u_l(P) = P(F_l)$. Then the sequence (u_l) uniformly converges to zero on $\mathcal{F}$.*

Proof : For positive ε, let us define $D_{l,\varepsilon} = \{P \in \mathcal{F} : \ u_l(P) \geq \varepsilon\}$. We first verify that $D_{l,\varepsilon}$ is closed. The topology of weak convergence on probability measures over S being metrizable, this can be done by means of sequences. Let $(P_n,\ n \in \mathbb{N})$ be a sequence in $D_{l,\varepsilon}$, weakly convergent to some P. By the portmanteau theorem we have:

$$u_l(P) = P(F_l) \geq \limsup_{n \to +\infty} P_n(F_l) = \limsup_{n \to +\infty} u_l(P_n) \geq \varepsilon,$$

so P is in $D_{l,\varepsilon}$, which is then closed. The monotone continuity of probability measures implies clearly $\bigcap_{l \in \mathbb{N}} D_{l,\varepsilon} = \emptyset$. In view of the compactness of $\mathcal{F}$, we can find an l_0 such that $\bigcap_{l \leq l_0} D_{l,\varepsilon} = \emptyset$. As the sequence $(D_{l,\varepsilon})$ decreases, we have: $u_l(P) < \varepsilon$ for all $l \geq l_0$, and all $P \in \mathcal{F}$. ■

Now we apply this lemma with F_j taken as the closure of:

$$A_j = \{x \in \mathcal{X},\ \sup_{i \geq j} \|x - E_i x\| \geq \varepsilon\}.$$

Clearly $(A_j,\ j \in \mathbb{N})$ is decreasing and so is $(F_j,\ j \in \mathbb{N})$. Using (5), we have:

$$\forall x \in A_j, \quad \|x - E_j x\| \geq \frac{\varepsilon}{1 + C}.$$

By continuity of $Id - E_j$, this remains true for all x in F_j. As for each x in $\mathcal{X}$, $\lim_{j\to+\infty} E_j x = x$, this imply $\bigcap_{j\in\mathbb{N}} F_j = \emptyset$. Since $\mu(x \in \mathcal{X},\ \|x - E_j x\| \geq \varepsilon) \leq \mu(F_j)$, the lemma 1 give us the expected conclusion. ∎

When the subspaces of the Schauder decomposition are finite dimensional, the theorem 3 has the following more tractable version:

Theorem 4 *Assume the Banach space $\mathcal{X}$ has the Schauder decomposition $\mathcal{X} = \bigoplus_{j\in\mathbb{N}} G_j$ where each G_j is of finite dimension. Then $\mathcal{F}$ is tight if and only if the condition (ii) of theorem 3 and the following condition (i') hold:*

(i') $\lim_{A\to+\infty} \sup_{\mu\in\mathcal{F}} \mu\big(x \in \mathcal{X} :\ \|x\| > A\big) = 0.$

Proof : Clearly (i') holds if $\mathcal{F}$ is tight. On the other hand, $\|x\|$ and $\sup_j \|E_j x\|$ being equivalent norms in $\mathcal{X}$ (Singer [20], prop. 15.3 b) p. 488), (i') implies:

$$\lim_{A\to+\infty} \sup_{\mu\in\mathcal{F}} \mu\big(x \in \mathcal{X} :\ \|E_j x\| > A\big) = 0, \quad j = 0, 1, 2, \ldots$$

As the V_j are finite dimensional, the tightness of $E_j\mathcal{F}$ follows. ∎

The theorem 4 applies in particular to the decompositions coming from a Schauder basis. As a result, combining theorems 3 and 4, we can treat the case of multiresolution analysis. Indeed, by an elementary topological argument, tightness in V_j reduces to tightness in V_0 and in the W_i $(0 \leq i < j)$. And these spaces have a wavelets Schauder basis.

By Markov's inequality, the sufficient conditions for tightness in theorem 4 admit the following moment form.

Theorem 5 *Assume the Banach space $\mathcal{X}$ has the Schauder decomposition $\mathcal{X} = \bigoplus_{j\in\mathbb{N}} G_j$ where each G_j is of finite dimension. Then $\mathcal{F}$ is tight if:*

(i) $\exists\alpha > 0,\ \sup_{\mu\in\mathcal{F}} \mathbb{E}_\mu \|\xi\|^\alpha < +\infty,$

(ii) $\exists\beta > 0,\ \lim_{j\to+\infty} \sup_{\mu\in\mathcal{F}} \mathbb{E}_\mu \|\xi - E_j\xi\|^\beta = 0,$

where ξ denotes a random element in $\mathcal{X}$ with distribution μ and $\mathbb{E}_\mu$ the expectation with respect to μ.

This last theorem is a generalization of the theorem 1.13 of Prohorov [18] for the Hilbertian case. According to this theorem (see also Parthasarathy [17] th. 2.2 p. 154), when $\mathcal{X}$ is a separable Hilbert space, $\mathcal{F}$ is tight if:

$$\lim_{j\to+\infty} \sup_{\mu\in\mathcal{F}} \int_{\mathcal{X}} r_j^2(x)\,\mu(dx) = 0, \tag{15}$$

where $r_j^2(x) = \sum_{i\geq j} \langle x, e_i\rangle^2$ and $(e_i,\ i \in \mathbb{N})$ is an orthonormal basis of $\mathcal{X}$. That is exactly the condition (ii) above with $\beta = 2$ for the Schauder decomposition associated to the basis $(e_i,\ i \in \mathbb{N})$.

In [21] the Prohorov-Parthasarathy's statement was shown to be incomplete according to the following counter example: take $\mathcal{F}$ as the canonical image in $V_d = \mathrm{span}[e_0, e_1, \ldots, e_d]$ of a non-tight family in $\mathbb{R}^d$. Clearly (15) is satisfied but $\mathcal{F}$ is not tight. The author proposed a rectified statement by adding to (15) the condition (i) of theorem 5 with $\alpha = 2$. So the theorem 5 is effectively a generalization of the rectified Prohorov's theorem.

Let us consider a sequence of stochastic processes as a sequence of random elements in $\mathcal{X}$. The conditions (i) or (i') of the theorems above are related to the size of the paths (in the norm of $\mathcal{X}$). As it will be illustrated in the following sections, in many cases, the conditions (ii) involve the oscillations of the processes (in a sense depending upon the norm of $\mathcal{X}$). Our next result roughly says that, if the sample path's size is well controlled, there is no real loss in investigating the tightness of the sequence by mean of the moment condition (ii) of theorem 5.

Theorem 6 *Suppose condition (i) of theorem 5 holds. Then condition (ii) with $\beta < \alpha$ is necessary for the tightness of $\mathcal{F}$.*

Without loss of generality, we can suppose $\mathcal{F}$ compact. Then condition (ii) will result from the following stronger one:

$$\lim_{j\to+\infty} \sup_{\mu\in\mathcal{F}} \mathbb{E}_\mu \Big(\sup_{i\geq j} \|\xi - E_i\xi\|^\beta \Big) = 0. \tag{16}$$

The first step is to prove the continuity of the functionals:

$$T_j : \mu \longmapsto \mathbb{E}_\mu \Big(\sup_{i\geq j} \|\xi - E_i\xi\|^\beta \Big), \quad j \in \mathbb{N}.$$

To this end, consider a sequence (μ_n) weakly converging to the probability measure μ and the random elements ξ_n, ξ with respective distributions μ_n, μ. We introduce the non negative random variables:

$$X_n = \sup_{i\geq j} \|\xi_n - E_i\xi_n\|^\beta, \quad X = \sup_{i\geq j} \|\xi - E_i\xi\|^\beta.$$

Putting $\alpha/\beta = 1 + \varepsilon$, we have by (3) a constant A such that:

$$\mathbb{E}_\mu \big(X_n^{1+\varepsilon} \big) \leq A\, \mathbb{E}_\mu \|\xi_n\|^\alpha, \quad n \in \mathbb{N}.$$

So (i) implies the uniform integrability of the sequence $(X_n,\, n \in \mathbb{N})$. Now we observe that the function $f_j : x \mapsto \sup_{i\geq j} \|x - E_i x\|^\beta$ is continuous on $\mathcal{X}$. This easily follows from (3) and (5). Thus, as ξ_n converges to ξ in distribution, the same is true for $X_n = f_j(\xi_n)$ and $X = f_j(\xi)$. So $\mathbb{E}\, X_n$ converges to $\mathbb{E}\, X$, which proves the continuity of T_j.

