



HAL
open science

Asymptotic Normality in Banach Spaces via Lindeberg Method

Alfredas Račkauskas, Charles Suquet

► **To cite this version:**

Alfredas Račkauskas, Charles Suquet. Asymptotic Normality in Banach Spaces via Lindeberg Method. *Journal of Theoretical Probability*, 2022, 36 (1), pp.409-455. 10.1007/s10959-022-01177-x. hal-04504095

HAL Id: hal-04504095

<https://hal.science/hal-04504095v1>

Submitted on 14 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic normality in Banach spaces via Lindeberg method

Alfredas Račkauskas · Charles Suquet

Received: date / Accepted: date

Abstract The relation between weak convergence of probabilities on a smooth Banach space and uniform convergence over certain class of smooth functions is established. This leads to an extension of Lindeberg's proof of the central limit theorem in Banach space framework. As a result, asymptotic normality is proved for sums of Banach space random variables including triangular arrays and weighted linear processes.

Keywords Lindeberg method · smooth Banach space · triangular array

Mathematics Subject Classification (2020) MSC 60B10 · 60B12 · 60F05

1 Introduction

In this paper we consider sums of Banach space valued random variables with the aim to establish conditions for their asymptotic normality using a generalization of Lindeberg's elementary proof of the central limit theorem of 1922 [18].

Lindeberg's method is simple and elegant. It is based on replacing one by one non-Gaussian random variables by Gaussian ones and then using Taylor expansion to get approximation bounds. This principle has been applied for proving central limit theorems for sums of independent random variables with values in a Hilbert space by Giné and León [12], or to estimate rates of convergence in a central limit theorem in Banach spaces (see, e.g., Bentkus *et al.* [4], Paulauskas and Račkauskas [19]) as well.

The research is supported by the Research Council of Lithuania, grant No. S-MIP-17-76.

A. Račkauskas
Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, LT-03225 Vilnius, Lithuania.
E-mail: alfredas.rackauskas@mif.vu.lt

Ch. Suquet
Univ. Lille, CNRS, UMR 8524 - Laboratoire Paul Painlevé, F-59000 Lille, France.
E-mail: charles.suquet@univ-lille.fr

Its potential for proving more general invariance results has been discovered by many researchers and led to results on matrices with exchangeable entries (Chatterjee [7]), on the universality of local laws (Tao and Vu [24], [25]) on matrices with correlated entries (Banna [1], Banna and Merlevède [2], Banna, Merlevède and Peligrad [3]) and many others. Various kernel type density (regression function) estimators produce sums of an array of random functions to which Lindeberg CLT can be applied to help solve various statistical problems (see, e.g. [13] and references therein).

The idea of Lindeberg was carefully examined and generalized by Zolotarev [26] through the introduction of the so-called ζ metrics, which metrizes weak convergence in the case of distributions in Hilbert space as later proved by Giné and León [12]. Although for Banach spaces such metrization of weak convergence is not possible in general however it is possible to connect weak convergence of probability measures with uniform convergence over a suitable class of differentiable functions in the case where the norm of Banach space is smooth enough (see Section 2). This leads to an extension of Lindeberg method for smooth Banach spaces. In turn we prove central limit theorem for a triangular array of random elements with values in such Banach spaces (see Section 4) and establish asymptotic normality of sums of Banach space valued linear processes as well of weighted sums of independent identically distributed B -valued random variables (Section 5). In Section 3 we present some remarks concerning differentiability of norm and some examples of smooth Banach spaces.

Abstract theory of smoothness in infinite-dimensional real Banach spaces and its connections with geometrical properties have been investigated by many authors. For a very detailed exposition of the theory, we refer to the book by Hájek and Johannis [14]. The existence of a p -smooth bump function for $1 < p \leq 2$, was shown to be equivalent to certain martingale moment inequality and appeared as sufficient property to some probability limit theorems in Banach spaces, see e.g., Pisier [20] and Rosiński [22].

2 Weak convergence via smooth functions

In what follows B denotes a real *separable* Banach space. The norm of an element $x \in B$ is denoted by $\|x\|_B$, or, if no confusion can arise, simply by $\|x\|$. The Banach space topological dual of B is B^* and we shall use the notation $\langle x, y^* \rangle := y^*(x)$ for the duality pairing of the elements $x \in B$ and $y^* \in B^*$. Let $L(B)$ be the Banach space of all continuous linear operators $u : B \rightarrow B$, endowed with the norm $\|u\| = \sup\{\|ux\| : x \in B, \|x\| \leq 1\}$, $I_B \in L(B)$ is the identity operator. By $\mathcal{L}_k(B)$ we denote the Banach space of bounded k -linear operators $T : B^k \rightarrow \mathbb{R}$ with the supremum norm

$$\|T\| = \sup\{|T(h_1, \dots, h_k)| : \|h_1\| \leq 1, \dots, \|h_k\| \leq 1\}.$$

To simplify the writing, the overloading of the notation $\|z\|$ for the norm is used throughout the text whenever the nature of the argument z dispels any doubt on the Banach space involved : B or one of the associated spaces of continuous operators $L(B)$, $\mathcal{L}_k(B)$, $k > 1$.

The set of all probability distributions on the measurable space (B, \mathcal{B}_B) is denoted by $\mathcal{P}(B)$, where \mathcal{B}_B is the σ -algebra of Borel subsets of B . Throughout we use

$$Pf := \int_B f(x)P(dx)$$

for any probability $P \in \mathcal{P}(B)$ and any P -integrable function $f : B \rightarrow \mathbb{R}$. Let us recall here that a sequence $(P_n) \subset \mathcal{P}(B)$ converges weakly to $P \in \mathcal{P}(B)$ (denoted $P_n \xrightarrow[n \rightarrow \infty]{w} P$) if

$$\lim_{n \rightarrow \infty} P_n f = Pf \text{ for each } f \in C_b(B),$$

where $C_b(B)$ is the class of all bounded continuous functions $f : B \rightarrow \mathbb{R}$. It is sometimes convenient to prove weak convergence of probability measures by showing that $P_n f \rightarrow Pf$ for a class \mathcal{F} of functions $f : B \rightarrow \mathbb{R}$ which is smaller than the class $C_b(B)$. In this case it is said that \mathcal{F} determines the weak convergence of probabilities. A well known example is provided by the class of bounded Lipschitz functions. Recall the function $f : B \rightarrow \mathbb{R}$ is bounded Lipschitz if

$$\|f\|_{\text{Lip}} := \sup_{x \in B} |f(x)| + \sup_{x \neq y} \frac{|f(x) - f(y)|}{\|x - y\|} < \infty.$$

Moreover, the bounded Lipschitz distance

$$d_{\text{BL}}(P, Q) = \sup \left\{ |Pf - Qf| : \|f\|_{\text{Lip}} \leq 1 \right\}, \quad P, Q \in \mathcal{P}(B),$$

metrizes weak convergence that is $P_n \xrightarrow[n \rightarrow \infty]{w} P$ if and only if $\lim_{n \rightarrow \infty} d_{\text{BL}}(P_n; P) = 0$.

Another example is known in the case where $B = \mathcal{H}$ is a separable Hilbert space. As proved by Giné and León [12], for $(P, P_n, n \in \mathbb{N}) \subset \mathcal{P}(\mathcal{H})$, in order to check weak convergence $P_n \xrightarrow[n \rightarrow \infty]{w} P$ it is enough to show that $P_n f \xrightarrow[n \rightarrow \infty]{} Pf$ for every $f : \mathcal{H} \rightarrow \mathbb{R}$ continuous, bounded and with bounded derivatives of all orders. So the situation in a separable Hilbert space is just as in the case of a finite dimensional space.

In what follows for a number $p \geq 1$ we denote by $\lfloor p \rfloor$ the unique integer satisfying $p - 1 \leq \lfloor p \rfloor < p$ and agree that $\{p\} := p - \lfloor p \rfloor$. The reader is warned about the difference with the classical ‘‘floor’’ and ‘‘fractional part’’ functions, e.g. $\lfloor 3.9 \rfloor = 3$, but $\lfloor 4 \rfloor = 3$ and $\{n\} = 1$ for any integer n . This is motivated by our wish to interpolate between spaces C_b^{n-1} and C_b^n of functions with respectively bounded $(n-1)$ derivatives or n derivatives by spaces of functions whose $(n-1)$ th derivatives satisfy a Hölder condition with exponent $0 < \alpha \leq 1$, the special case $\alpha = 1$ giving Lipschitz $(n-1)$ derivatives.

More precisely we introduce for any real $p \geq 1$ the class $C_b^{(p)}(B)$ of functions $f \in C_b(B)$ that are $\lfloor p \rfloor$ -times continuously Fréchet differentiable and such that

$$\|f\|_{(p)} := \sum_{k=0}^{\lfloor p \rfloor} \sup_{x \in B} \|f^{(k)}(x)\| + \sup_{x \neq y} \frac{\|f^{(\lfloor p \rfloor)}(x) - f^{(\lfloor p \rfloor)}(y)\|}{\|x - y\|^{\{p\}}} < \infty, \quad (1)$$

where $f^{(k)}$ denotes the k^{th} Fréchet derivative of the function f with $f^{(0)} := f$. For the definition of Fréchet derivatives and properties of Fréchet differentiable functions in infinite dimensional Banach spaces we refer to [6]. Clearly $C_b^{(1)}(\mathbb{B})$ coincides with the class of Lipschitz functions, and $\|f\|_{(1)} = \|f\|_{\text{Lip}}$ for $f \in C_b^{(1)}(\mathbb{B})$. By $C_b^\infty(\mathbb{B})$ we denote the class of infinitely many times differentiable functions with bounded derivatives.

Define for $P, Q \in \mathcal{P}(\mathbb{B})$,

$$\zeta_p(P, Q) := \sup \left\{ |Pf - Qf| : f \in C_b^{(p)}(\mathbb{B}), \|f\|_{(p)} \leq 1 \right\}. \quad (2)$$

A natural question is then: for which Banach spaces \mathbb{B} , does the class $C_b^{(p)}(\mathbb{B}), p \geq 1$, determine the weak convergence of probability measures? Roughly speaking this is true in the case where the norm of \mathbb{B} is sufficiently smooth. To be more precise, we define first what we mean by smoothness of a norm.

Definition 1 Let $p \geq 1$. We say that a Banach space \mathbb{B} is p -smooth if its norm $\psi(x) := \|x\|_{\mathbb{B}}, x \in \mathbb{B}$, is $\lfloor p \rfloor$ -times continuously Fréchet differentiable on the set $\mathbb{B} \setminus \{0\}$, and

$$\sum_{i=1}^{\lfloor p \rfloor} \sup_{\|x\|=1} \|\psi^{(i)}(x)\| + \sup_{x \neq y, \|x\|=\|y\|=1} \frac{\|\psi^{(\lfloor p \rfloor)}(x) - \psi^{(\lfloor p \rfloor)}(y)\|}{\|x - y\|^{\{p\}}} < \infty. \quad (3)$$

Evidently every Banach space is 1-smooth. If \mathbb{B} is q -smooth for some $q > 1$, it is also p -smooth for $1 \leq p \leq q$. Examples of p -smooth spaces, where $p > 1$, are given below (see Examples 12 and 14).

Remark 2 Our definition of p -smoothness, tailored for the Lindeberg method, looks different of the p -smoothability as in e.g. Rosiński [22, p.159] where \mathbb{B} is said to be p -smoothable ($1 \leq p \leq 2$), if there exists an equivalent norm $|\cdot|$ on \mathbb{B} such that the modulus of smoothness

$$\rho_{|\cdot|}(t) := \sup \left\{ \frac{|x+ty| + |x-ty|}{2} - 1 : |x| = |y| = 1 \right\} = O(t^p), \quad \text{as } t \rightarrow 0.$$

By Lemma 19 p.246 in [14], this condition means that $|\cdot|$ is $C^{1,p-1}$ smooth in the sense of [14, Def.124, p.55]. So for $1 \leq p \leq 2$, p -smoothability and p smoothness in the sense of Definition 1 are similar up to the equivalence of norms. But the use of $\rho_{|\cdot|}(t)$ to define the p -smoothability induces the restriction $p \leq 2$ while we need $p > 2$ for the Lindeberg method.