Now $(T_j,\, j \in \mathbb{N})$ is a decreasing sequence of continuous functions on the compact $\mathcal{F}$. According to Dini's theorem, $(T_j,\, j \in \mathbb{N})$ converges uniformly on

$\mathcal{F}$. To identify the limit, observe that:

$$\lim_{j\to+\infty} T_j(\mu) = \lim_{j\to+\infty} \int_{\mathcal{X}} \sup_{i\geq j} \|x - E_i x\|^\beta \, d\mu(x) = 0,$$

by (3), (i) and the dominated convergence theorem. ■

Before closing this section, we test our tightness criterion in the well known case of the space $C(0,1)$. Classically, tightness in this space is characterized via the Arzela-Ascoli theorem by ([4], th. 8.2):

(a) *For each positive η, there is an a such that*

$$\mu\big(x \in C(0,1) : \, |x(0)| > a\big) \leq \eta, \quad \forall \mu \in \mathcal{F}$$

(b) *For each positive ε, $\lim_{\delta\to 0} \sup_{\mu\in\mathcal{F}} \mu\big(x \in C(0,1) : \; w(\delta, x) \geq \varepsilon\big) = 0$, where $w(\delta, x)$ is the modulus of continuity.*

This criterion can be rederived as a corollary of theorem 4:

Proposition 1 *For the Schauder decomposition associated to the Faber-Schauder basis of $C(0,1)$, the conditions (i') and (ii) of theorem 4 are equivalent to the conditions (a) and (b).*

Proof : Recalling that for the Faber-Schauder basis, $E_j x$ is the linear interpolation of x with knots $\big(k2^{-j}, x(k2^{-j})\big)$, the following inequalities are easily checked:

$$\|x - E_j x\|_\infty \leq w(2^{-j}, x), \quad x \in C(0,1), \tag{17}$$

$$w(\delta, x) \leq 2\|x - E_j x\|_\infty + 2\|x\|_\infty 2^j \delta, \quad 0 < \delta < 2^{-j}, \; x \in C(0,1). \tag{18}$$

Now, using (18), it follows that (i') and (ii) imply (b). Moreover (i') implies obviously (a). On the other hand, by (17), (b) implies (ii). Finally, since $|x(t)| \leq |x(0)| + \frac{1}{\delta} w(\delta, x)$, ($i'$) follows from ($a$) and ($b$). ■

4 Some applications in L^p spaces

From now, we specialize in the case of stochastic processes considered as random elements in functions spaces. By definition a sequence of such processes is tight when the sequence of corresponding distributions is tight.

4.1 A sufficient condition in $\mathrm{L}^p(\mathbb{R})$

We shall use a r-regular multiresolution analysis of $\mathrm{L}^p(\mathbb{R})$ to obtain a sufficient condition for tightness of stochastic processes sequences in $\mathrm{L}^p(\mathbb{R})$. Let us first recall how the multiresolution analysis works in $\mathrm{L}^p(\mathbb{R})$. From a r-regular multiresolution analysis $(V_j,\, W_j;\ j \in \mathbb{Z})$ of $L^2(\mathbb{R})$, we define the spaces $V_j(p)$, $W_j(p)$ in the following way (Meyer [13] p. 31 and 45):

> *Case* $1 < p < 2$: $V_0(p) = V_0 \cap \mathrm{L}^p(\mathbb{R})$, $W_j(p) = W_j \cap \mathrm{L}^p(\mathbb{R})$,
>
> *Case* $2 < p < +\infty$: $V_0(p)$ and $W_j(p)$ are the completions of V_0 and W_j in the L^p norm.

In both cases, $V_j(p)$ is defined by change of scale: $f(t) \in V_0(p)$ if and only if $f(2^j t) \in V_j(p)$. Moreover we have $V_{j+1}(p) = V_j(p) \oplus W_j(p)$ and

$$\mathrm{L}^p(\mathbb{R}) = V_0(p) \oplus \bigoplus_{j=0}^{+\infty} W_j(p).$$

All these sums are direct and topological (i.e. the projections on the components are continuous in the L^p norm), so we have a Schauder decomposition defined by: $G_0 = V_0(p)$, $G_j = W_{j-1}(p)$.

Theorem 7 *A sequence* $(\xi_n,\, n \in \mathbb{N})$ *of stochastic processes with paths in* $\mathrm{L}^p(\mathbb{R})$ *(*$1 < p < +\infty$*) is tight in* $\mathrm{L}^p(\mathbb{R})$ *if:*

(i) $\displaystyle \sup_{n\in\mathbb{N}} \int_{\mathbb{R}} \mathbb{E}\,|\xi_n(t)|^p\, dt < +\infty,$

(ii) $\displaystyle \lim_{A\to+\infty} \sup_{n\in\mathbb{N}} \int_{\{|t|\geq A\}} \mathbb{E}\,|\xi_n(t)|^p\, dt = 0,$

(iii) There is a $\gamma > 0$ *and a function* $g \in \mathrm{L}^1(\mathbb{R})$ *such that:*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\,|\xi_n(t+u) - \xi_n(t)|^p \leq |u|^\gamma g(t), \qquad u, t \in \mathbb{R}.$$

Proof : The condition (ii) of theorem 3 will follow from the sufficient moment condition:

$$\lim_{j\to+\infty} \sup_{n\in\mathbb{N}} \mathbb{E}\,\|\xi_n - E_j\xi_n\|_p^p = 0 \tag{19}$$

An easy adaptation of the proof of theorem 9.1.6. in Daubechies [7] shows that if φ is C^1 and $|\varphi(t)|$, $|\varphi'(t)| \leq a(1+|t|)^{-1-\varepsilon}$, then $\{\varphi_{j,k},\ k \in \mathbb{Z}\}$ is an unconditional basis of $V_j(p)$. So we can use the theorem 5 to investigate the tightness of $(E_j\xi_n, n \in \mathbb{N})$ in $V_j(p)$. From now, let us choose the scaling function φ real, compactly supported and C^1.

Applying theorem 5 with $\alpha = \beta = p$, we have to check:

$$\sup_{n \in \mathbb{N}} \mathbb{E} \, \|E_j \xi_n\|_p^p < +\infty \tag{20}$$

and

$$\lim_{l \to +\infty} \sup_{n \in \mathbb{N}} \mathbb{E} \, \|E_j \xi_n - E_{j,l} \xi_n\|_p^p = 0, \tag{21}$$

where $E_{j,l}$ is the projection from $V_j(p)$ onto the finite dimensional subspace $\mathrm{span}[\varphi_{j,k},\ |k| \leq l]$.

By an elementary density argument, the integral representation for E_j remains true in $\mathrm{L}^p(\mathbb{R})$:

$$E_j f(t) = \int_{\mathbb{R}} E_j(s,t) f(s) \, ds, \qquad f \in \mathrm{L}^p(\mathbb{R}). \tag{22}$$

Write $K_j(u)$ for $2^j K(2^j u)$ where K is the majorizing convolution kernel involved by (11). We have for each $f \in \mathrm{L}^p(\mathbb{R})$:

$$\|E_j f\|_p^p \leq \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} K_j(s-t) |f(s)| \, ds \Big|^p \, dt = \| \, K_j * |f| \, \|_p^p \leq \|K_j\|_1^p \, \|f\|_p^p$$

and hence (20) follows from the hypothesis (i).

Since each $\varphi_{j,k}$ ($k \in \mathbb{Z}$) has a compact support localized around $k2^{-j}$, there is a sequence $(A_l)_{l \in \mathbb{N}}$ increasing to $+\infty$ such that:

$$E_j(s,t) - E_{j,l}(s,t) = 0, \qquad |s| \leq A_l, \ t \in \mathbb{R}.$$

As a result, we have for each $f \in \mathrm{L}^p(\mathbb{R})$:

$$\begin{aligned} \mathbb{E} \, \|E_j f - E_{j,l} f\|_p^p &= \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} \big(E_j(s,t) - E_{j,l}(s,t)\big) f(s) \, ds \Big|^p \, dt \\ &\leq \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} K_j(s-t) |f(s)| \mathbb{1}_{\{|s| > A_l\}}(s) \, ds \Big|^p \, dt \\ &\leq \|K_j\|_1^p \int_{\{|s| > A_l\}} |f(s)|^p \, ds. \end{aligned}$$

So (21) follows from the hypothesis (ii).

It remains to verify the condition (19). For this part of the proof, the compacity of the support of φ is not needed. It follows from the proof of theorem 4 p. 33 in Meyer [13] that:

$$\int_{\mathbb{R}} E_j(s,t) \, dt = 1, \qquad s \in \mathbb{R}.$$

Consequently, for each $f \in \mathrm{L}^p(\mathbb{R})$, we have:

$$\begin{aligned} \|f - E_j f\|_p^p &= \int_{\mathbb{R}} \Big| \int_{\mathbb{R}} E_j(s,t)\big(f(s) - f(t)\big)\, dt \Big|^p ds \\ &\leq \int_{\mathbb{R}} b_j(s)^p \left(\int_{\mathbb{R}} \frac{|E_j(s,t)|}{b_j(s)} |f(s) - f(t)|\, dt \right)^p ds, \end{aligned}$$

where $b_j(s) = \int_{\mathbb{R}} |E_j(s,t)|\, dt$. By Jensen's inequality it follows:

$$\|f - E_j f\|_p^p \leq \int_{\mathbb{R}^2} b_j(s)^{p-1} |E_j(s,t)| |f(s) - f(t)|^p \, ds dt.$$

We now check that $b_j(s)$ is bounded uniformly in j and s. Recalling that for each integer m there is a constant a_m such that: $\varphi(t) \leq a_m(1 + |t|)^{-m}$, we get:

$$b_j(s) = \int_{\mathbb{R}} |E_0(2^j s, t)|\, dt \leq a_m^2 \int_{\mathbb{R}} \frac{dt}{(1 + |t|)^m} \sum_{k \in \mathbb{Z}} \frac{1}{(1 + |2^j s - k|)^m}. \tag{23}$$

Let us write $\theta_j(s)$ for the series in (23). Clearly $\theta_j(s) = \theta_0(2^j s)$ and hence $\|\theta_j\|_\infty = \|\theta_0\|_\infty$. Now, θ_0 is periodic with period 1 and continuous (assuming $m \geq 2$) so $\|\theta_0\|_\infty < +\infty$.

Going back to the stochastic processes ξ_n, we have then a constant B such that:

$$\mathbb{E}\, \|\xi_n - E_j \xi_n\|_p^p \leq B \int_{\mathbb{R}^2} |E_j(s,t)|\, \mathbb{E}\, |\xi_n(s) - \xi_n(t)|^p \, ds dt.$$

By the properties of the majorizing kernel K, there is for each integer m a constant c_m such that:

$$|E_j(s,t)| \leq \frac{c_m 2^j}{(1 + 2^j |s - t|)^m}, \qquad (s,t) \in \mathbb{R}^2.$$

The hypothesis (iii) yields now a constant B_m such that:

$$\mathbb{E}\, \|\xi_n - E_j \xi_n\|_p^p \leq B_m \int_{\mathbb{R}} g(t)\, dt \int_{\mathbb{R}} \frac{2^j |u|^\gamma}{(1 + 2^j |u|)^m}\, du.$$

Choosing $m \geq \gamma + 2$, we have by the change of variable $v = 2^j u$:

$$\mathbb{E}\, \|\xi_n - E_j \xi_n\|_p^p \leq B_m \|g\|_1 2^{-j\gamma} \int_{\mathbb{R}} \frac{|v|^\gamma}{(1 + |v|)^m}\, dv,$$

so (19) is satisfied and the proof is complete. ∎

4.2 Donsker random step functions

Let $(X_i)_{i\geq 1}$ be a sequence of centered random variables. Here the X_i need not have any special properties as independence or mixing. Write $S_0 = 0$, $S_n = \sum_{i=1}^n X_i$ and consider the Donsker random step functions ξ_n and the random broken lines ζ_n defined by:

$$\xi_n(t) = \sum_{i=0}^{n-1} \frac{S_i}{s_n} \mathbb{1}_{I_{n,i}}(t), \text{ where } I_{n,i} = \Big[\frac{i}{n}, \frac{i+1}{n}\Big[, \quad t \in [0,1], \tag{24}$$

$$\zeta_n(t) = \frac{S_{[nt]}}{s_n} + \big(nt - [nt]\big)\frac{X_{[nt]+1}}{s_n}, \quad t \in [0,1]. \tag{25}$$

Here the normalizing sequence (s_n) is of the form:

$$s_n = n^\delta L(n), \quad 0 < \delta < 1,$$

where L is a slowly varying function verifying $\inf_{x\geq 1} L(x) > 0$. The study of the tightness of $(\xi_n)_{n\geq 1}$ in $D(0,1)$ and of $(\zeta_n)_{n\geq 1}$ in $C(0,1)$ has been carried out to prove invariance principles. Here the tightness of $(\xi_n)_{n\geq 1}$ in $\mathrm{L}^p(0,1)$ $(1 < p < +\infty)$ is considered in itself, as an indicator of the more or less chaotic asymptotic behavior of $(\xi_n)_{n\geq 1}$.

Theorem 8 *The sequence $(\xi_n)_{n\geq 1}$ of random functions defined by (24) is tight in $\mathrm{L}^p(0,1)$ $(1 < p < +\infty)$ if there is a constant A such that:*

$$\mathbb{E}\left|\frac{S_{k+m} - S_k}{s_m}\right|^p \leq A, \quad m \in \mathbb{N}^*, \ k \in \mathbb{N}. \tag{26}$$

Corollary 1 *If the sequence $(X_i)_{i\geq 1}$ is stationary and verify:*

$$\mathbb{E}\left|\frac{S_m}{s_m}\right|^p \leq A, \quad m \in \mathbb{N}^*, \tag{27}$$

then $(\xi_n)_{n\geq 1}$ is tight in $\mathrm{L}^p(0,1)$ $(1 < p < +\infty)$.

Remark: Consider the linear injection T from $C(0,1)$ into $\mathrm{L}^p(0,1)$:

$$T: \ f \longmapsto \sum_{i=0}^{n-1} f\Big(\frac{i}{n}\Big)\mathbb{1}_{I_{n,i}}.$$

We have:

$$\|Tf\|_p = \bigg(\frac{1}{n}\sum_{i=0}^{n-1}\Big|f\Big(\frac{i}{n}\Big)\Big|^p\bigg)^{1/p} \leq \|f\|_\infty,$$

so T is continuous. Hence tightness of $(\zeta_n)_{n\geq 1}$ in $C(0,1)$ implies tightness of $(\xi_n)_{n\geq 1}$ in $\mathrm{L}^p(0,1)$. This bounds the field of interest of our result in the stationary case. Indeed, it easily follows from theorem 12.2 p. 94 in [4], that if (27) holds with $p > 2$ and $s_m = \sqrt{m}$ then:

$$P\big(\max_{1\le i\le m} |S_i| \ge \lambda\sqrt{m}\big) \le \frac{C}{\lambda^p},$$

and so $(\zeta_n)_{n\ge 1}$ is tight in $C(0,1)$.

Proof : Using the Schauder decomposition associated to the Haar basis (see section 2), we will verify the conditions (i) and (ii) of theorem 5 with $\alpha = \beta = p$. Observe first that:

$$\mathbb{E}\,\|\xi_n\|_p^p = \mathbb{E}\int_0^1 |\xi_n(t)|^p\,dt = \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left|\frac{S_i}{s_n}\right|^p \le A,$$

so (i) is satisfied.

Before treating the condition (ii), it seems convenient to recall here some useful facts about the E_j associated to the Haar basis. The projection operator E_j has an integral kernel: $E_j f(s) = \int_0^1 E_j(s,t)f(t)\,dt$, $f \in \mathrm{L}^p(0,1)$. It is easily verified that:

$$E_j(s,t) = 2^j\sum_{k=0}^{2^j-1}\mathbb{I}_{C_{j,k}}(s,t), \text{ where } C_{j,k} = \left[\frac{k}{2^j}, \frac{k+1}{2^j}\right[^2, \quad (s,t)\in[0,1]^2.$$

In other words, $(s,t)\mapsto E_j(s,t)$ is the uniform density over its support, the union C_j of the diagonal squares $C_{j,k}$ $(0\le k<2^j)$. It follows that $t\mapsto E_j(s,t)$ is the uniform density over the segment $[k2^{-j},\,(k+1)2^{-j}[$ which contains s. As $\int_0^1 E_j(s,t)\,dt = 1$, we have the following representation:

$$\|\xi_n - E_j\xi_n\|_p^p = \int_0^1\Big|\int_0^1 E_j(s,t)\sum_{i=0}^{n-1}\frac{S_i}{s_n}\big(\mathbb{I}_{I_{n,i}}(s) - \mathbb{I}_{I_{n,i}}(t)\big)\,dt\Big|^p ds. \quad (28)$$

We will find an upper bound for $\|\xi_n - E_j\xi_n\|_p^p$ in the two cases $n > 2^j$ and $n < 2^j$. If $n = 2^j$, by the mean value interpretation of E_j, we have $\xi_n = E_j\xi_n$.

The case $n > 2^j$: Recasting (28) as:

$$\|\xi_n - E_j\xi_n\|_p^p = \sum_{i=0}^{n-1}\int_{\frac{i}{n}}^{\frac{i+1}{n}}\Big|\sum_{l=0}^{n-1}\int_{\frac{l}{n}}^{\frac{l+1}{n}} E_j(s,t)\frac{S_i - S_l}{s_n}\Big|^p ds,$$

we get by convexity:

$$\begin{aligned}\|\xi_n - E_j\xi_n\|_p^p &\le \sum_{i,l=0}^{n-1}\Big|\frac{S_i-S_l}{s_n}\Big|^p\int_{\frac{i}{n}}^{\frac{i+1}{n}}\int_{\frac{l}{n}}^{\frac{l+1}{n}} E_j(s,t)\,dsdt & (29)\\ &\le \frac{2^j}{n^2}\sum_{i,l=0}^{n-1}\Big|\frac{S_i-S_l}{s_n}\Big|^p a(n,i,l), & (30)\end{aligned}$$

where $a(n,i,l)=1$ if the double integral in (29) is positive, $a(n,i,l)=0$ else. So the sum in (30) is composed of 2^j square blocks:

$$\sum_{i,l=0}^{n-1}\Big|\frac{S_i-S_l}{s_n}\Big|^p a(n,i,l)=\sum_{k=0}^{2^j-1}\sum_{(i,l)\in Q_{j,k,n}}\Big|\frac{S_i-S_l}{s_n}\Big|^p,$$

where the square index sets $Q_{j,k,n}=\Big\{(i,l)\ :\ \frac{k}{2^j}-\frac{1}{n}<\frac{i}{n},\ \frac{l}{n}<\frac{k+1}{2^j}\Big\}$, have side lenght $q_{j,k,n}$ bounded by: $n2^{-j}\le q_{j,k,n}\le n2^{-j}+2$. Renormalizing $|S_i-S_l|$ by some $m^\delta L(m)$, with $m=q_{j,k,n}$, we obtain:

$$\|\xi_n-E_j\xi_n\|_p^p\le\frac{2^j}{n^2}\Big(2^{-j}+\frac{2}{n}\Big)^{p\delta}\sum_{k=0}^{2^j-1}\Big(\frac{L(m)}{L(n)}\Big)^p\sum_{(i,l)\in Q_{j,k,n}}\Big|\frac{S_i-S_l}{m^\delta L(m)}\Big|^p.$$