Remark 3 It seems worth noticing here the two following facts about ψ .

- a) For any Banach space \mathbb{B} , there is no Fréchet derivative of ψ at 0. Indeed should ψ be Fréchet derivable at 0, the same should hold for its restriction to the one dimensional subspace $D = \{tu, t \in \mathbb{R}\}$, for some fixed $u \in \mathbb{B} \setminus \{0\}$. This in turn would imply the derivability at 0 (in the classical elementary sense) of the function $t \mapsto |t|$, which clearly fails.

b) If ψ is Fréchet differentiable on $B \setminus \{0\}$, then

$$\psi^{(1)}(x) = \psi^{(1)}\left(\frac{x}{\|x\|}\right), \quad x \in B \setminus \{0\}. \quad (4)$$

In particular, $\psi^{(1)}$ is bounded on $B \setminus \{0\}$ if and only if it is bounded on $\{x \in B, \|x\| = 1\}$. Another obvious consequence of (4) is that

$$\psi^{(1)}(cx) = \psi^{(1)}(x), \quad c \in \mathbb{R} \setminus \{0\}, x \in B \setminus \{0\}. \quad (5)$$

To check (4), we note that the Fréchet differentiability of ψ means that for each fixed $x \neq 0$,

$$\|x+h\| - \|x\| - \psi^{(1)}(x)(h) = o(\|h\|), \quad h \rightarrow 0.$$

Putting $y := x/\|x\|$, $u := h/\|x\|$ and recalling that $\psi^{(1)}(x)$ is a linear operator $B \rightarrow \mathbb{R}$, lead to

$$|\psi(y+u) - \psi(y) - \psi^{(1)}(x)(u)| = o(\|u\|), \quad u \rightarrow 0,$$

whence $\psi^{(1)}(x) = \psi^{(1)}(y)$.

The main result in this section is the following theorem.

Theorem 4 *Let $p \geq 1$. If the Banach space B is p -smooth, then for $(P, P_n, n \in \mathbb{N}) \subset \mathcal{P}(B)$, the following statements are equivalent*

- (i) $P_n \xrightarrow[n \rightarrow \infty]{w} P$;
- (ii) $P_n f \xrightarrow[n \rightarrow \infty]{} P f$ for any $f \in C_b^{(p)}(B)$;
- (iii) $\lim_{n \rightarrow \infty} \zeta_p(P_n, P) = 0$.

The proof of the theorem will be achieved by establishing the cycle of implications:

$$(i) \Rightarrow (iii) \Rightarrow (ii) \Rightarrow (i).$$

For $(i) \Rightarrow (iii)$ we note that for $p \geq 1$, the unit ball U_p of $C_b^{(p)}(B)$ is an equicontinuous family in $C_b(B)$, uniformly bounded by the constant 1. Then the convergence $P_n f \rightarrow P f$ is uniform on U_p by Theorem 3.1 in Ranga Rao [21].

$(iii) \Rightarrow (ii)$ is obvious for f in U_p and extends to any f in $C_b^{(p)}(B)$ by linearity of the integral.

The hard part is $(ii) \Rightarrow (i)$ which we detail now.

Proof of (ii) \Rightarrow (i) To this aim it is enough to prove that if (ii) holds, then for each finite intersection A of open balls, we have $P_n(A) \xrightarrow{n \rightarrow \infty} P(A)$, provided A is a P-continuity set (see, e.g., Billingsley [5], Corollary 2 to Th. 2.2, p.15). Recall, $A \in \mathcal{B}_B$ is P-continuity set if $P(\partial A) = 0$, where ∂A denotes the boundary of A , that is $\partial A = \bar{A} \setminus \overset{\circ}{A}$ where \bar{A} and $\overset{\circ}{A}$ are the closure and the interior of A respectively.

Let for $x \in B$ and $r > 0$, $B(x, r) = \{y \in B : \|y - x\| < r\}$. Set

$$A = \bigcap_{i=1}^m B(x_i, r_i), \quad x_1, \dots, x_m \in B, \quad r_1, \dots, r_m > 0.$$

It suffices to prove, that

$$P_n(A) \xrightarrow{n \rightarrow \infty} P(A) \quad \text{whenever} \quad P(\partial A) = 0. \quad (6)$$

To this aim, define for $0 < \varepsilon < r_0 := \min_{1 \leq i \leq m} r_i$,

$$A^\varepsilon := \bigcap_{i=1}^m B(x_i, r_i + \varepsilon), \quad A_\varepsilon := \bigcap_{i=1}^m B(x_i, r_i - \varepsilon).$$

It is easily seen that

$$\bigcup_{0 < \varepsilon < r_0} A_\varepsilon = A, \quad (7)$$

$$\bigcap_{0 < \varepsilon < r_0} A^\varepsilon = A' := \{x \in B : \|x - x_i\| \leq r_i, 1 \leq i \leq m\}.$$

As intersection of closed balls, A' is closed and since $A \subset A'$, the closure \bar{A} of A is included in A' . It is not difficult to find examples where this inclusion is strict when A is empty. Of course this special case may be discarded since with $A = \emptyset$, the convergence (6) is trivial. When A is non empty, one can check that $A' = \bar{A}$ as follows. Let x be an arbitray element in A' . There is at least one element y_0 in A . Then we define $y_1 := \frac{x+y_0}{2}$. For $i = 1, \dots, m$, $\|y_1 - x_i\| \leq \frac{1}{2}\|x - x_i\| + \frac{1}{2}\|y_0 - x_i\|$. As for $i = 1, \dots, m$, $\|x - x_i\| \leq r_i$ and $\|y_0 - x_i\| < r_i$, we see that $\|y_1 - x_i\| < r_i$ hence y_1 is in A . Iterating this argument we construct the sequence (y_n) in A such that $y_n = \frac{x+y_{n-1}}{2}$. Since $\|x - y_n\| \leq 2^{-n}\|x - y_0\|$, $n \geq 1$, x belongs to \bar{A} as limit of a sequence of points of A . Therefore, if $A \neq \emptyset$,

$$\bigcap_{0 < \varepsilon < r_0} A^\varepsilon = \bar{A}. \quad (8)$$

Next we construct for $A \neq \emptyset$ and $0 < \varepsilon < r_0$, the functions $f^\varepsilon, f_\varepsilon \in C_b^{(p)}(B)$ such that

$$\mathbf{1}_{A_\varepsilon}(x) \leq f_\varepsilon(x) \leq \mathbf{1}_A(x) \leq \mathbf{1}_{\bar{A}}(x) \leq f^\varepsilon(x) \leq \mathbf{1}_{A^\varepsilon}(x) \quad (9)$$

for all $x \in B$, where $\mathbf{1}_A$ is the indicator function of A . Assume for a moment that these functions are already constructed. From (7) and (8) we obtain by monotone sequential continuity of the probability measure P ,

$$\lim_{\varepsilon \downarrow 0} P(A_\varepsilon) = P(A), \quad \lim_{\varepsilon \downarrow 0} P(A^\varepsilon) = P(\bar{A}).$$

In view of (9) and recalling that A is a P -continuity set, this gives

$$\lim_{\varepsilon \downarrow 0} P f_\varepsilon = P(A) = P(\overline{A}) = \lim_{\varepsilon \downarrow 0} P(f^\varepsilon), \quad (10)$$

From (9) and the hypothesis (ii) applied with $f_\varepsilon, f^\varepsilon$,

$$\limsup_{n \rightarrow \infty} P_n(A) \leq \limsup_{n \rightarrow \infty} P_n f^\varepsilon = \lim_{n \rightarrow \infty} P_n f^\varepsilon = P f^\varepsilon \quad (11)$$

and

$$\liminf_{n \rightarrow \infty} P_n(A) \geq \liminf_{n \rightarrow \infty} P_n f_\varepsilon = \lim_{n \rightarrow \infty} P_n f_\varepsilon = P f_\varepsilon \quad (12)$$

for each $0 < \varepsilon < r_0$. Taking into account (10), from (11) and (12) we deduce (6). So, it remains to construct the functions $f^\varepsilon, f_\varepsilon$.

We begin with a lemma on the Fréchet derivatives of the norm of a p -smooth space which completes Remark 3. It quantifies the explosion of the successive derivatives of the norm near 0.

Lemma 5 *When B is p -smooth, its norm $\psi(x) = \|x\|$, satisfies for $i = 1, \dots, \lfloor p \rfloor$,*

$$\psi^{(i)}\left(\frac{x}{\|x\|}\right) = \|x\|^{i-1} \psi^{(i)}(x), \quad x \in B \setminus \{0\}. \quad (13)$$

Moreover there exists constants $c_1, \dots, c_{\lfloor p \rfloor}, c_p$ such that

$$\|\psi^{(i)}(x)\| \leq c_i \|x\|^{1-i}, \quad i = 1, \dots, \lfloor p \rfloor \quad (14)$$

and

$$\|\psi^{(\lfloor p \rfloor)}(x) - \psi^{(\lfloor p \rfloor)}(y)\| \leq c_p (\|x\|^{1-p} + \|y\|^{1-p}) \|x - y\|^{\{p\}}, \quad (15)$$

for all $x, y \in B, x, y \neq 0$.

Proof of lemma 5. To prove (13), we proceed by finite induction on $1 \leq i < \lfloor p \rfloor$. The initialization step is (4) already checked. Define for $x \in B \setminus \{0\}, y \in B$ and $1 \leq i < \lfloor p \rfloor$,

$$T_{i+1}(x, y) := \psi^{(i)}\left(\frac{x}{\|x\|} + y\right) \cdot y^{\otimes i} - \psi^{(i)}\left(\frac{x}{\|x\|}\right) \cdot y^{\otimes i} - \|x\|^i \psi^{(i+1)}(x) \cdot y^{\otimes(i+1)},$$

where for $y \in B, y^{\otimes i} = (y, \dots, y)$ denotes the element of B^i with all components equal to y and for a i -linear form L on $B^i, L \cdot w$ stands for $L(w), w = (w_1, \dots, w_i) \in B^i$. To complete the proof of (13), it suffices to prove that under the induction assumption (13) for some $i, |T_{i+1}(x, y)| = o(\|y\|^{i+1})$ for any fixed $x \in B \setminus \{0\}$ when $y \rightarrow 0$ in B . Indeed then both symmetric $(i+1)$ -linear forms $\psi^{(i+1)}\left(\frac{x}{\|x\|}\right)$ and $\|x\|^i \psi^{(i+1)}(x)$ are equal on the diagonal of B^{i+1} and this equality extends to the whole space B^{i+1} by symmetry.

By the induction assumption, restricting to $\|y\| < 1$ to avoid a possible vanishing of $\frac{x}{\|x\|} + y$,

$$\begin{aligned} \psi^{(i)}\left(\frac{x}{\|x\|} + y\right) &= \left\| \frac{x}{\|x\|} + y \right\|^{1-i} \psi^{(i)}\left(\frac{\frac{x}{\|x\|} + y}{\left\| \frac{x}{\|x\|} + y \right\|}\right) \\ &= \|x\|^{i-1} \|x + \|x\|y\|^{1-i} \psi^{(i)}\left(\frac{x + \|x\|y}{\|x + \|x\|y\|}\right) \\ &= \|x\|^{i-1} \psi^{(i)}(x + \|x\|y) \end{aligned}$$

and

$$\psi^{(i)}\left(\frac{x}{\|x\|}\right) = \|x\|^{i-1} \psi^{(i)}(x).$$

By the multilinearity of the Fréchet derivatives, it follows that

$$\begin{aligned} \|x\| T_{i+1}(x, y) &= \\ &= \psi^{(i)}(x + \|x\|y) \cdot (\|x\|y)^{\otimes i} - \psi^{(i)}(x) \cdot (\|x\|y)^{\otimes i} - \psi^{(i+1)}(x) \cdot (\|x\|y)^{\otimes(i+1)}. \end{aligned}$$

Putting $h := \|x\|y$ we deduce from the existence of $\psi^{(i+1)}$ that $\|x\| |T_{i+1}(x, y)| = o(\|h\|^{i+1})$ as $h \rightarrow 0$, whence recalling that x is fixed, $|T_{i+1}(x, y)| = o(\|y\|^{i+1})$ when y tends to 0, as expected.

Clearly (14) follows from (13) with $c_i := \sup\{\|\psi^{(i)}(y)\| : \|y\| = 1\}$, recalling that the definition of p -smoothness includes the boundedness of each $\psi^{(i)}$, $1 \leq i \leq \lfloor p \rfloor$, on $\{y \in B : \|y\| = 1\}$, see (3).

Now we prove (15). Denoting $t = \|x\|^{-1}$, $s = \|y\|^{-1}$ for $x \neq 0, y \neq 0$, and using (13), we start from

$$\begin{aligned} \|\psi^{(\lfloor p \rfloor)}(x) - \psi^{(\lfloor p \rfloor)}(y)\| &= \|\psi^{(\lfloor p \rfloor)}(tx)t^{\lfloor p \rfloor - 1} - \psi^{(\lfloor p \rfloor)}(sy)s^{\lfloor p \rfloor - 1}\| \\ &\leq \|\psi^{(\lfloor p \rfloor)}(tx) - \psi^{(\lfloor p \rfloor)}(sy)\| t^{\lfloor p \rfloor - 1} + \\ &\quad \|\psi^{(\lfloor p \rfloor)}(sy)\| |t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}|. \end{aligned} \quad (16)$$

The first term in (16) is bounded by

$$\|\psi^{(\lfloor p \rfloor)}(tx) - \psi^{(\lfloor p \rfloor)}(sy)\| t^{\lfloor p \rfloor - 1} \leq c'_p \|tx - sy\|^{\{p\}} t^{\lfloor p \rfloor - 1},$$

where

$$c'_p := \sup_{\substack{\|w\|=\|z\|=1 \\ w \neq z}} \frac{\|\psi^{(\lfloor p \rfloor)}(w) - \psi^{(\lfloor p \rfloor)}(z)\|}{\|w - z\|^{\{p\}}},$$

is finite by the p -smoothness assumption (3). Now

$$\|tx - sy\|^{\{p\}} = \| \|y\|x - \|x\|y \|^{\{p\}} (\|x\| \|y\|)^{-\{p\}} = \left\| x - \frac{\|x\|}{\|y\|} y \right\|^{\{p\}} \|x\|^{-\{p\}}$$

and writing $x - \frac{\|x\|}{\|y\|}y = x - y + \frac{\|y\| - \|x\|}{\|y\|}y$, the triangle inequality gives

$$\left\| x - \frac{\|x\|}{\|y\|}y \right\| \leq \|x - y\| + \left\| \frac{\|y\| - \|x\|}{\|y\|}y \right\| \leq 2\|x - y\|,$$

whence, noticing that $2^{\{p\}} \leq 2$,

$$\begin{aligned} \|\Psi^{(\lfloor p \rfloor)}(tx) - \Psi^{(\lfloor p \rfloor)}(sy)\| t^{\lfloor p \rfloor - 1} &\leq 2c'_p \|x - y\|^{\{p\}} \|x\|^{-\{p\} - \lfloor p \rfloor + 1} \\ &= 2c'_p \|x - y\|^{\{p\}} \|x\|^{1-p}. \end{aligned} \quad (17)$$

To estimate the second term in (16), choose $a = \max\{\|x\|, \|y\|\}$ and consider first the case where $\|x - y\| \geq a$. In this case

$$|t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}| \leq [t^{\lfloor p \rfloor - 1} + s^{\lfloor p \rfloor - 1}] a^{-\{p\}} \|x - y\|^{\{p\}} \leq [\|x\|^{1-p} + \|y\|^{1-p}] \|x - y\|^{\{p\}}.$$

If $\|x - y\| \leq a$, we claim that

$$|t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}| \leq 2p(\|x\|^{1-p} + \|y\|^{1-p}) \|x - y\|^{\{p\}}. \quad (18)$$

Let us check (18). Put for simplification $m := \lfloor p \rfloor - 1$. Applying the elementary bound

$$|u^m - v^m| = |u - v| |u^{m-1} + u^{m-2}v + \dots + v^{m-1}| \leq |u - v| m (\max(|u|, |v|))^{m-1}, \quad u, v \in \mathbb{R},$$

which is optimal when v tends to u , gives

$$|t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}| = \frac{|\|x\|^m - \|y\|^m|}{\|x\|^m \|y\|^m} \leq \frac{ma^{m-1}}{\|x\|^m \|y\|^m} \|x - y\|^{\{p\}}.$$

As $m < p$ and $\|x - y\|^{1-\{p\}} \leq (2a)^{1-\{p\}} \leq 2a^{1-\{p\}}$, this leads to

$$|t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}| \leq 2p \frac{a^{\lfloor p \rfloor - 1 - \{p\}}}{\|x\|^{\lfloor p \rfloor - 1} \|y\|^{\lfloor p \rfloor - 1}} \|x - y\|^{\{p\}}.$$

If $a = \|x\|$, this provides

$$\frac{|t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}|}{\|x - y\|^{\{p\}}} \leq 2p \frac{a^{-\{p\}}}{\|y\|^{\lfloor p \rfloor - 1}} = 2p \|y\|^{1-p} \left(\frac{\|y\|}{a} \right)^{\{p\}} \leq 2p \|y\|^{1-p}.$$

Obviously the same estimate holds replacing $\|y\|$ by $\|x\|$ when $a = \|y\|$, and adding both estimates to have a common bound gives (18).

Gathering the estimates, we obtain for the second term in (16),

$$\Psi^{(\lfloor p \rfloor)}(sy) |t^{\lfloor p \rfloor - 1} - s^{\lfloor p \rfloor - 1}| \leq 2pc_{\lfloor p \rfloor} (\|x\|^{1-p} + \|y\|^{1-p}) \|x - y\|^{\{p\}}.$$

Accounting (17), this complete the proof of (15) with $c_p := 2 \max(c'_p, pc_{\lfloor p \rfloor})$.

Let us go back to the construction of f^ε , f_ε . Lemma 5 quantifies in some way the non membership of the norm of B in the space $C_b^{(p)}(B)$. To remedy this drawback, the idea is to modify ψ inside the ball $B(0, \varepsilon)$ by flatening to zero the peak of ψ in the ball $B(0, \varepsilon/2)$ and use a connection through $B(0, \varepsilon) \setminus B(0, \varepsilon/2)$ smooth enough to

obtain an approximation of ψ by a function g_ε in $C_b^{(p)}(\mathbb{B})$. To this aim, let us choose a function $u \in C_b^{(\infty)}([0, \infty))$ such that $0 \leq u \leq 1$, $u = 0$ on $[0, 1/2]$, $u = 1$ on $[1, \infty)$. Set

$$g_\varepsilon(x) = \psi(x)u(\varepsilon^{-1}\psi(x)), \quad x \in \mathbb{B}. \quad (19)$$

More explicitly,

$$g_\varepsilon(x) = \begin{cases} 0 & \text{if } 0 \leq \|x\| \leq \frac{\varepsilon}{2}, \\ \|x\| u\left(\frac{\|x\|}{\varepsilon}\right) & \text{if } \frac{\varepsilon}{2} < \|x\| < \varepsilon, \\ \|x\| & \text{if } \|x\| \geq \varepsilon. \end{cases}$$

The function g_ε is uniformly approximating the norm. Indeed

$$\sup_{x \in \mathbb{B}} |\psi(x) - g_\varepsilon(x)| = \sup_x |\psi(x)| |1 - u(\varepsilon^{-1}\psi(x))| \leq \varepsilon, \quad (20)$$

since $u(\varepsilon^{-1}\psi(x)) = 1$, if $\psi(x) \geq \varepsilon$.

The next lemma establishes the membership of g_ε in $C_b^{(p)}(\mathbb{B})$ and provides some control of its norm $\|g_\varepsilon\|_p$, defined by (1), in terms of the parameter ε .

Lemma 6 *There is a constant $C > 0$ such that*

$$\sup_{x \in \mathbb{B}} \|g_\varepsilon^{(i)}(x)\| \leq C\varepsilon^{1-i}, \quad i = 1, \dots, [p] \quad (21)$$

and

$$\sup_{x, y \in \mathbb{B}} \|g_\varepsilon^{([p])}(x) - g_\varepsilon^{([p])}(y)\| \leq C\varepsilon^{1-p} \|x - y\|^{[p]}. \quad (22)$$

Proof of lemma 6. Denoting

$$t_x := \varepsilon^{-1}\psi(x), \quad v(t) := tu(t),$$

we rewrite

$$g_\varepsilon(x) = \varepsilon v(t_x).$$

An immediate induction provides the following formula for the successive derivatives of v

$$v^{(m)}(t) = tu^{(m)}(t) + mu^{(m-1)}(t), \quad t \geq 0, m \geq 1.$$

We note also that $u^{(m)}(1/2) = u^{(m)}(1) = 0$ since u is C^∞ and null at the right of 1 and at the left of $1/2$. The values of $v^{(m)}$ on $[0, \frac{1}{2}] \cup [1, \infty)$ are then

$$v'(t) = \begin{cases} 0 & \text{if } t \leq 1/2, \\ 1 & \text{if } t \geq 1, \end{cases} \quad v^{(m)}(t) = \begin{cases} 0 & \text{if } t \leq 1/2, \\ 0 & \text{if } t \geq 1, \end{cases} \quad m > 1. \quad (23)$$

Together with the infinite derivability of v , this implies that for any integer $m \geq 1$,

$$C_m := \sup_{t \geq 0} |v^{(m)}(t)| < \infty, \quad D_m := \sup_{\substack{s, t \geq 0 \\ t \neq s}} \frac{|v^{(m)}(t) - v^{(m)}(s)|}{|t - s|^{[p]}} < \infty. \quad (24)$$

Using differentiation of composite functions on Banach spaces (see, e.g., Fraenkel [11]), we find

$$\frac{d^m}{dx^m} v(t_x)(h_1, \dots, h_m) = \sum_{j=1}^m \sum_{\substack{\beta \in \langle j \rangle_+ \\ |\beta|=m}} \sum_{\sigma} \frac{v^{(j)}(t_x)}{j! \beta!} I_{\beta, \sigma}(x)(h_1, \dots, h_m) \quad (25)$$

with

$$I_{\beta, \sigma}(x)(h_1, \dots, h_m) := t_x^{(\beta_1)}(h_{\sigma_1}, \dots, h_{\sigma_{\beta_1}}) \cdots t_x^{(\beta_j)}(h_{\sigma_{m-\beta_j+1}}, \dots, h_{\sigma_m}),$$

where $\beta \in \langle j \rangle_+$ means that $\beta = (\beta_1, \dots, \beta_j)$ with integers $\beta_i \geq 1, i = 1, \dots, j$, $|\beta| = \beta_1 + \dots + \beta_j$ and $\beta! = \beta_1! \cdots \beta_j!$. In \sum_{σ} , the summation runs over the $m!$ permutations σ of $\{1, \dots, m\}$.

To prove (21), we need to bound $\|I_{\beta, \sigma}(x)\|$ only for $\|x\| > \frac{\varepsilon}{2}$ since for $m \geq 1$, $v^{(m)}(t_x) = 0$ for $t_x \in [0, 1/2]$, that is for $\|x\| \leq \frac{\varepsilon}{2}$. So assuming $\|x\| > \frac{\varepsilon}{2}$ and accounting (14), we obtain

$$\begin{aligned} \|I_{\beta, \sigma}(x)\| &\leq \|t_x^{(\beta_1)}\| \cdots \|t_x^{(\beta_j)}\| \leq \left(\frac{1}{\varepsilon} c_{\beta_1} \|x\|^{1-\beta_1}\right) \cdots \left(\frac{1}{\varepsilon} c_{\beta_j} \|x\|^{1-\beta_j}\right) \\ &= c_{\beta_1} \cdots c_{\beta_j} \varepsilon^{-j} \|x\|^{j-m} \\ &\leq 2^m c_{\beta_1} \cdots c_{\beta_j} \varepsilon^{-m}. \end{aligned} \quad (26)$$

Consequently $\left\| \frac{d^m}{dx^m} v(t_x) \right\| \leq K_m \varepsilon^{-m}$ and as $g_{\varepsilon}(x) = \varepsilon v(t_x)$, (21) is checked.