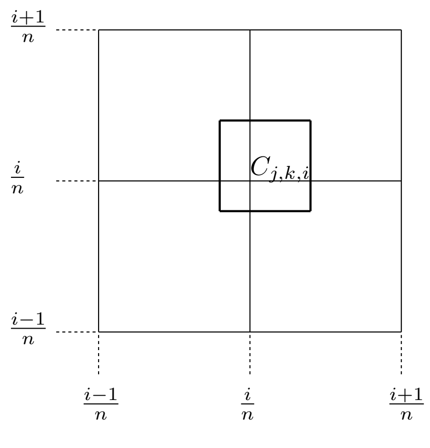
By the properties of slowly varying functions, we have $L(x)/L(y)\le a(y/x)^{\delta/2}$ $(1\le x\le y)$ for some constant a. Using the hypothesis (26), we have then some constant A' such that:

$$\mathbb{E}\,\|\xi_n-E_j\xi_n\|_p^p\le A'2^{-jp\delta/2},\quad (n>2^j).\tag{31}$$

The case $n<2^j$*:* In view of the diagonal localization of the support C_j, (28) can be recast as:

$$\|\xi_n-E_j\xi_n\|_p^p=\sum_{k=0}^{2^j-1}\int_{\frac{k}{2^j}}^{\frac{k+1}{2^j}}\Big|\int_{\frac{k}{2^j}}^{\frac{k+1}{2^j}}2^j\sum_{i=0}^{n-1}\frac{S_i}{s_n}\big(\mathbb{1}_{I_{n,i}}(s)-\mathbb{1}_{I_{n,i}}(t)\big)\,dt\Big|^p\,ds.$$

If $C_{j,k}$ is included in some $I_{n,i}^2$, the corresponding term in the summation over k above vanishes. Hence we have only to consider the "small" squares $C_{j,k,i}$ which cover the junction of two consecutives "big" squares, $I_{n,i-1}^2$ and $I_{n,i}^2$.



So we have:

$$\|\xi_n - E_j\xi_n\|_p^p = \sum_{i=1}^{n-1}\int_{\frac{k_i}{2^j}}^{\frac{k_i+1}{2^j}}\Big|\int_{\frac{k_i}{2^j}}^{\frac{k_i+1}{2^j}} 2^j\Big[\big(\mathbb{1}_{I_{n,i-1}}(s)-\mathbb{1}_{I_{n,i-1}}(t)\big)\frac{S_{i-1}}{s_n}+ \big(\mathbb{1}_{I_{n,i}}(s)-\mathbb{1}_{I_{n,i}}(t)\big)\frac{S_i}{s_n}\Big]\,ds\Big|^p\,dt \tag{32}$$

$$\leq 2^j\sum_{i=1}^{n-1}\int_{C_{j,k,i}}\Big|\big(\mathbb{1}_{I_{n,i-1}}(s)-\mathbb{1}_{I_{n,i-1}}(t)\big)\frac{S_{i-1}}{s_n}+ \big(\mathbb{1}_{I_{n,i}}(s)-\mathbb{1}_{I_{n,i}}(t)\big)\frac{S_i}{s_n}\Big|^p\,dsdt, \tag{33}$$

by the Jensen inequality. As the function under the integral in (33) vanishes on the traces of $I^2_{n,i-1}$ and $I^2_{n,i}$ over $C_{j,k,i}$, the integral in (33) is bounded by $s_n^{-p}|S_i - S_{i-1}|^p$ times the area of $C_{j,k,i}$. Hence:

$$\|\xi_n - E_j\xi_n\|_p^p \leq \frac{1}{2^j}\sum_{i=1}^{n-1}\Big|\frac{S_i - S_{i-1}}{s_n}\Big|^p, \qquad n < 2^j.$$

Taking the expectations, we get for some constant A'':

$$\mathbb{E}\,\|\xi_n - E_j\xi_n\|_p^p \leq \frac{A''}{2^j n^{p\delta-1}},$$

which provides the following uniform bounds in the case $n < 2^j$:

$$\mathbb{E}\,\|\xi_n - E_j\xi_n\|_p^p \leq \begin{cases} A''2^{-j} & \text{if} \quad p \geq 1/\delta \\ A''2^{-jp\delta} & \text{if} \quad 1 < p < 1/\delta. \end{cases} \tag{34}$$

Finally, condition (ii) of theorem 5 follows from (31) and (34). ■

4.3 Donsker random step functions in $\mathrm{L}^2(0,1)$

The following theorem about the particular case $p = 2$ was first published by Oliveira [14]. The proof used the Prohorov's theorem 1.13 [18], the reproducing space H_K with kernel $K(s,t) = 1 - \max(s,t)$ and some trigonometric basis of H_K. Oliveira and Suquet [15] presented a new version of the proof taking account of the rectification in Prohorov's theorem and working with any hilbertian basis of H_K. Now we are giving a third proof of the same result, in a more natural way without use of the reproducing kernel Hilbert spaces.

Theorem 9 *The sequence $(\xi_n)_{n\geq 1}$ defined by (24) is tight in $\mathrm{L}^2(0,1)$ if there is a constant B such that:*

$$\frac{1}{s_n^2}\sum_{i,l=1}^{n}\big|\mathbb{E}\,X_iX_l\big| \leq B, \quad n \geq 1. \tag{35}$$

Proof : We work again with the Haar basis. We use the representation of ξ_n involving directly the random variables X_i:

$$\xi_n(t) = \sum_{i=1}^{n} \frac{X_i}{s_n} \mathbb{I}_{J_{n,i}}(t), \text{ where } J_{n,i} = \left[\frac{i}{n}, 1 \right], \; t \in [0,1].$$

Put $h_{n,i}(s,t) = \mathbb{I}_{J_{n,i}}(s) - \mathbb{I}_{J_{n,i}}(t)$. By the properties of $E_j(s,t)$ recalled in the proof of theorem 8, we have:

$$\begin{aligned}
\|\xi_n - E_j\xi_n\|_2^2 &= \int_0^1 \Big| \sum_{i=1}^{n} \frac{X_i}{s_n} \int_0^1 E_j(u,t) h_{n,i}(u,t)\, dt \Big|^2 du \\
&= \sum_{i,l=1}^{n} \frac{X_i X_l}{s_n^2} \int_0^1 du \int_0^1 \int_0^1 E_j(u,s) E_j(u,t) h_{n,i}(u,s) h_{n,l}(u,t) ds dt \\
&= \frac{1}{s_n^2} \sum_{i,l=1}^{n} X_i X_l b(n,i,l),
\end{aligned}$$

where:

$$b(n,i,l) = \sum_{k=0}^{2^j-1} 2^{2j} \int_{[\frac{k}{2^j}, \frac{k+1}{2^j}[^3} h_{n,i}(u,s) h_{n,l}(u,t)\, ds dt du. \tag{36}$$

Observe that if $\frac{i+1}{n} \leq \frac{k}{2^j}$ or $\frac{i}{n} \geq \frac{k+1}{2^j}$, $h_{n,i}(u,s)$ vanishes on $[\frac{k}{2^j}, \frac{k+1}{2^j}[^2$ and else $|h_{n,i}|$ is bounded by 1. We see so there is at most one non null term in the right hand side of (36) under the constraints $\frac{k}{2^j} \leq \frac{i}{n} \leq \frac{k+1}{2^j}$ and $\frac{k}{2^j} \leq \frac{l}{n} \leq \frac{k+1}{2^j}$. Hence we always have $0 \leq b(n,i,l) \leq 2^{-j}$. Taking expectations, we get:

$$\mathbb{E}\, \|\xi_n - E_j\xi_n\|_2^2 = \frac{1}{s_n^2} \sum_{i,l=1}^{n} \mathbb{E}\, X_i X_l\, b(n,i,l) \leq B 2^{-j}.$$

So condition (ii) of theorem (5) is satisfied.