To prove (22), we have to find a bound of the form $c \varepsilon^{-p} \|x - y\|^{[p]}$ for each

$$\Delta_{\beta, \sigma}(x, y) := v^{(j)}(t_x) I_{\beta, \sigma}(x) - v^{(j)}(t_y) I_{\beta, \sigma}(y), \quad \beta \in \langle j \rangle_+, \quad 1 \leq j \leq [p], \quad (27)$$

because of the decomposition

$$\frac{d^{[p]}}{dx^{[p]}} v(t_x) - \frac{d^{[p]}}{dy^{[p]}} v(t_y) = \sum_{j=1}^{[p]} \sum_{\substack{\beta \in \langle j \rangle_+ \\ |\beta|=[p]}} \sum_{\sigma} \frac{1}{j! \beta!} \Delta_{\beta, \sigma}(x, y). \quad (28)$$

In view of (23), the discussion is naturally ordered according to the various configurations of $\|x\|$, $\|y\|$ and the open interval $(\frac{\varepsilon}{2}, \varepsilon)$.

Case 1: $\|x\|$ and $\|y\|$ are both outside $(\frac{\varepsilon}{2}, \varepsilon)$. As t_x, t_y are both outside $(\frac{1}{2}, 1)$, it is clear from (23) that for $j \geq 2$, $v^{(j)}(t_x) = v^{(j)}(t_y) = 0$ whence $\Delta_{\beta, \sigma}(x, y) = 0$. For the same reason, $\Delta_{\beta, \sigma}(x, y) = 0$ when $j = 1$ and $\|x\|, \|y\| \leq \frac{\varepsilon}{2}$. If $j = 1$ and $\|x\|, \|y\| \geq \varepsilon$,

$$\begin{aligned} \Delta_{\beta, \sigma}(x, y)(h_1, \dots, h_{[p]}) &= (I_{\beta, \sigma}(x) - I_{\beta, \sigma}(y))(h_1, \dots, h_{[p]}) \\ &= (t_x^{([p])} - t_y^{([p])})(h_{\sigma_1}, \dots, h_{\sigma_{[p]}}), \end{aligned}$$

whence recalling (15),

$$\begin{aligned} \|\Delta_{\beta,\sigma}(x,y)\| &\leq \left\| t_x^{(lp)} - t_y^{(lp)} \right\| = \varepsilon^{-1} \left\| \psi^{(lp)}(x) - \psi^{(lp)}(y) \right\| \\ &\leq c_p \varepsilon^{-1} (\|x\|^{1-p} + \|y\|^{1-p}) \|x-y\|^{\{p\}} \\ &\leq 2c_p \varepsilon^{-p} \|x-y\|^{\{p\}}. \end{aligned} \quad (29)$$

Case 2: only one of $\|x\|$, $\|y\|$ is inside $(\frac{\varepsilon}{2}, \varepsilon)$. By symmetry, it suffices to treat the configurations where $\|x\|$ is inside $(\frac{\varepsilon}{2}, \varepsilon)$ and $\|y\|$ outside.

Case 2.a: $0 < \|y\| \leq \frac{\varepsilon}{2} < \|x\| < \varepsilon$. In this configuration, for $j \geq 1$ and $\beta \in \langle j \rangle_+$,

$$\|\Delta_{\beta,\sigma}(x,y)\| = |v^{(j)}(t_x)| \|I_{\beta,\sigma}(x)\| = |v^{(j)}(t_x) - v^{(j)}(t_y)| \|I_{\beta,\sigma}(x)\|. \quad (30)$$

From (24),

$$|v^{(j)}(t_x) - v^{(j)}(t_y)| \leq D_j |t_x - t_y|^{\{p\}} = D_j \varepsilon^{-\{p\}} \|\|x\| - \|y\|\|^{\{p\}} \leq D_j \varepsilon^{-\{p\}} \|x-y\|^{\{p\}}$$

which together with (26), leads to

$$\|\Delta_{\beta,\sigma}(x,y)\| \leq (2^p c_{\beta_1} \cdots c_{\beta_j} D_j) \varepsilon^{-p} \|x-y\|^{\{p\}}, \quad \beta \in \langle j \rangle_+, j \geq 1.$$

Case 2.b: $\frac{\varepsilon}{2} < \|x\| < \varepsilon \leq \|y\|$. For $j \geq 2$, (30) still holds and exactly the same argument as above gives

$$\|\Delta_{\beta,\sigma}(x,y)\| \leq (2^p c_{\beta_1} \cdots c_{\beta_j} D_j) \varepsilon^{-p} \|x-y\|^{\{p\}}, \quad \beta \in \langle j \rangle_+, j \geq 2.$$

In the special case where $j = 1$, as $v'(t_y) = 1$,

$$\Delta_{\beta,\sigma}(x,y) = v'(t_x) I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y),$$

whence

$$\|\Delta_{\beta,\sigma}(x,y)\| \leq |v'(t_x) - v'(t_y)| \|I_{\beta,\sigma}(x)\| + \|I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y)\|.$$

Bounding the first term in the right-hand side exactly as in case 2.a and referring to (29) in case 1 for the second one, we obtain

$$\|\Delta_{\beta,\sigma}(x,y)\| \leq (2^p c_{\beta_1} \cdots c_{\beta_j} D_j + 2c_p) \varepsilon^{-p} \|x-y\|^{\{p\}}, \quad \beta \in \langle 1 \rangle_+.$$

Case 3: $\frac{\varepsilon}{2} < \|x\|, \|y\| < \varepsilon$. We start from

$$\|\Delta_{\beta,\sigma}(x,y)\| \leq |v^{(j)}(t_x) - v^{(j)}(t_y)| \cdot \|I_{\beta,\sigma}(x)\| + |v^{(j)}(t_y)| \cdot \|I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y)\|. \quad (31)$$

The first term in the right-hand side of (31) is bounded exactly as in case 2.a:

$$|v^{(j)}(t_x) - v^{(j)}(t_y)| \cdot \|I_{\beta,\sigma}(x)\| \leq (2^p c_{\beta_1} \cdots c_{\beta_j} D_j) \varepsilon^{-p} \|x-y\|^{\{p\}}, \quad \beta \in \langle j \rangle_+, j \geq 1.$$

For the second term, let us treat first the special case where $j = 1$. Arguing as in case 1, we just have to replace ε by $\frac{\varepsilon}{2}$ in (29), so accounting (24),

$$|v^{(1)}(t_y)| \cdot \|I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y)\| \leq 2^{p+1} C_1 c_p \varepsilon^{-p} \|x-y\|^{\{p\}}, \quad \beta \in \langle 1 \rangle_+.$$

Assume now that $\beta \in \langle j \rangle_+$ with $2 \leq j \leq \lfloor p \rfloor$. As $|\beta| = \lfloor p \rfloor$, for each component β_i of β , $\beta_i < \lfloor p \rfloor$. Using telescopic summation where at each step one factor $t_x^{(\beta_i)}$ is replaced by $t_y^{(\beta_i)}$ gives

$$I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y) = \sum_{i=1}^j \left(\prod_{1 \leq k < i} t_x^{(\beta_k)} \right) \left(t_x^{(\beta_i)} - t_y^{(\beta_i)} \right) \left(\prod_{i < k \leq j} t_y^{(\beta_k)} \right),$$

with the usual convention that a product indexed by \emptyset equals 1. Therefore

$$\|I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y)\| \leq \sum_{i=1}^j \Pi_{\beta,i} \|t_x^{(\beta_i)} - t_y^{(\beta_i)}\|, \quad (32)$$

where

$$\Pi_{\beta,i} := \left(\prod_{1 \leq k < i} \|t_x^{(\beta_k)}\| \right) \left(\prod_{i < k \leq j} \|t_y^{(\beta_k)}\| \right)$$

It is easy to bound $\Pi_{\beta,i}$ since by (14), $\|t_x^{(\beta_k)}\| = \varepsilon^{-1} \|\psi^{(\beta_k)}(x)\| \leq \varepsilon^{-1} c_{\beta_k} \|x\|^{1-\beta_k}$ and as $\beta_k \geq 1$ and $\|x\| > \frac{\varepsilon}{2}$, $\|x\|^{1-\beta_k} \leq 2^{\beta_k-1} \varepsilon^{1-\beta_k}$. Obviously the same holds for $\|t_y^{(\beta_k)}\|$ and all this gives

$$\Pi_{\beta,i} \leq 2^{\lfloor p \rfloor - \beta_i} \varepsilon^{\beta_i - \lfloor p \rfloor} \prod_{\substack{1 \leq k \leq j \\ k \neq i}} c_{\beta_k}. \quad (33)$$

Recalling that $\beta_i < \lfloor p \rfloor$, $1 \leq i \leq j$, it remains to estimate for any $1 \leq m < \lfloor p \rfloor$,

$$\delta_m(x, y) := \|t_x^{(m)} - t_y^{(m)}\| = \varepsilon^{-1} \|\psi^{(m)}(x) - \psi^{(m)}(y)\|.$$

In view of (14) it seems relevant to apply the mean-value theorem for derivatives to the function $\psi^{(m)} : \mathbb{B} \setminus \{0\} \rightarrow \mathcal{L}_m(\mathbb{B})$. But then, care must be taken of the inclusion of the segment $[x, y]$ in the open set $\mathbb{B} \setminus \{0\}$. If 0 belongs to $[x, y]$, then there exists $s \in [0, 1]$ such that $(1-s)x + sy = 0$, that is $x = s(x-y)$ whence $\|x\| = s\|x-y\|$. If $\|x-y\| \leq \frac{\varepsilon}{2}$, this equality is impossible since $s \in [0, 1]$ and $\|x\| > \frac{\varepsilon}{2}$. Accordingly we separate the cases $\|x-y\| \leq \frac{\varepsilon}{2}$ and $\|x-y\| > \frac{\varepsilon}{2}$.

If $\|x-y\| \leq \frac{\varepsilon}{2}$, by the mean-value theorem, (14) and the convexity of the function $t \mapsto t^{-m}$,

$$\begin{aligned} \delta_m(x, y) &\leq \varepsilon^{-1} \sup_{s \in [0, 1]} \left\| \psi^{(m+1)}((1-s)x + sy) \right\| \|x-y\| \\ &\leq \varepsilon^{-1} c_{m+1} \sup_{s \in [0, 1]} \|(1-s)x + sy\|^{-m} \left(\frac{\varepsilon}{2}\right)^{1-\{p\}} \|x-y\|^{\{p\}} \\ &\leq 2^{\{p\}-1} c_{m+1} \sup_{s \in [0, 1]} \left((1-s)\|x\|^{-m} + s\|y\|^{-m} \right) \varepsilon^{-\{p\}} \|x-y\|^{\{p\}} \\ &\leq 2^{m+\{p\}-1} c_{m+1} \varepsilon^{-m-\{p\}} \|x-y\|^{\{p\}}. \end{aligned}$$

If $\|x - y\| > \frac{\varepsilon}{2}$ it is enough to use (14) as follows.

$$\begin{aligned} \left\| \Psi^{(m)}(x) - \Psi^{(m)}(y) \right\| &\leq \left\| \Psi^{(m)}(x) \right\| + \left\| \Psi^{(m)}(y) \right\| \\ &\leq c_m (\|x\|^{1-m} + \|y\|^{1-m}) \leq 2^m c_m \varepsilon^{1-m}. \end{aligned}$$

Noticing that here $\varepsilon < 2\|x - y\|$, this gives

$$\delta_m(x, y) \leq 2^{m+\{p\}} c_m \varepsilon^{-m-\{p\}} \|x - y\|^{\{p\}}.$$

So we can retain from both cases the common bound:

$$\delta_m(x, y) \leq 2^{m+\{p\}} \max(c_m, c_{m+1}) \varepsilon^{-m-\{p\}} \|x - y\|^{\{p\}},$$

for $\frac{\varepsilon}{2} < \|x\|, \|y\| < \varepsilon$, $1 \leq m < \lfloor p \rfloor$. This together with (33) enables us to bound the i^{th} term in (32) as

$$\begin{aligned} \Pi_{\beta, i} \left\| t_x^{(\beta_i)} - t_y^{(\beta_i)} \right\| &\leq 2^{\lfloor p \rfloor - \beta_i} \varepsilon^{\beta_i - \lfloor p \rfloor} \prod_{\substack{1 \leq k \leq j \\ k \neq i}} c_{\beta_k} 2^{\beta_i + \{p\}} \max(c_{\beta_i}, c_{\beta_{i+1}}) \varepsilon^{-\beta_i - \{p\}} \|x - y\|^{\{p\}} \\ &= 2^p \prod_{\substack{1 \leq k \leq j \\ k \neq i}} c_{\beta_k} \max(c_{\beta_i}, c_{\beta_{i+1}}) \varepsilon^{-p} \|x - y\|^{\{p\}}. \end{aligned}$$

Finally, accounting (24),

$$|v^{(j)}(t_y)| \cdot \|I_{\beta, \sigma}(x) - I_{\beta, \sigma}(y)\| \leq c \varepsilon^{-p} \|x - y\|^{\{p\}}, \quad \beta \in \langle j \rangle_+, \quad 2 \leq j \leq \lfloor p \rfloor.$$

This completes the proof of (22) and Lemma 6 \square

Next for each $\varepsilon > 0$ we construct a $\lfloor p \rfloor$ -times Fréchet differentiable function $\phi_\varepsilon : \mathbb{B} \rightarrow \mathbb{R}$ such that

$$\phi_\varepsilon(x) = \begin{cases} 1 & \text{if } \|x\| \leq 1 \\ 0 & \text{if } \|x\| > 1 + \varepsilon \end{cases} \quad (34)$$

and

$$\sum_{j=0}^{\lfloor p \rfloor} \sup_{\varepsilon > 0} \varepsilon^j \sup_{x \in \mathbb{B}} \|\phi_\varepsilon^{(j)}(x)\| + \sup_{\varepsilon > 0} \varepsilon^p \sup_{x \neq y} \frac{\|\phi_\varepsilon^{(\lfloor p \rfloor)}(x) - \phi_\varepsilon^{(\lfloor p \rfloor)}(y)\|}{\|x - y\|^{\{p\}}} < \infty. \quad (35)$$

To this aim, let the function $q \in C_b^{(\infty)}(\mathbb{R})$ be such that $0 \leq q \leq 1$, $q(t) = 1$, if $t < 1/8$, $q(t) = 0$, if $t > 7/8$. Set

$$\phi_\varepsilon(x) = q(\varepsilon^{-1}(g_{\varepsilon/8}(x) - 1)), \quad x \in \mathbb{B}.$$

If $\|x\| \leq 1$, then $(\|x\| - 1)\varepsilon^{-1} < 0$ and

$$\frac{g_{\varepsilon/8}(x) - 1}{\varepsilon} = \frac{g_{\varepsilon/8}(x) - 1}{\varepsilon} - \frac{\|x\| - 1}{\varepsilon} + \frac{\|x\| - 1}{\varepsilon} \leq \frac{g_{\varepsilon/8}(x) - \|x\|}{\varepsilon} \leq 1/8$$

therefore $\phi_\varepsilon(x) = 1$. If $\|x\| > 1 + \varepsilon$, then

$$\begin{aligned} \frac{g_{\varepsilon/8}(x) - 1}{\varepsilon} &= \frac{g_{\varepsilon/8}(x) - 1}{\varepsilon} - \frac{\|x\| - 1}{\varepsilon} + \frac{\|x\| - 1}{\varepsilon} > 1 + \frac{g_{\varepsilon/8}(x) - 1}{\varepsilon} - \frac{\|x\| - 1}{\varepsilon} \\ &\geq 1 - \left| \frac{g_{\varepsilon/8}(x) - 1}{\varepsilon} - \frac{\|x\| - 1}{\varepsilon} \right| \geq 1 - 1/8 = 7/8, \end{aligned}$$

therefore $\phi(x) = 0$ and (34) is confirmed. There remains to evaluate the derivatives of the function ϕ_ε . This can be done in much the same way as we proved (21) and (22). Finally we use $\phi_\varepsilon, \varepsilon > 0$ to define the required functions

$$f^\varepsilon(x) = \prod_{i=1}^m \phi_{\varepsilon r_i^{-1}}(r_i^{-1}(x - x_i)), \quad f_\varepsilon(x) = \prod_{i=1}^m \phi_{\varepsilon(r_i - \varepsilon)^{-1}}((r_i - \varepsilon)^{-1}(x - x_i)).$$

It is straightforward to check that $\|f^\varepsilon\|_p < \infty$ and $\|f_\varepsilon\|_p < \infty$. This completes the proof of (ii) \Rightarrow (i) and Theorem 4 \square

Theorem 7 *If the Banach space B is ∞ -smooth, then for $(P, P_n, n \in \mathbb{N}) \subset \mathcal{P}(B)$, the following statements are equivalent*

- (i) $P_n \xrightarrow[n \rightarrow \infty]{w} P$;
- (ii) $P_n f \xrightarrow[n \rightarrow \infty]{} P f$ for any $f \in C_b^{(\infty)}(B)$.

Proof. The proof of Theorem 4 is adapted for the case of ∞ -smooth Banach space as well. One needs to follow its lines having in mind that p -differentiability of functions involved are ∞ -differentiable (or p -differentiable for any $p > 1$). So the finally constructed functions f^ε and f_ε belongs to $C_b^{(\infty)}(B)$. \square

Remark 8 From the proof of Theorem 4 we see that the differentiability of the norm could be substituted by its smooth approximation in the sense that for each $\varepsilon > 0$, there exists a $\lfloor p \rfloor$ -times Fréchet differentiable function $\psi_\varepsilon : B \rightarrow \mathbb{R}$ such that

- (a) for any $\varepsilon > 0$,

$$\sup_{x \in B} |\psi_\varepsilon(x) - \psi(x)| \leq \varepsilon;$$

- (b) with some constant $C > 0$,

$$\sup_{x \in B} \|\psi_\varepsilon^{(i)}(x)\| \leq C \varepsilon^{1-i}, \quad i = 1, \dots, \lfloor p \rfloor.$$

- (c) with some constant $C > 0$,

$$\sup_{x \neq y, x, y \in B} \frac{\|\psi_\varepsilon^{(\lfloor p \rfloor)}(x) - \psi_\varepsilon^{(\lfloor p \rfloor)}(y)\|}{\|x - y\|^{(\lfloor p \rfloor)}} \leq C \varepsilon^{1-p}.$$

Remark 9 Since $C_b^{(p)}(B) \subset C_b^{(p')}(B)$, if $p > p'$ it holds

$$\zeta_p(P, Q) \leq \zeta_{p'}(P, Q), \quad P, Q \in \mathcal{P}(B).$$

Remark 10 For B -valued random variables X, Y we set

$$\zeta_p(X, Y) := \zeta_p(P_X, P_Y),$$

where P_X denotes the distribution of X . Hence, if B is p -smooth then in order to check convergence in distribution of a sequence $(X_n, n \in \mathbb{N})$ of B -valued random variables to a B -valued random variable X it is enough to prove $\zeta_p(X_n, X) \xrightarrow{n \rightarrow \infty} 0$. The use of ζ_p in proving convergence in distribution of random variables is attractive due to the following simple but powerful properties of ζ_p :

(a) for each $c \in \mathbb{R}$,

$$\zeta_p(cX, cY) \leq \max\{1, |c|^p\} \zeta_p(X, Y);$$

(b) if the B -valued random element Z is independent of (X, Y) , then

$$\zeta_p(X + Z, Y + Z) \leq \zeta_p(X, Y);$$

(c) for independent B -valued random elements $X_1, \dots, X_n; Y_1, \dots, Y_n$,

$$\zeta_p\left(\sum_{k=1}^n X_k, \sum_{k=1}^n Y_k\right) \leq \sum_{k=1}^n \zeta_p(X_k, Y_k). \quad (36)$$

These properties of ζ_p were discovered by Zolotarev [26], but actually are easy to prove. The statement (a) follows directly from the definition of ζ_p . To prove (b) one needs to use Fubini theorem and invariance by shifts of the function space $C_b^{(p)}(B)$ that is under the transformations $T_x : C_b^{(p)}(B) \rightarrow C_b^{(p)}(B)$, $f \mapsto T_x f := f(x + \cdot)$, $x \in B$. Finally (c) follows from (b) and the triangle inequality.

3 Some remarks on smooth Banach spaces

Various aspects of differentiability of Banach space norm are discussed in Sundaresan [23].

3.1 Smoothness and type 2

Recall a Banach space B is said to be of type 2 if there is a constant $K > 0$ such that for any finite set of elements x_1, \dots, x_n in B and Rademacher sequence $\varepsilon_1, \dots, \varepsilon_n$ (the ε_i being independent and such that $P(\varepsilon_i = -1) = P(\varepsilon_i = 1) = 1/2$),

$$\left(\mathbb{E} \|\varepsilon_1 x_1 + \dots + \varepsilon_n x_n\|^2\right)^{1/2} \leq K \left(\|x_1\|^2 + \dots + \|x_n\|^2\right)^{1/2}. \quad (37)$$

By the Khintchine-Kahane inequality giving the equivalence of moments of Rademacher sums $\sum_i \varepsilon_i x_i$ see e.g. [17, Th.4.7], the second moment the left-hand side of (37) may be replaced by the first one, leading to the equivalent definition of type 2 by the inequality

$$\mathbb{E} \|\varepsilon_1 x_1 + \dots + \varepsilon_n x_n\| \leq K' \left(\|x_1\|^2 + \dots + \|x_n\|^2\right)^{1/2}. \quad (38)$$

Moreover, by e.g. [17, Prop.9.11], if the separable Banach space B is of type 2, there is a constant $K > 0$ depending only on B such that for any finite set of mean zero independent B -valued random elements X_1, \dots, X_n ,

$$\mathbb{E} \|X_1 + \dots + X_n\|^2 \leq K^2 (\mathbb{E} \|X_1\|^2 + \dots + \mathbb{E} \|X_n\|^2). \quad (39)$$

This obviously implies that

$$\mathbb{E} \|X_1 + \dots + X_n\| \leq K (\mathbb{E} \|X_1\|^2 + \dots + \mathbb{E} \|X_n\|^2)^{1/2}. \quad (40)$$

Conversely, if in a separable Banach space B , any finite set of mean zero independent B -valued random elements X_1, \dots, X_n , satisfies (40), then choosing $X_i = \varepsilon_i x_i$ shows that B satisfies (38) hence B is of type 2.

Proposition 11 *If the Banach space B is 2-smooth then B is of type 2.*

Proof. We just have to prove that (40) is satisfied in B . We use the functions g_ε , $\varepsilon > 0$, defined in (19) which are in $C_b^{(2)}(B)$ by Lemma 6.