To check the condition (i), an elementary calculation yields:

$$\mathbb{E}\, \|\xi_n\|_2^2 = \frac{1}{s_n^2} \sum_{i,l=1}^{n} \mathbb{E}(X_i X_l) \Big(1 - \frac{\max(i,l)}{n}\Big) \leq B, \qquad n \geq 1.$$

∎

If the sequence $(X_i)_{i\geq 1}$ is independent or stationary or positively dependent, then (35) reduces to $s_n^{-2}\, \mathbb{E}\, S_n^2 \leq B$. But in the general case the absolute value cannot be dropped in (35) without extra assumption on the $(X_i)_{i\geq 1}$. We give now an example (with $s_n = n^{1/2}$) of sequence $(X_i)_{i\geq 1}$ verifying $n^{-1}\, \mathbb{E}\, S_n^2 \leq B$ but with $(\xi_n)_{n\geq 1}$ not tight in $\mathrm{L}^2(0,1)$.

Let X_1 be a random variable following the uniform distribution on $[-1,+1]$. Define for $n > 1$, $X_n = (-1)^{n+1}\big(\sqrt{n} + \sqrt{n+1}\big) X_1$. Thus we have $S_n =$

$(-1)^{n+1}\sqrt{n}X_1$ and $n^{-1}\,\mathbb{E}\,S_n^2 \leq 1$. Intuitively, the oscillations of (ξ_n) are too violent. To precise this idea, we will show the necessary condition (ii) in theorem 3 is violated by the subsequence $(\xi_{n_j})_{j\geq 1}$ where $n_j = 2^j - 1$. Going back to (32) with $p = 2$ and expanding the square instead of applying the Jensen inequality, we obtain:

$$\begin{aligned}\|\xi_{n_j} - E_j\xi_{n_j}\|_2^2 &= \sum_{i=1}^{n_j-1}\Big(\frac{S_i - S_{i-1}}{\sqrt{n_j}}\Big)^2 2^{2j}\Big[\Big(\frac{i}{n_j} - \frac{k_i}{2^j}\Big)\Big(\frac{k_i+1}{2^j} - \frac{i}{n_j}\Big)^2 + \\ &\qquad \Big(\frac{k_i+1}{2^j} - \frac{i}{n_j}\Big)\Big(\frac{i}{n_j} - \frac{k_i}{2^j}\Big)^2\Big] \\ &= \frac{2^j}{n_j}X_1^2\sum_{i=1}^{n_j-1}\big(\sqrt{i} + \sqrt{i-1}\big)^2\Big(\frac{i}{n_j} - \frac{k_i}{2^j}\Big)\Big(\frac{k_i+1}{2^j} - \frac{i}{n_j}\Big) \quad (37)\end{aligned}$$

Restricting the summation in (37) to the i verifying $2^j/4 \leq k_i \leq 3\,2^j/4$, we get a lower bound. Observing now that $\big|i(2^j-1)^{-1} - k2^{-j}\big|$ is minimal for $i - k = -1$ or 0, we have:

$$\min_{\frac14 2^j \leq k \leq \frac34 2^j}\Big|\frac{i}{2^j - 1} - \frac{k}{2^j}\Big| = \min_{\frac14 2^j \leq k \leq \frac34 2^j}\frac{\min(k, 2^j - k)}{(2^j-1)2^j} \geq \frac{1}{4(2^j-1)}\,.$$

Hence for $\frac14 2^j \leq k_i \leq \frac34 2^j$:

$$\Big(\frac{i}{n_j} - \frac{k_i}{2^j}\Big)\Big(\frac{k_i+1}{2^j} - \frac{i}{n_j}\Big) \geq \frac{C}{2^{2j}}$$

and

$$\|\xi_{n_j} - E_j\xi_{n_j}\|_2^2 \geq \frac{C}{2^{2j}}X_1^2\sum_{i:\frac14 2^j \leq k_i \leq \frac34 2^j} i.$$

Recall the integers k_i verify:

$$\frac{k_i}{2^j} \leq \frac{i}{2^j - 1} < \frac{k_i+1}{2^j}, \qquad i = 1, \ldots, n-1.$$

So we have for $j \geq 2$:

$$\sum_{i:\frac14 2^j \leq k_i \leq \frac34 2^j} i \;\geq \sum_{i:\frac14 2^j < i < \frac34 2^j} i \;\geq C' 2^{2j}.$$

Finally,

$$\|\xi_{n_j} - E_j\xi_{n_j}\|_2^2 \geq C'' X_1^2, \qquad j \geq 2.$$

Clearly, this contradicts condition (ii) of theorem 3.

4.4 Empirical process under strong mixing in $\mathrm{L}^p(0,1)$

Let $(X_n)_{n\geq 1}$ be a strictly stationary sequence of uniform variables on $[0,1]$. The strong mixing coefficients of this sequence are defined by:

$$\alpha_n = \sup\big\{|\mathrm{P}(A\cap B) - \mathrm{P}(A)\,\mathrm{P}(B)|\,,\ A\in\mathcal{F}_0^{k+1},\ B\in\mathcal{F}_{n+k}^{+\infty},\ k\in\mathbb{N}\big\}, \qquad (38)$$

where $\mathcal{F}_j^l$ stands for the σ-field generated by the variables $(X_i,\ j<i<l)$. We define the empirical processes related to $(X_n)_{n\geq 1}$ by:

$$\xi_n(t) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big(\mathbb{I}_{[X_i,1]}(t) - t\big), \qquad t\in[0,1],\ n\geq 1. \qquad (39)$$

In the $D(0,1)$ setting, under the assumption:

$$\alpha_n = O\big(n^{-3-\varepsilon}\big), \quad \varepsilon>0. \qquad (40)$$

Yoshihara [23] proved the weak convergence of the empirical process ξ_n to the centered gaussian process with covariance:

$$\Gamma(s,t) = s\wedge t - st + 2\sum_{k=2}^{+\infty}\big(P(X_1\leq s,\, X_k\leq t) - st\ \big). \qquad (41)$$

Recently, Oliveira and Suquet [16] proved the tightness of $(\xi_n)_{n\geq 1}$ in $\mathrm{L}^2(0,1)$ and its convergence to a gaussian process with covariance Γ given by (41) under the assumption:

$$\sum_{n\geq 1}\alpha_n < +\infty \qquad (42)$$

Thus there is some interest in investigating the tightness of $(\xi_n)_{n\geq 1}$ between the two conditions (40) and (42) and the two spaces $\mathrm{L}^2(0,1)$ and $D(0,1)$.

Theorem 10 *Suppose the mixing coefficients of $(X_i)_{i\geq 1}$ verify:*

$$|\alpha_i| \leq \frac{C}{i^{3-\varepsilon}}, \qquad i\geq 1, \qquad (43)$$

for some ε $(0\leq\varepsilon<2)$ and some constant C. Then the sequence $(\xi_n)_{n\geq 1}$ of empirical processes defined by (39) is tight in $\mathrm{L}^p(0,1)$ for each $p<6-2\varepsilon$.

Regarding ξ_n as a random function defined on $\mathbb{R}$ with support $[0,1]$, we can use the theorem 7. Of course, in this particular case, the hypothesis (ii) of this later theorem is trivially verified. Our main tool will be the following theorem.

Theorem 11 (Yokoyama[24]) *Let $(Y_i)_{i\geq 1}$ be a strictly stationary strong mixing sequence with $\mathbb{E}\,Y_1=0$ and $\mathbb{E}\,|Y_1|^{p+\delta}<+\infty$ for some $p>2$ and $\delta>0$. If its strong mixing coefficients $\alpha_i(Y)$, $i\geq 1$, verify:*

$$\sum_{i=1}^{+\infty}(i+1)^{p/2-1}\alpha_i(Y)^{\delta/(p+\delta)} < +\infty, \qquad (44)$$

then there is a constant K such that:

$$\mathbb{E}\Big|\sum_{i=1}^{n} Y_i\Big|^p \leq K n^{p/2}, \qquad n \geq 1. \tag{45}$$

A close examen of the proof shows we can choose: $K = A \max(1, \mathbb{E}\,|Y_1|^{p+\delta})$, where the constant A depends only on the mixing coefficients $(\alpha_i(Y),\, i \geq 1)$.

To verify condition (i) of theorem 7, observe that:

$$\mathbb{E}\big|\xi_n(t)\big|^p = n^{-p/2}\,\mathbb{E}\Big|\sum_{i=1}^{n}\big(\mathbb{I}_{[X_i,1]}(t) - t\big)\Big|^p.$$

Since the strong mixing coefficients of $Y_i = \mathbb{I}_{[X_i,1]}(t) - t$ are dominated by the α_i, the hypothesis (43) and the Yokohama's theorem give us, for a choice of p to be precised later, a constant C' such that $\mathbb{E}\,|\xi_n(t)|^p \leq C'$ uniformly in n, t.