For any fixed $a, b \in B$, the map $f : [0, 1] \rightarrow \mathbb{R}$, $t \mapsto f(t) := g_\varepsilon(a + tb)$ has clearly a continuous first derivative $f'(t) = g_\varepsilon^{(1)}(a + tb) \cdot b$, so $f(1) - f(0) = \int_0^1 f'(t) dt$, that is :

$$g_\varepsilon(a + b) - g_\varepsilon(a) = \int_0^1 g_\varepsilon^{(1)}(a + tb) \cdot b dt, \quad a, b \in B. \quad (41)$$

Denoting $S_0 = 0$, $S_j = X_1 + \dots + X_j$, $j = 1, \dots, n$, we have by (20),

$$\mathbb{E} \|S_n\| \leq \varepsilon + \mathbb{E} g_\varepsilon(S_n). \quad (42)$$

Recalling that $g_\varepsilon(0) = 0$ and applying (41) gives

$$\begin{aligned} \mathbb{E} g_\varepsilon(S_n) &= \mathbb{E} (g_\varepsilon(S_n) - g_\varepsilon(0)) = \sum_{j=1}^n \mathbb{E} (g_\varepsilon(S_j) - g_\varepsilon(S_{j-1})) \\ &= \sum_{j=1}^n \int_0^1 \mathbb{E} (g_\varepsilon^{(1)}(S_{j-1} + tX_j) \cdot X_j) dt. \end{aligned}$$

It is well known that if φ is a continuous linear form on B and X a random element in B which is Bochner or Pettis integrable, $\mathbb{E} \varphi(X_j) = \varphi(\mathbb{E} X_j)$. Combining this property with the independence of S_{j-1} and X_j gives via an obvious Fubini argument that

$$\mathbb{E} (g_\varepsilon^{(1)}(S_{j-1}) \cdot X_j) = (\mathbb{E} g_\varepsilon^{(1)}(S_{j-1})) \cdot (\mathbb{E} X_j) = 0.$$

This enables us to rewrite the above decomposition of $\mathbb{E} g_\varepsilon(S_n)$ as

$$\mathbb{E} g_\varepsilon(S_n) = \sum_{j=1}^n \int_0^1 \mathbb{E} \left((g_\varepsilon^{(1)}(S_{j-1} + tX_j) - g_\varepsilon^{(1)}(S_{j-1})) \cdot X_j \right) dt.$$

As g_ε satisfies Lemma 6 with $[p] = \{p\} = 1$, we can use (22) to obtain

$$\begin{aligned} \mathbb{E} g_\varepsilon(S_n) &\leq \sum_{j=1}^n \int_0^1 \mathbb{E} \left\| (g_\varepsilon^{(1)}(S_{j-1} + tX_j) - g_\varepsilon^{(1)}(S_{j-1})) \cdot X_j \right\| dt \\ &\leq \sum_{j=1}^n \int_0^1 \mathbb{E} \left(\frac{C}{\varepsilon} \|tX_j\| \|X_j\| \right) dt \end{aligned}$$

Going back to (42), this gives

$$\mathbb{E} \|S_n\| \leq \varepsilon + \frac{C}{\varepsilon} \sum_{j=1}^n \mathbb{E} \|X_j\|^2. \quad (43)$$

Minimizing in ε this upper bound, yields (40) with $K = 2C^{1/2}$. \square

3.2 The case of Hilbert spaces

Example 12 Let \mathcal{H} be a separable Hilbert space with the inner product $\langle x, y \rangle$ and the norm $\|x\| = \sqrt{\langle x, x \rangle}$, $x, y \in \mathcal{H}$. Then $\psi(x) = \|x\|$ satisfies $c_j := \sup_{\|x\|=1} \|\psi^{(j)}(x)\| < \infty$ for any $j \geq 1$. This can be seen from $\psi(x) = (\langle x, x \rangle)^{1/2}$ and the fact that the inner product is a bilinear function, hence its first derivative is a linear function whereas its second one is a constant. So in Hilbert space the convergence $P_n \xrightarrow[n \rightarrow \infty]{w} P$ is equivalent to $P_n f \rightarrow P f$ for any $f \in C_b^\infty(\mathcal{H})$. As well the weak convergence is metrizable by $\zeta_d(P_n, P)$ for $d \geq 1$. The following result proved by Giné and León [12] is also a corollary of Theorems 4 and 7.

Theorem 13 *Let \mathcal{H} be a separable Hilbert space. Then for $(P, P_n, n \in \mathbb{N}) \subset \mathcal{P}(\mathcal{H})$ the following statements are equivalent:*

- (i) $P_n \xrightarrow[n \rightarrow \infty]{w} P$,
- (ii) $P_n f \xrightarrow[n \rightarrow \infty]{} P f$ for every $f \in C_b^{(\infty)}(\mathcal{H})$;
- (iii) for at least one $d > 1$, $\lim_{n \rightarrow \infty} \zeta_d(P_n, P) = 0$.

3.3 Smoothness of L_p spaces

Example 14 Let $(\mathbb{S}, \mathcal{S}, \nu)$ be a σ -finite measure space, $p \geq 1$. By $\mathcal{L}_p(\mathbb{S}, \nu; \mathbb{R})$ we denote the set of measurable functions $x : \mathbb{S} \rightarrow \mathbb{R}$ such that $\int_{\mathbb{S}} |x(s)|^p \nu(ds) < \infty$. The corresponding Banach space is denoted by $L_p(\mathbb{S}, \mathcal{S}, \nu; \mathbb{R})$ or shortly $L_p(\mathbb{S}, \nu)$ and is endowed with the norm

$$\|x\|_{L_p} := \left(\int_{\mathbb{S}} |x(s)|^p \nu(ds) \right)^{1/p}.$$

Throughout we assume that the spaces $L_p(\mathbb{S}, \nu)$, $p \geq 1$, are separable. This is the case if \mathcal{S} is countably generated or if $(\mathbb{S}, \mathcal{S}, \nu)$ is ν -countably generated: there

exists a sequence $(S_n, n \geq 1) \subset \mathcal{S}$, consisting of sets of finite ν -measure, which ν -essentially generates \mathcal{S} in the sense that for all $A \in \mathcal{S}$ we can find a set A_0 in the σ -algebra generated by $(S_n, n \geq 1)$ such that $\nu(A \Delta A_0) = 0$, see Proposition 1.49 in Hytönen *et al.* [16].

As proved in [19, Prop. 2.23], the norm $\psi(x) = \|x\|_{L_p}$ is $\lfloor p \rfloor$ times continuously differentiable on $L_p(\mathbb{S}, \nu) \setminus \{0\}$ and satisfies $\sum_{k=0}^{\lfloor p \rfloor} \sup_{\|x\|=1} \|\psi^{(k)}(x)\| < \infty$. We use here the following notations : $\psi_1(x) := (\psi(x))^p$, $g(t) := |t|^p$, $f(t) := |t|^{1/p}$. The method used in the proof of [19, Prop. 2.23] is to establish the $\lfloor p \rfloor$ -continuous differentiability of ψ_1 by a Taylor formula techniques and as $\psi = f \circ \psi_1$ and f is infinitely differentiable on $\mathbb{R} \setminus \{0\}$, the $\lfloor p \rfloor$ times continuous differentiability of ψ on $L_p(\mathbb{S}, \nu) \setminus \{0\}$ follows. In what follows we adopt the Toscano notation for the falling factorial, that is for any real number r and any integer $k \geq 1$,

$$r^{\underline{k}} := \prod_{i=0}^{k-1} (r - i).$$

With this notation the derivatives of f and g are conveniently expressed as

$$f^{(k)}(t) = \left(\frac{1}{p}\right)^k |t|^{1/p-k} \operatorname{sgn}(t)^k, \quad g^{(k)}(t) = p^k |t|^{p-k} \operatorname{sgn}(t)^k, \quad t \neq 0, k \geq 1. \quad (44)$$

In the proof of [19, Prop. 2.23], it is shown that $\psi_1^{(k)}(x) = A_k(x)$, $k = 1, \dots, r$, where the k -linear form $A_k(x)$ is defined by

$$A_k(x)(h_1, \dots, h_k) = \int_{\mathbb{S}} g^{(k)}(x(s)) h_1(s) \cdots h_k(s) \nu(ds), \quad h_1, \dots, h_k \in L_p. \quad (45)$$

For our aim it is useful to explicit here the iterated use of Hölder inequality mentioned in [19] to check the continuity of the k -linear operator $A_k(x)$. This way we obtain a bound of the norm $\|A_k(x)\|_{\mathcal{L}_k(L_p)}$ in terms of p , k and $\|x\|_{L_p}$. In view of (44), the problem is reduced to the successive “extractions” of $\|h_1\|_{L_p}, \dots, \|h_k\|_{L_p}$ via Hölder inequality applied iteratively along an *ad hoc* sequence $(p_1, q_1), \dots, (p_k, q_k)$ of conjugate exponents, starting from the integral $J_1 := \int_{\mathbb{S}} |x|^{p-k} |h_1| \dots |h_k| \nu$. To this aim we choose $p_i = p - i + 1$, $q_i = p_i / (p_i - 1) = (p - i + 1) / (p - i)$, $1 \leq i \leq k$. It is easily seen that

$$q_1 \cdots q_i = \frac{p}{p-i}, \quad q_1 \cdots q_{i-1} p_i = p, \quad i = 1, \dots, k. \quad (46)$$

The step $i \rightarrow i + 1$ of this procedure consist in applying Hölder inequality as follows :

$$\begin{aligned} J_i &:= \left(\int_{\mathbb{S}} |x|^{q_1 \cdots q_{i-1} (p-k)} |h_i \cdots h_k|^{q_1 \cdots q_{i-1}} \nu \right)^{\frac{1}{q_1 \cdots q_{i-1}}} \\ &\leq \left(\int_{\mathbb{S}} |x|^{q_1 \cdots q_i (p-k)} |h_{i+1} \cdots h_k|^{q_1 \cdots q_i} \nu \right)^{\frac{1}{q_1 \cdots q_i}} \times \left(\int_{\mathbb{S}} |h_i|^{q_1 \cdots q_{i-1} p_i} \nu \right)^{\frac{1}{q_1 \cdots q_{i-1} p_i}} \\ &= J_{i+1} \|h_i\|_{L_p}. \end{aligned}$$

At the end of this procedure we obtain $J_1 \leq J_{k+1} \|h_1\|_{L_p} \cdots \|h_k\|_{L_p}$, where

$$J_{k+1} = \left(\int_{\mathbb{S}} |x|^{q_1 \cdots q_k (p-k)} d\mathbf{v} \right)^{\frac{1}{q_1 \cdots q_k}} = \left(\int_{\mathbb{S}} |x|^p d\mathbf{v} \right)^{\frac{p-k}{p}} = \|x\|_{L_p}^{p-k},$$

using (46). From this bound for J_1 we deduce that for every $x \in B \setminus \{0\}$ the integral in (45) is well defined, that the k -linear operator $A_k(x) : (L_p)^k \rightarrow \mathbb{R}$ is continuous and satisfies

$$\left\| \psi_1^{(k)}(x) \right\| = \|A_k(x)\|_{\mathcal{L}_k(L_p)} \leq p^k \|x\|_{L_p}^{p-k}, \quad x \neq 0, \quad 1 \leq k \leq \lfloor p \rfloor. \quad (47)$$

To prove the p -smoothness of L_p we have to check (3). Recalling that $\psi = f \circ \psi_1$, we can use differentiation of composite functions on Banach spaces as in the proof of Lemma 6:

$$\begin{aligned} \psi^{(m)}(x)(h_1, \dots, h_m) &= \frac{d^m f(\psi_1(x))}{dx^m}(h_1, \dots, h_m) \\ &= \sum_{j=1}^m \sum_{\substack{\beta \in \langle j \rangle_+ \\ |\beta|=m}} \sum_{\sigma} \frac{f^{(j)}(\psi_1(x))}{j! \beta!} I_{\beta, \sigma}(x)(h_1, \dots, h_m), \quad 1 \leq m \leq \lfloor p \rfloor, \end{aligned} \quad (48)$$

with the same summation conventions as in (25) and

$$I_{\beta, \sigma}(x)(h_1, \dots, h_m) := \psi_1^{(\beta_1)}(x)(h_{\sigma_1}, \dots, h_{\sigma_{\beta_1}}) \cdots \psi_1^{(\beta_j)}(x)(h_{\sigma_{m-\beta_j+1}}, \dots, h_{\sigma_m}). \quad (49)$$

Write $\mathbb{U} := \{x \in L_p(\mathbb{S}, \mathbf{v}) : \|x\|_{L_p} = 1\}$ for the unit sphere of $L_p(\mathbb{S}, \mathbf{v})$. As $\psi_1(x) = 1$ for $x \in \mathbb{U}$, it follows from (44) that

$$|f^{(j)}(\psi_1(x))| = \left| \left(\frac{1}{p} \right)^j \right|, \quad x \in \mathbb{U}. \quad (50)$$

Moreover $\beta \in \langle j \rangle_+$, have all its components $\beta_i \geq 1$, so by (47),

$$\left\| \psi_1^{(\beta_i)}(x) \right\| \leq p^{\beta_i}, \quad x \in \mathbb{U}. \quad (51)$$

Gathering (48) to (51) we obtain

$$\sum_{k=1}^{\lfloor p \rfloor} \sup_{x \in \mathbb{U}} \left\| \psi^{(k)}(x) \right\| < \infty.$$

It remains to check that $\psi^{(\lfloor p \rfloor)}$ satisfies

$$\sup_{\substack{x, y \in \mathbb{U} \\ x \neq y}} \frac{\left\| \psi^{(\lfloor p \rfloor)}(x) - \psi^{(\lfloor p \rfloor)}(y) \right\|}{\|x - y\|_{\{p\}}} < \infty. \quad (52)$$

As a preliminary, we check the following inequality

$$\left| |a|^\alpha \operatorname{sgn}(a) - |b|^\alpha \operatorname{sgn}(b) \right| \leq 2^{1-\alpha} |a-b|^\alpha, \quad a, b \in \mathbb{R}, \alpha \in (0, 1). \quad (53)$$

To this aim, we put $c := \max(|a|, |b|)$, $d := \min(|a|, |b|)$ and use the elementary inequalities

$$1 \leq t^\alpha + (1-t)^\alpha \leq 2^{1-\alpha}, \quad 0 \leq t \leq 1. \quad (54)$$

If $\operatorname{sgn}(a) = \operatorname{sgn}(b)$, the choice of $t = d/c$ in (54) leads to $c^\alpha - d^\alpha \leq (c-d)^\alpha$, whence

$$\left| |a|^\alpha \operatorname{sgn}(a) - |b|^\alpha \operatorname{sgn}(b) \right| = \left| |a|^\alpha - |b|^\alpha \right| = c^\alpha - d^\alpha \leq (c-d)^\alpha = |a-b|^\alpha.$$

If $\operatorname{sgn}(a) \neq \operatorname{sgn}(b)$, the choice of $t = c/(c+d)$ in (54) gives $c^\alpha + d^\alpha \leq 2^{1-\alpha}(c+d)^\alpha$, whence

$$\left| |a|^\alpha \operatorname{sgn}(a) - |b|^\alpha \operatorname{sgn}(b) \right| = c^\alpha + d^\alpha \leq 2^{1-\alpha}(c+d)^\alpha = 2^{1-\alpha}|a-b|^\alpha.$$

Proof of (52).

Case $1 < p < 2$. Here $\lfloor p \rfloor = 1$ and for $x, y \in L_p$ with $\|x\|_{L_p} = \|y\|_{L_p} = 1$ we have

$$\begin{aligned} |\psi'(x)(h) - \psi'(y)(h)| &= \left| \frac{1}{p} (\psi_1(x))^{(1-p)/p} \psi_1'(x)(h) - \psi_1(y)^{(1-p)/p} \psi_1'(y)(h) \right| \\ &= \frac{1}{p} |\psi_1'(x)(h) - \psi_1'(y)(h)|. \end{aligned}$$

As ψ_1' is the linear form A_1 , (44), (45) and (53) with $\alpha = p-1$ give

$$\begin{aligned} |\psi'(x)(h) - \psi'(y)(h)| &= \left| \int_{\mathbb{S}} \left[|x(s)|^{p-1} \operatorname{sgn}(x(s)) - |y(s)|^{p-1} \operatorname{sgn}(y(s)) \right] h(s) \nu(ds) \right| \\ &\leq \int_{\mathbb{S}} \left| |x(s)|^{p-1} \operatorname{sgn}(x(s)) - |y(s)|^{p-1} \operatorname{sgn}(y(s)) \right| |h(s)| \nu(ds) \\ &\leq \int_{\mathbb{S}} 2^{2-p} |x(s) - y(s)|^{p-1} |h(s)| \nu(ds) \end{aligned}$$

Applying Hölder inequality with exponents p and $q = p/(p-1)$ we obtain

$$|\psi'(x)(h) - \psi'(y)(h)| \leq 2^{2-p} \|x-y\|_{L_p}^{p-1} \|h\|_{L_p}.$$

This inequality being valid for every h in $L_p(\mathbb{S}, \nu)$ and as $\{p\} = p-1$ here, it follows that $\|\psi'(x) - \psi'(y)\| \leq 2^{2-p} \|x-y\|_{L_p}^{\{p\}}$, so (52) is satisfied when $1 < p < 2$.

Case $p \geq 2$. By (48) and (50), we have for $x, y \in \mathbb{U}$,

$$\left\| \psi^{(\lfloor p \rfloor)}(x) - \psi^{(\lfloor p \rfloor)}(y) \right\| \leq \sum_{j=1}^{\lfloor p \rfloor} \sum_{\substack{\beta \in \langle j \rangle_+ \\ |\beta| = \lfloor p \rfloor}} \sum_{\sigma} \frac{|(1/p)^j|}{j! \beta!} \|\Delta_{\beta, \sigma}(x, y)\|, \quad (55)$$

where $\Delta_{\beta,\sigma}(x,y) := I_{\beta,\sigma}(x) - I_{\beta,\sigma}(y)$. Using telescopic summation as in the proof of Lemma 6 and (51), we get

$$\begin{aligned} \|\Delta_{\beta,\sigma}(x,y)\| &\leq \sum_{i=1}^j \left(\prod_{1 \leq k < i} \|\psi_1^{(\beta_k)}(x)\| \right) \|\psi_1^{(\beta_i)}(x) - \psi_1^{(\beta_i)}(y)\| \left(\prod_{i < k \leq j} \|\psi_1^{(\beta_k)}(y)\| \right) \\ &\leq \sum_{i=1}^j \|\psi_1^{(\beta_i)}(x) - \psi_1^{(\beta_i)}(y)\| \prod_{\substack{1 \leq k \leq j \\ k \neq i}} p^k. \end{aligned} \quad (56)$$

Now it remains to find a suitable control of each increment $\|\psi_1^{(\beta_i)}(x) - \psi_1^{(\beta_i)}(y)\|$.

If $2 \leq j \leq \lfloor p \rfloor$, the multiindex β has at least two components, so $1 \leq \beta_i < \lfloor p \rfloor$. Then $\psi_1^{\beta_i}$ has a continuous derivative on $\mathbb{B} \setminus \{0\}$. So if $0 \notin [x,y]$, recalling (47), we get

$$\begin{aligned} \|\psi_1^{(\beta_i)}(x) - \psi_1^{(\beta_i)}(y)\| &\leq \sup_{z \in [x,y]} \|\psi_1^{(\beta_i+1)}(z)\| \|x-y\| \\ &\leq p^{\beta_i+1} \sup_{t \in [0,1]} \|(1-t)x + ty\|^{p-\beta_i-1} \|x-y\| \\ &\leq 2^{p-\beta_i-1} p^{\beta_i+1} \|x-y\|. \end{aligned} \quad (57)$$

To complete the case $j \geq 2$, notice that $0 \in [x,y]$ if and only if $y = -x$. In this special case, $\|x-y\| = 2$ and accounting (51), we can simply write

$$\|\psi_1^{(\beta_i)}(x) - \psi_1^{(\beta_i)}(y)\| \leq \|\psi_1^{(\beta_i)}(x)\| + \|\psi_1^{(\beta_i)}(-x)\| \leq 2p^{\beta_i} = p^{\beta_i} \|x-y\|. \quad (58)$$

Now, $\|x-y\| \leq 2^{1-\lfloor p \rfloor} \|x-y\|^{\lfloor p \rfloor}$ for $x,y \in \mathbb{U}$, so from (56)–(58), there is a constant K dependent only on the space $L_p(\mathbb{S}, \nu)$ such that

$$\sum_{j=2}^{\lfloor p \rfloor} \sum_{\substack{\beta \in \langle j \rangle_+ \\ |\beta| = \lfloor p \rfloor}} \sum_{\sigma} \frac{\lfloor (1/p)^j \rfloor}{j! \beta!} \|\Delta_{\beta,\sigma}(x,y)\| \leq K \|x-y\|^{\lfloor p \rfloor}, \quad x,y \in \mathbb{U}. \quad (59)$$

It remains to treat the sum of terms for which $j = 1$ in (55). Here β is a mono-index necessarily equal to $\lfloor p \rfloor$ and by (56), one can bound this remaining sum R as

$$R := \sum_{\sigma} \frac{1}{p \lfloor p \rfloor!} \|\Delta_{\lfloor p \rfloor, \sigma}(x,y)\| \leq \frac{1}{p} \|\psi_1^{(\lfloor p \rfloor)}(x) - \psi_1^{(\lfloor p \rfloor)}(y)\|.$$

Recalling (45) and (44), we have for $h_1, \dots, h_{\lfloor p \rfloor} \in L_p(\mathbb{S}, \nu)$,

$$(\psi_1^{(\lfloor p \rfloor)}(x) - \psi_1^{(\lfloor p \rfloor)}(y))(h_1, \dots, h_{\lfloor p \rfloor}) = p^{\lfloor p \rfloor} \int_{\mathbb{S}} (|x|^{\lfloor p \rfloor} \operatorname{sgn} x - |y|^{\lfloor p \rfloor} \operatorname{sgn} y) h_1 \cdots h_{\lfloor p \rfloor} \, d\nu$$

Using iteratively Hölder inequality exactly as in the proof of (47), we obtain

$$\begin{aligned} \left\| \psi_1^{(\lfloor p \rfloor)}(x) - \psi_1^{(\lfloor p \rfloor)}(y) \right\| &\leq p^{\lfloor p \rfloor} \left(\int_{\mathbb{S}} \left| |x|^{\lfloor p \rfloor} \operatorname{sgn} x - |y|^{\lfloor p \rfloor} \operatorname{sgn} y \right|^{\frac{p}{\lfloor p \rfloor}} d\nu \right)^{\frac{\lfloor p \rfloor}{p}} \\ &\leq 2^{1-\lfloor p \rfloor} p^{\lfloor p \rfloor} \left(\int_{\mathbb{S}} |x-y|^p d\nu \right)^{\frac{\lfloor p \rfloor}{p}}, \end{aligned}$$

thanks to (53). Finally $R \leq K' \|x-y\|^{\lfloor p \rfloor}$ with a constant K' depending only on p and recalling (59), this completes the proof of (52). \square

Recalling that by [19, Prop. 2.23], if $p = 2\ell$ is an even integer number, then the norm $\psi(x)$ is infinitely many times Fréchet differentiable, we can summarize about L_p smoothness by the following proposition.

Proposition 15

- (a) For any $p > 1$, the space $L_p(\mathbb{S}, \nu)$ is p -smooth.
- (b) If $p = 2\ell$ is an even integer, then the norm of $L_p(\mathbb{S}, \nu)$ is infinitely many times Fréchet differentiable on $L_p(\mathbb{S}, \nu) \setminus \{0\}$ and has bounded derivatives on the unit circle, so that $L_p(\mathbb{S}, \nu)$ is d -smooth for any integer $d \geq 1$.

Theorem 4 and Proposition 15 yield the following results.

Theorem 16 Let $p \geq 1$. For $(P, P_n, n \in \mathbb{N}) \subset \mathcal{P}(L_p(\mathbb{S}, \nu))$, the following statements are equivalent:

- (i) $P_n \xrightarrow[n \rightarrow \infty]{w} P$,
- (ii) $P_n f \xrightarrow[n \rightarrow \infty]{} P f$ for any $f \in C_b^{(p)}(L_p(\mathbb{S}, \nu))$;
- (iii) $\lim_{n \rightarrow \infty} \zeta_p(P_n, P) = 0$.

Theorem 17 If $p \geq 2$ is an even integer, then for $(P, P_n, n \in \mathbb{N}) \subset \mathcal{P}(L_p(\mathbb{S}, \nu))$ the following statements are equivalent:

- (i) $P_n \xrightarrow[n \rightarrow \infty]{w} P$,
- (ii) $P_n f \xrightarrow[n \rightarrow \infty]{} P f$ for any $f \in C_b^{(\infty)}(L_p(\mathbb{S}, \nu))$;
- (iii) for at least one $d > 1$, $\lim_{n \rightarrow \infty} \zeta_d(P_n, P) = 0$.

4 Lindeberg CLT in p -smooth Banach spaces

First we implement in Th. 21 below the main principle of Lindeberg method and compare the sums $\sum_{k=1}^n X_{nk}$ with sums of independent Gaussian random variables. Beforehand, it seems convenient to recall the notion of B -valued stochastic integral with respect to a white noise which plays a key role in our proof of Th. 21.