Now, ξ_n being compactly supported by $[0,1]$, to verify the condition (iii) of theorem 7, it suffices to prove that for some constant B:

$$\mathbb{E}\big|\xi_n(t+u) - \xi_n(t)\big|^p \leq B|u|^\gamma, \qquad t,\ t+u \in [0,1],\ n \geq 1.$$

We have for t and $t+u$ in $[0,1]$:

$$\mathbb{E}\big|\xi_n(t+u) - \xi_n(t)\big|^p = n^{-p/2}\,\mathbb{E}\Big|\sum_{i=1}^{n} Z_i\Big|^p,$$

where

$$Z_i = \mathbb{I}_{[X_i,1]}(t+u) - \mathbb{I}_{[X_i,1]}(t) - u.$$

An elementary calculation yields:

$$\mathbb{E}\,|Z_i|^p = (1-|u|)|u|\big(|u|^{p-1} + (1-|u|)^{p-1}\big)$$

and then $\mathbb{E}\,|Z_i|^p \leq 2|u|$. Therefore, applying the Yokohama's theorem to the variables:

$$Y_i' = Z_i |u|^{-1/(p+\delta)},$$

we obtain for a choice of p to be precised below:

$$\mathbb{E}\big|\xi_n(t+u) - \xi_n(t)\big|^p \leq C''|u|^{p/(p+\delta)}, \qquad t \in [0,1].$$

It remains to examine the constraints on p. Clearly there is no upper bound in the choice of δ. Under (43), the general term of the series in (44) is an $O(i^\eta)$ with:

$$\eta = \frac{p}{2} - 1 - \frac{\delta(3-\varepsilon)}{\delta + p}. \tag{46}$$

As δ goes to $+\infty$, η goes to $p/2 - 4 + \varepsilon$, so the convergence condition $\eta < -1$ will be satisfied with $p < 6 - 2\varepsilon$ for δ large enough. ■

Theorem 12 *Under the hypotheses of theorem 10, the empirical process ξ_n converges weakly to ζ in $\mathrm{L}^p(0,1)$ for each $p < 6 - 2\varepsilon$.*

Proof : It remains only to check the convergence of characteristic functionals. For fixed p $(2 < p < 6 - 2\varepsilon)$, define q by $1/p + 1/q = 1$. We want to prove that for each f in $\mathrm{L}^q(0,1)$ and each t in $\mathbb{R}$:

$$\varphi_{f,n}(t) = \mathbb{E}\exp\Big(it\int_0^1 f(s)\xi_n(s)\,ds\Big) \longrightarrow \mathbb{E}\exp\Big(it\int_0^1 f(s)\zeta(s)\,ds\Big), \quad n \to +\infty \tag{47}$$

By the theorem of Oliveira Suquet [16] for the $\mathrm{L}^2(0,1)$ case, (47) holds for functions belonging to $\mathrm{L}^2(0,1)$ under the condition (42) which is weaker than (43). The general case follows from the density of $\mathrm{L}^2(0,1)$ in $\mathrm{L}^q(0,1)$ if we are able to keep control on the quantities:

(a) $\sup_{n\in\mathbb{N}}\big|\varphi_{f,n}(t) - \varphi_{g,n}(t)\big|$, $t \in \mathbb{R}$,

(b) $\Big|\mathbb{E}\Big(\exp it\int_0^1 \big(f(s) - g(s)\big)\zeta(s)\,ds\Big)\Big|$, $t \in \mathbb{R}$,

where $f \in \mathrm{L}^q(0,1)$ is approximated in the L^q norm by $g \in \mathrm{L}^2(0,1)$.

Using the inequality $|\mathrm{e}^{iu} - \mathrm{e}^{iv}| \le |u - v|$, we have:

$$\begin{aligned}\big|\varphi_{f,n}(t) - \varphi_{g,n}(t)\big| &\le |t|\,\mathbb{E}\int_0^1 |f(s) - g(s)||\xi_n(s)|\,ds\\ &\le |t|\;\|f - g\|_q \int_0^1 \mathbb{E}\,|\xi_n(s)|^p\,ds.\end{aligned}$$

As shown in the proof of theorem 10, $\mathbb{E}\,|\xi_n(s)|^p$ is bounded uniformly in n, s.

To bound (b), write:

$$\begin{aligned}\Big|\mathbb{E}\Big(\exp it\int_0^1 \big(f(s) - g(s)\big)\zeta(s)\,ds\Big)\Big| &\le |t|\int_0^1 |f(s) - g(s)|\,\mathbb{E}\,|\zeta(s)|\,ds\\ &\le |t|\;\|f - g\|_1 \sup_{s\in[0,1]} \mathbb{E}\,|\zeta(s)|\\ &\le |t|\;\|f - g\|_q \sup_{s\in[0,1]} \big(\mathbb{E}\,\zeta(s)^2\big)^{1/2}.\end{aligned}$$

Now under (42), the continuity of Γ on $[0,1]^2$ follows from the uniform convergence of the series in (41) and then $\Gamma(s,s) = \mathbb{E}\,\zeta(s)^2$ is bounded on $[0,1]$. ■

5 Application to the Hölderian spaces H_α^0

We turn now to the Hölderian spaces H_α^0. For the isomorphic properties (in Banach's sense) of these spaces, we refer to Ciesielski [5]. As shown by Baldi and Roynette [2], Kerkyacharian and Roynette [11], the spaces H_α^0 are more tractable than $C(0,1)$ to handle stochastic processes, like the brownian motion, whose paths have a regularity going beyond the simple continuity. For $0<\alpha<1$, H_α is the space of functions vanishing at zero and verifying:

$$\|f\|_{\mathrm{H}_\alpha} = \sup_{s,t\in[0,1]} \frac{|f(t)-f(s)|}{|t-s|^\alpha} < +\infty. \tag{48}$$

H_α^0 is the subspace of elements f in H_α satisfying the additional condition:

$$\lim_{\delta\to 0} \sup_{\substack{|s-t|\le\delta \\ s,t\in[0,1]}} \frac{|f(t)-f(s)|}{|t-s|^\alpha} = 0. \tag{49}$$

H_α and H_α^0 equipped with the norm $\| \ \|_{\mathrm{H}_\alpha}$ are both Banach spaces but only H_α^0 is separable.

It follows from theorem 2 in [5] that the Faber-Schauder basis is a Schauder basis for H_α^0 and hence H_α^0 has the same Schauder decomposition as $C(0,1)$ (of course with a different topology). Moreover $f\in\mathrm{H}_\alpha$ if and only if there exists a constant C such that the Faber-Schauder coefficients $a_{j,k}$ of f verify:

$$|a_0|\le C, \qquad 2^{j\alpha}|a_{j,k}|\le C, \qquad j\in\mathbb{N},\ 0\le k<2^j \tag{50}$$

and $f\in\mathrm{H}_\alpha^0$ if and only if it verify (50) and:

$$\lim_{j\to+\infty} \max_{0\le k<2^j} 2^{j\alpha}|a_{j,k}| = 0. \tag{51}$$

Here the $a_{j,k}$ are defined by:

$$a_0 = f(1), \quad a_{j,k} = f\Big(\frac{2k+1}{2^{j+1}}\Big) - \frac{1}{2}\Big[f\Big(\frac{k}{2^j}\Big) + f\Big(\frac{k+1}{2^j}\Big)\Big]. \tag{52}$$

For $f\in\mathrm{H}_\alpha$, define $\|f\|$ as the infimum of constants C verifying (50). Then $\|f\|$ is a norm on H_α equivalent to the initial norm $\|f\|_{\mathrm{H}_\alpha}$.

The theorem 3 provides the following tightness criterion in H_α^0:

Theorem 13 *The sequence $(\xi_n)_{n\ge1}$ of random elements in H_α^0 is tight if and only if:*

(a) $\lim_{A\to+\infty} \sup_{n\ge1} P\big(\|\xi_n\|_\infty \ge A\big) = 0,$

(b) $\forall\varepsilon>0,\ \lim_{\delta\to0} \sup_{n\ge1} P\left(\sup_{|t-s|\le\delta} \frac{|\xi_n(t)-\xi_n(s)|}{|t-s|^\alpha} \ge \varepsilon\right) = 0.$

Proof : *Sufficiency of* (a) *and* (b): Write $a^n_{j,k}$ for the Faber-Schauder coefficients of ξ_n. Clearly we have:

$$\begin{aligned} \|E_j\xi_n\| &= \max_{i\le j}\max_{0\le k<2^i}|a^n_{i,k}|2^{i\alpha} \\ \|\xi_n - E_j\xi_n\|, &= \sup_{i>j}\max_{0\le k<2^i}|a^n_{i,k}|2^{i\alpha}. \end{aligned}$$

Hence, by (52), for fixed j, $\|E_j\xi_n\| \le 2\,2^{j\alpha}\|\xi_n\|_\infty$ so the condition (i) of theorem 3 follows from (a). Next writing:

$$a^n_{i,k} = \frac{1}{2}\Big[\xi_n\Big(\frac{2k+1}{2^{i+1}}\Big) - \xi_n\Big(\frac{2k}{2^{i+1}}\Big)\Big] + \frac{1}{2}\Big[\xi_n\Big(\frac{2k+1}{2^{i+1}}\Big) - \xi_n\Big(\frac{2k+2}{2^{i+1}}\Big)\Big],$$

we have:

$$|a^n_{i,k}|2^{i\alpha} \le 2^{-\alpha}\sup_{|t-s|\le 2^{-i-1}}\frac{|\xi_n(t)-\xi_n(s)|}{|t-s|^\alpha}$$

and

$$\|\xi_n - E_j\xi_n\| \le 2^{-\alpha}\sup_{|t-s|\le 2^{-j}}\frac{|\xi_n(t)-\xi_n(s)|}{|t-s|^\alpha},$$

which shows that (b) implies the condition (ii) of theorem 3.

Necessity of (a) *and* (b): The necessity of (a) follows obviously from the continuity of the canonical injection from H_α into $C(0,1)$ (for each $f \in \mathrm{H}_\alpha$, $\|f\|_\infty \le \|f\|_{\mathrm{H}_\alpha}$). To verify the necessity of (b), let us define the functional:

$$T_\delta :\ \mathrm{H}_\alpha \longrightarrow \mathbb{R}^+, \qquad T_\delta(f) = \sup_{|t-s|\le\delta}\frac{|f(t)-f(s)|}{|t-s|^\alpha}.$$

It is easily checked that for fixed δ, $T_\delta(f)$ is a norm on H_α, equivalent to $\|f\|_{\mathrm{H}_\alpha}$. As a result, T_δ is continuous and we can apply the lemma 1 with the sequence of closed subsets F_l of H^0_α (itself closed in H_α) defined by:

$$F_l = \{f \in \mathrm{H}^0_\alpha,\ T_{\delta_l}(f) \ge \varepsilon\},$$

where δ_l decreases to zero. This yields the necessity of (b). ■

As in the $C(0,1)$ case (Billingsley [4] th. 8.3), the supremum over δ in (b) can be somewhat localized.

Theorem 14 *The sequence* $(\xi_n)_{n\ge1}$ *of random elements in* H^0_α *is tight if:*

(a) $\lim_{A\to+\infty}\sup_{n\ge1} P\big(\|\xi_n\|_\infty \ge A\big) = 0,$

(c) *For each positive* ε *and* η*, there exists a* δ*, with* $0<\delta<1$*, and an integer* n_0 *such that:*

$$\frac{1}{\delta}P\Big(\sup_{u\le s,t\le u+\delta}\frac{|\xi_n(t)-\xi_n(s)|}{|t-s|^\alpha} \ge \varepsilon\Big) \le \eta, \quad n \ge n_0,$$

for all u *in* $[0,1[$.

Of course when $u > 1 - \delta$, the supremum is restricted to $u \le s, t \le 1$. The proof is obtained by an easy adaptation of the arguments given in the $C(0,1)$ case and will be omitted.

Another sufficient condition for tightness in H^0_α is the following proved by Kerkyacharian and Roynette using the Faber Schauder coefficients. The moment form of this condition was obtained first by Lamperti [12] by an another method.

Theorem 15 (Kerkyacharian Roynette [11]) *Let $(\xi_n)_{n\ge 1}$ be a sequence of processes, vanishing at 0 and verifying:*

$$\forall \lambda > 0, \quad P\big(|\xi_n(t) - \xi_n(s)| > \lambda\big) \le \frac{c}{\lambda^\gamma}|t-s|^{1+\tau}, \quad \textit{for some } \tau > 0,\ \gamma > 0.$$

Then $(\xi_n)_{n\ge 1}$ is tight in H^0_α for $\alpha < \tau/\gamma$.

From a practical point of view, this condition is more tractable than theorem 13. The usefulness of theorem 13 seems be in its necessary part. Let us consider, for instance, the following result of Lamperti about the convergence of Gaussian processes.

Proposition 2 (Lamperti [12]) *Let $(\xi_n)_{n\ge 1}$ be a sequence of centered Gaussian processes with covariance functions $\rho_n(s,t) = \mathbb{E}\big(\xi_n(s)\xi_n(t)\big)$. Suppose $\rho_n(s,t)$ converges to $\rho(s,t)$ and there exists constants $\alpha \in (0,1]$ and $B < +\infty$ such that:*

$$\big|\rho_n(s,s) - 2\rho_n(s,t) + \rho_n(t,t)\big| \le B|t-s|^{2\alpha}, \quad s,t \in [0,1],\ n \ge 1. \tag{53}$$

Then there is a separable centered Gaussian process ξ with covariance $\rho(s,t)$ and paths belonging a.s. to H^0_γ for every $\gamma < \alpha$ such that $(\xi_n)_{n\ge 1}$ converges weakly to ξ in H^0_γ.

Applying theorem 13, we can see there is not a great loss in using condition (53) to prove the weak convergence of ξ_n in H^0_γ.

Proposition 3 *With the notations of proposition 2, suppose $(\xi_n)_{n\ge 1}$ converges weakly to ξ in H^0_γ, then:*

$$\big|\rho_n(s,s) - 2\rho_n(s,t) + \rho_n(t,t)\big| = o\big(|t-s|^{2\gamma}\big), \quad \textit{uniformly in } s,\ t.$$

Proof : For notational convenience, define $F(x) = \dfrac{2}{\sqrt{2\pi}} \displaystyle\int_x^{+\infty} \exp\big(-u^2/2\big)\, du$ and:

$$Z(n,s,\delta) = \frac{\xi_n(s) - \xi_n(s+\delta)}{\delta^\gamma}, \quad \sigma^2(n,s,\delta) = \mathbb{E}\, Z^2(n,s,\delta).$$

From condition (b) of theorem 13 it follows:

$$\lim_{\delta\to 0} \sup_{n,s} P\big(|Z(n,s,\delta)| > \varepsilon\big) = \lim_{\delta\to 0} \sup_{n,s} F\Big(\frac{\varepsilon}{\sigma(n,s,\delta)}\Big) = 0.$$

As F decreases, this implies:

$$\lim_{\delta\to 0} \sup_{n,s} \sigma^2(n,s,\delta) = 0,$$

from which the conclusion follows. ■

We close this section with an example illustrating the use of the isomorphism between H_α^0 and the space c_0 of sequences converging to zero equipped with the norm of supremum. Define the normalized Ciesielski basis of H_α^0 by:

$$\Delta_i^{(\alpha)}(t) = \|\Delta_i\|_{\mathrm{H}_\alpha}^{-1} \Delta_i(t), \qquad t\in[0,1],\ i\in\mathbb{N}.$$

Then $f\in \mathrm{H}_\alpha^0$ if and only if:

$$\forall t\in[0,1], \quad f(t) = \sum_{i=0}^{+\infty} a_i \Delta_i^{(\alpha)}(t), \quad \text{with } (a_i)_{i\in\mathbb{N}} \in \mathrm{c}_0.$$

The operator $T:\ f\mapsto (a_i)_{i\in\mathbb{N}}$ is an isomorphism of Banach spaces between H_α^0 and c_0 [5].

Proposition 4 *Consider the sequence $(\xi_n)_{n\geq 1}$ of random elements in H_α^0 defined by:*

$$\xi_n(t) = \sum_{i=0}^{+\infty} X_{n,i}\, \Delta_i^{(\alpha)}(t), \qquad t\in[0,1],$$

where for each $n\geq 1$, $\big(X_{n,i}\big)_{i\in\mathbb{N}}$ is a sequence of independent random variables which converges to zero almost surely. Then $(\xi_n)_{n\geq 1}$ is tight in H_α^0 if and only if:

$$(a) \quad \lim_{A\to+\infty} \sup_{n\geq 1} \sum_{i=0}^{+\infty} P\big(|X_{n,i}| > A\big) = 0,$$

$$(b)\ \forall\varepsilon, \quad \lim_{j\to+\infty} \sup_{n\geq 1} \sum_{i=j}^{+\infty} P\big(|X_{n,i}| > \varepsilon\big) = 0.$$

Proof : The isomorphism T shift the problem to the tightness of $\big(T(\xi_n)\big)_{n\geq 1}$ in c_0. Using the Schauder decomposition associated to the canonical basis of c_0 and the independence of the $X_{n,i}$ ($i\in\mathbb{N}$), we can write the conditions (i') and (ii) of theorem 4 under the form:

$$(i') \quad \lim_{A\to+\infty} \prod_{i=0}^{+\infty} \big(1 - P(|X_{n,i}| > A)\big) = 1, \text{ uniformly in } n,$$

$$(ii)\ \forall\varepsilon, \quad \lim_{j\to+\infty} \prod_{i=j}^{+\infty} \big(1 - P(|X_{n,i}| > \varepsilon)\big) = 1, \text{ uniformly in } n.$$

The result follows then by elementary arguments about the comparison of infinite products and series. ■

6 Sequences spaces and Besov spaces

In all the preceeding examples, but for the last, no explicit computation of $\|x - \mathbb{E}_j x\|$ $(x \in \mathcal{X})$ was needed. We only made use of the connection between $\|x - \mathbb{E}_j x\|$ and the oscillations of x (in a sense depending on the functional space $\mathcal{X}$). We close the paper with some cases where the computation of $\|x - \mathbb{E}_j x\|$ is easy: the sequences spaces $\ell^p(\mathbb{N})$ and the Besov spaces.

The canonical basis of $\ell^p(\mathbb{N})$ $(1 \leq p < +\infty)$ being a Schauder one, the theorems 3 and 5 have the following obvious translation in the case of discrete time process ξ_n with paths in $\ell^p(\mathbb{N})$:

$$\xi_n : (\omega, i) \mapsto \xi_n(\omega, i), \qquad (\omega, i) \in \Omega \times \mathbb{N}, \quad \sum_{i=0}^{+\infty} |\xi_n(\omega, i)|^p < +\infty.$$

Theorem 16 *The sequence of random elements $(\xi_n)_{n\geq 1}$ is tight in $\ell^p(\mathbb{N})$ $(1 \leq p < +\infty)$ if and only if:*

(*i*) *For each $i \in \mathbb{N}$, $\big(\xi_n(i)\big)_{n\geq 1}$ is tight in $\mathbb{R}$,*

(*ii*) *For each positive ε, $\lim_{j\to+\infty} \sup_{n\geq 1} P\Big(\sum_{i>j} |\xi_n(i)|^p > \varepsilon\Big) = 0.$*

Corollary 2 *The sequence $(\xi_n)_{n\geq 1}$ is tight in $\ell^p(\mathbb{N})$ $(1 \leq p < +\infty)$ if:*

(*i*) $\sup_{n\geq 1} \sum_{i\in\mathbb{N}} \mathbb{E}\,|\xi_n(i)|^p < +\infty,$

(*ii*) $\lim_{j\to+\infty} \sup_{n\geq 1} \sum_{i>j} \mathbb{E}\,|\xi_n(i)|^p = 0.$

Remarks: Of course we have excluded the case $p = +\infty$ since $\ell^\infty(\mathbb{N})$ is not separable.

The case of $\ell^p(\mathbb{Z})$ $(1 \leq p < +\infty)$ is similar to $\ell^p(\mathbb{N})$. The canonical basis $(e_j, j \in \mathbb{Z})$ of $\ell^p(\mathbb{Z})$ is an unconditional Schauder basis. If we choose the Schauder decomposition with $V_j = \operatorname{span}[e_i,\ -j \leq i \leq j]$ the theorem 16 and its corollary remain valid replacing $\mathbb{N}$ by $\mathbb{Z}$ in (*i*) and $i > j$ by $|i| > j$ in (*ii*).

The Besov spaces $B_p^{s,q}$ provide a ladder of spaces generalizing the Sobolev and Hölder spaces. Their usefulness in functional estimation has been illustrated by the recents papers of Kerkyacharian and Picard [9], [10].

Several equivalent definitions of the Besov spaces are available. We follow here Meyer [13] p. 49. Some other definitions can be found in Bergh and

Löfstrom [3], Triebel [22]. In particular, theorem 6.2.5. in Bergh and Löfstrom presents an alternative definition in terms of derivatives and L^p-moduli of continuity.

Consider now a multiresolution analysis of $\mathrm{L}^2(\mathbb{R})$ with regularity r. According to Meyer, for $1 \le p \le +\infty$, $1 \le q \le +\infty$, $0 < s < r$, the Besov space $B_p^{s,q} = B_p^{s,q}(\mathbb{R})$ can be viewed as a function space equipped with one of the two equivalent norms:

$$\begin{aligned} J_p^{s,q} &= \|E_0 f\|_{\mathrm{L}^p(\mathbb{R})} + \left(\sum_{j \ge 0} \left(2^{js} \|D_j f\|_{\mathrm{L}^p(\mathbb{R})} \right)^q \right)^{1/q} \\ K_p^{s,q} &= \|\alpha_{0,\cdot}\|_{\ell^p(\mathbb{Z})} + \left(\sum_{j \ge 0} \left(2^{j(s+1/2-1/p)} \|\beta_{j,\cdot}\|_{\ell^p(\mathbb{Z})} \right)^q \right)^{1/q} \end{aligned}$$

where the sequences $\alpha_{0,\cdot} = \left(\alpha_{0,k}\right)_{k \in \mathbb{Z}}$ and $\beta_{j,\cdot} = \left(\beta_{j,k}\right)_{k \in \mathbb{Z}}$ are sequences of wavelets coefficients given by:

$$E_0 f = \sum_{k \in \mathbb{Z}} \alpha_{0,k} \varphi_{0,k}, \qquad D_j f = \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k}, \quad j \in \mathbb{N}.$$

This definition is intrinsic: for each multiresolution analysis with regularity $r' > s$, we find the same space $B_p^{s,q}$ of functions equipped with norms ${J'}_p^{s,q}$ and ${K'}_p^{s,q}$ equivalent to $J_p^{s,q}$ and $K_p^{s,q}$.

The use of the norm $K_p^{s,q}$ reduces the Besov space $B_p^{s,q}$ to a sequence space. This allows us to give a tightness criterion involving the wavelets coefficients.

Theorem 17 *The sequence of random elements $(\xi_n)_{n \ge 1}$ in the Besov space $B_p^{s,q}$ ($s > 0$, $1 \le p < +\infty$, $1 \le q < +\infty$) is tight if and only if for some multiresolution analysis with regularity $r > s$:*

(i) $\left(\alpha_{0,\cdot}(\xi_n)\right)_{n \ge 1}$ *is tight in* $\ell^p(\mathbb{Z})$,

(ii) For each $j \ge 0$, $\left(\beta_{j,\cdot}(\xi_n)\right)_{n \ge 1}$ *is tight in* $\ell^p(\mathbb{Z})$,

(iii) $\forall \varepsilon > 0$, $\lim_{j \to +\infty} \sup_{n \ge 1} P\left(\sum_{i > j} \left(2^{j(s+1/2-1/p)} \|\beta_{j,\cdot}(\xi_n)\|_{\ell^p(\mathbb{Z})} \right)^q > \varepsilon \right) = 0,$

where for each n, the sequences $\alpha_{0,\cdot}(\xi_n)$ and $\beta_{j,\cdot}(\xi_n)$ are given by:

$$\alpha_{0,k}(\xi_n) = \int_{\mathbb{R}} \xi_n(t) \varphi_{0,k}(t)\, dt, \qquad \beta_{j,k}(\xi_n) = \int_{\mathbb{R}} \xi_n(t) \psi_{j,k}(t)\, dt.$$

Proof : Obvious. ■

References

[1] Akhiezer, N., I. and Glazman, I., M., Theory of linear operators in Hilbert space. Pitman (1980)

[2] Baldi, P., Roynette, B., Some equivalents for the Brownian motion in Hölder norm. Probab. Theory Relat. Fields 93 (1992), 457–484.

[3] Bergh, J. and Löfstrom, J., Interpolation Spaces: an Introduction, Springer (1976).

[4] Billingsley, P., Convergence of probability measures, Wiley (1968).

[5] Ciesielski, Z., On the isomorphisms of the spaces H_α and m, Bull. Acad. Pol. Sci. Sr. Sci. Math. Astronom. Phys. 8 (1960), 217–222.

[6] Ciesielski, Z., Hölder conditions for realizations of Gaussian processes, Trans. Amer. Math. Soc. 99 (1961), 403–413.

[7] Daubechies, I., Ten lectures on wavelets. SIAM (1992).

[8] Gikhman, I., I. and Skorohod A., V., Theory of stochastic processes I. Springer (1974).

[9] Kerkyacharian, G., Picard, D., Density estimation in Besov spaces, Statist. Probab. Lett. 13 (1992), 15–24.

[10] Kerkyacharian, G., Picard, D., Density estimation by kernel and wavelets methods: Optimality of Besov spaces, Statist. Probab. Lett. 18 (1993), 327–336.

[11] Kerkyacharian, G., Roynette, B., Une dmonstration simple des thormes de Kolmogorov, Donsker et Ito-Nisio, C. R. Acad. Sci. Paris, t. 312, Srie I (1991), 877–882.

[12] Lamperti, J., On convergence of stochastic processes, Trans. Amer. Math. Soc., 104 (1962), 430–435.

[13] Meyer, Y., Ondelettes et opérateurs I, Hermann (1990).

[14] Oliveira, P., E., Invariance principles in $\mathrm{L}^2(0,1)$, Comment. Math. Univ. Carolinae 31(1990), 357-366.

[15] Oliveira, P., E., Suquet, Ch., Auto-reproducing spaces and invariance principles in $\mathrm{L}^2(0,1)$, Preprint, Publ. IRMA Lille 32(1993) - III.

[16] Oliveira, P., E., Suquet, Ch., $\mathrm{L}^2(0,1)$ weak convergence of the empirical process for dependent variables. Preprint, Publ. IRMA Lille 34 (1994) - XII.

[17] Parthasarathy K., R., Probability measures on metric spaces. Academic Press (1967).

[18] Prohorov Y., V., Convergence of random processes and limit theorems in probability theory. Theor. Prob. Appl. 1, 157–214 (1956).

[19] Singer, I., Bases in Banach Spaces I, Springer (1970).

[20] Singer, I., Bases in Banach Spaces II, Springer (1981).

[21] Suquet, Ch., Some remarks about weak compactness of sets of probability measures on Hilbert spaces, Preprint.

[22] Triebel, H., Theory of Function Spaces II, Birkhaüser (1992).

[23] Yoshihara, K., Billingsley's theorems on empirical processes of strong mixing sequences, Yokohama Math. J. 23 (1975), 1–7.

[24] Yokoyama, R., Moment bounds for stationary mixing sequences, Z. Wahrsch. Verw. Gebiete 52 (1980), 45–57.