Definition 18 Let $(\mathbb{S}, \mathcal{S}, \mu)$ be a measure space and $\mathcal{S}_0 := \{A \in \mathcal{S}; \mu(A) < \infty\}$. A white noise with variance μ is a stochastic process $W = (W(A); A \in \mathcal{S}_0)$ defined on some rich enough probability space $(\Omega', \mathcal{F}', P')$ such that

- (a) for each $A \in \mathcal{S}_0$, $W(A)$ is a real valued Gaussian random variable with mean zero and variance $\mu(A)$;
- (b) if $A_1 \in \mathcal{S}_0, \dots, A_j \in \mathcal{S}_0$ are disjoint, then $W(A_1), \dots, W(A_j)$ are independent and

$$W\left(\bigcup_{i=1}^j A_i\right) = \sum_{i=1}^j W(A_i), \quad j \geq 2.$$

Next, following Proposition 3.3 in Hoffmann Jørgensen and Pisier [15], one can construct a \mathbb{B} -valued stochastic integral with respect to W . Classically we define first this integral for functions g in the space $\mathcal{L}_0(\mu)$ of \mathcal{S}_0 -simple functions g , that is of the form $g = \sum_{i=1}^j x_i \mathbf{1}_{A_i}$, with $x_i \in \mathbb{B}$, $A_i \in \mathcal{S}_0$, $1 \leq i \leq j$, $j \geq 1$ and extend it to the whole space $L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$. The next proposition is essentially stated and proved in [15]. Our rewriting of its statement and proof is motivated by the need to explicit Corollary 20 in view of its role in Th. 21.

Proposition 19 *If \mathbb{B} is of type 2 and W is a white noise with variance μ on some probability space $(\Omega', \mathcal{F}', \mathbb{P}')$, then there exists a unique linear map*

$$I_W : L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B}) \longrightarrow L^2(\Omega', \mathcal{F}', \mathbb{P}', \mathbb{B}), \quad g \longmapsto I_W(g) = \int_{\mathbb{S}} g dW,$$

such that the following statements hold.

- (a) For every $g = \sum_{i=1}^j x_i \mathbf{1}_{A_i}$, where $x_1, \dots, x_j \in \mathbb{B}$, $A_1, \dots, A_j \in \mathcal{S}_0$, $j \geq 1$,

$$I_W(g) = \int_{\mathbb{S}} g dW := \sum_{i=1}^j x_i W(A_i). \quad (60)$$

- (b) There exists a constant C such that for every $g \in L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$,

$$\mathbb{E} \left\| \int_{\mathbb{S}} g dW \right\|^2 \leq C \int_{\mathbb{S}} \|g\|^2 d\mu. \quad (61)$$

- (c) For every $g \in L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$, $\int_{\mathbb{S}} g dW$ is a Gaussian mean zero random element in \mathbb{B} .
- (d) If $D', D'' \in \mathcal{B}_{\mathbb{B}}$ are disjoint, $\int_{D'} g dW$ and $\int_{D''} g dW$ are independent for every g in $L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$.

Proof. The coherence of the definition of $I_W(g)$ by (60) when g is a \mathcal{S}_0 -simple function is checked in a standard way using the additivity property (b) in Definition 18. Checking the linearity of I_W on the subspace $\mathcal{L}_0(\mu)$ of simple functions in $L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$ is then straightforward. Next, if $g \in \mathcal{L}_0(\mu)$, it can be represented as $g = \sum_{i=1}^j x_i \mathbf{1}_{A_i}$ where the $A_i \in \mathcal{S}_0$ are *disjoint*, which was not requested in (60). As \mathbb{B} is of type 2 and the $x_i W(A_i)$ are independent with mean zero and finite second moment, there is a constant C depending only on \mathbb{B} such that

$$\mathbb{E} \|I_W(g)\|^2 = \mathbb{E} \left\| \sum_{i=1}^j x_i W(A_i) \right\|^2 \leq C \sum_{i=1}^j \mathbb{E} \|x_i W(A_i)\|^2 = C \sum_{i=1}^j \|x_i\|^2 \mathbb{E} W(A_i)^2.$$

As the random variables $W(A_i)$ are mean zero with respective variances $\mu(A_i)$, this implies

$$\mathbb{E} \|I_W(g)\|^2 \leq C \sum_{i=1}^j \|x_i\|^2 \mu(A_i) = C \int_{\mathbb{S}} \|g\|^2 d\mu.$$

Therefore I_W is a *continuous* linear map $\mathcal{L}_0(\mu) \longrightarrow L^2(\Omega', \mathcal{F}', P', \mathbb{B})$ and by density of $\mathcal{L}_0(\mu)$ in $L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$, I_W has a unique continuous linear extension to this space, still denoted I_W , and satisfying (61) with the same constant C .

To prove (c), we check that for every $u \in \mathbb{B}^*$, $u(I_W(g))$ is a Gaussian random variable. This is clear for g simple since then $u(I_W(g))$ is a linear combination of independent Gaussian random variables. In the general case, g is the limit in $L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$ of a sequence (g_n) of simple functions. Combining the continuity of the linear functional u with (61) gives

$$\mathbb{E} |u(I_W(g)) - u(I_W(g_n))|^2 \leq C \|u\|^2 \int_{\mathbb{S}} \|g - g_n\|^2 d\mu,$$

which shows that $u(I_W(g))$ is a Gaussian random variable as limit in quadratic mean of a sequence of Gaussian random variables. Moreover, for every $u \in \mathbb{B}$, $\mathbb{E} u(I_W(g)) = \lim_{n \rightarrow \infty} \mathbb{E} u(I_W(g_n)) = 0$, whence $\mathbb{E} I_W(g) = 0$.

To prove (d), we first note that for g simple, $g = \sum_{i=1}^j x_i \mathbf{1}_{A_i}$ with the A_i disjoint, and $D \in \mathcal{B}_{\mathbb{B}}$, $\int_D g dW := \int_{\mathbb{S}} g \mathbf{1}_D dW = \sum_{i=1}^j x_i W(A_i \cap D)$. Since D' and D'' are disjoint, so are the sets $A_1 \cap D', \dots, A_j \cap D', A_1 \cap D'', \dots, A_j \cap D''$, which provides the independence of

$$Y' = \int_{D'} g dW = \sum_{i=1}^j x_i W(A_i \cap D') \quad \text{and} \quad Y'' = \int_{D''} g dW = \sum_{i=1}^j x_i W(A_i \cap D'').$$

This independence is preserved when g is the limit in $L^2(\mathbb{S}, \mathcal{S}, \mu, \mathbb{B})$ of a sequence (g_n) of simple functions since then, $Y'_n = \int_{D'} g_n dW$ and $Y''_n = \int_{D''} g_n dW$ converge in probability to $Y' = \int_{D'} g dW$ and $Y'' = \int_{D''} g dW$ respectively which implies the convergence in distribution of (Y'_n, Y''_n) to (Y', Y'') . Then the distribution of (Y', Y'') is the product of the distributions of Y' and Y'' which is equivalent to the independence of Y' and Y'' . \square

Now, denote by X a \mathbb{B} valued random element defined on a probability space (Ω, \mathcal{F}, P) with distribution $P_X = P \circ X^{-1}$ (which is a probability measure on $\mathcal{B}_{\mathbb{B}}$) and such that $\mathbb{E} X = 0$, $\mathbb{E} \|X\|^2 < \infty$. Let us denote by $Q = \text{cov}(X) \in L(\mathbb{B}^*, \mathbb{B})$ the covariance operator of X , that is the linear bounded operator from \mathbb{B}^* to \mathbb{B} defined by

$$Qu = \mathbb{E} (\langle u, X \rangle X), \quad u \in \mathbb{B}^*.$$

Since \mathbb{B} is of type 2 and $\mathbb{E} \|X\|^2 < \infty$, the operator Q is pregaussian (see, Theorem 3.5. in Hoffmann Jørgensen and Pisier [15]), so there exists a Gaussian mean zero random element Y in \mathbb{B} with covariance operator Q . One way to construct such an Y is to apply Prop. 19 with $\mathbb{S} = \mathbb{B}$, $\mathcal{S} = \mathcal{B}_{\mathbb{B}}$, $\mu = P_X$, which gives the following corollary.

Corollary 20 Let B be a separable type 2 Banach space and X be a random element in B defined on some probability space (Ω, \mathcal{F}, P) . Denote by $P_X := P \circ X^{-1}$ the distribution of X . Assume that $E\|X\|^2 < \infty$ and $EX = 0$. Let $W = (W(A), A \in \mathcal{B}_B)$ be a white noise with variance $\mu = P_X$ defined on some probability space $(\Omega', \mathcal{F}', P')$. As $E\|X\|^2 < \infty$ the identity map, $\text{Id}_B : B \rightarrow B, x \mapsto x$, is in $L^2(B, \mathcal{B}_B, P_X, B)$, so we can define a Gaussian mean zero random element Y in B by

$$Y := \int_B \text{Id}_B dW = \int_B x dW(x). \quad (62)$$

Then the following statements hold.

(a) With the constant C in (61),

$$E\|Y\|^2 \leq CE\|X\|^2. \quad (63)$$

(b) For every $g \in L^2(B, \mathcal{B}_B, P_X, B)$, the Gaussian mean zero random element $Z = \int_B g dW$ has the same covariance operator as $g(X)$. In particular, Y and X have the same covariance operator.

(c) For every symmetric $T \in \mathcal{L}_2(B)$ and every $g \in L^2(B, \mathcal{B}_B, P_X, B)$,

$$ET(g(X), g(X)) = ET(Z, Z) \quad (64)$$

and in particular $ET(X, X) = ET(Y, Y)$.

Proof. (a) is a simple translation of (61) in the special case under consideration. For (b), we have to check that $Q_Z = Q_{g(X)}$ which is equivalent to $E(u(Z)v(Z)) = E(u(g(X))v(g(X)))$ for every u, v in B^* . For $g = \sum_{i=1}^j x_i \mathbf{1}_{A_i}$ with the A_i 's disjoint, $u(Z) = \sum_{i=1}^j u(x_i)W(A_i)$, whence by independence of the A_i 's,

$$\begin{aligned} E(u(Z)v(Z)) &= \sum_{i=1}^j u(x_i)v(x_i)P_X(A_i) = \int_B u(g(x))v(g(x))dP_X(x) \\ &= E(u(g(X))v(g(X))). \end{aligned}$$

Valid for every g simple, this equality extends to the whole space $L^2(B, \mathcal{B}_B, P_X, B)$ by the continuity of I_W . In particular for $g = \text{Id}_B$ and $Y = \int_B \text{Id}_B dW$, $Q_Y = Q_X$.

The proof of (c) is similar and will be omitted. \square

Theorem 21 Assume that B is of type 2. Consider an array of B -valued random variables

$$X_{n1}, \dots, X_{nk}, \dots, X_{nr_n}, \quad 1 \leq k \leq r_n, n \geq 1,$$

where the probability space $(\Omega_n, \mathcal{F}_n, P_n)$ underlying the n^{th} line may vary with n and for each $n \geq 1$, the X_{nk} , $1 \leq k \leq r_n$, are mean zero independent and

$$M_n := \sum_{k=1}^{r_n} E\|X_{nk}\|^2 < \infty. \quad (65)$$

Then for each $n \geq 1$, one can construct on some probability space $(\Omega'_n, \mathcal{F}'_n, P'_n)$ independent mean zero Gaussian B -valued random variables Y_{n1}, \dots, Y_{nr_n} , such that for

$1 \leq k \leq r_n$, X_{nk} and Y_{nk} have the same covariance operator and for any $\delta \in (0, 1]$, and any $\varepsilon > 0$,

$$\zeta_{2+\delta} \left(\sum_{k=1}^{r_n} X_{nk}, \sum_{k=1}^{r_n} Y_{nk} \right) \leq c(\mathbb{B}, \delta) \left(M_n \varepsilon^\delta + (M_n + 1) \sum_{k=1}^{r_n} \mathbb{E} \|X_{nk}\|^2 \mathbf{1}_{\{\|X_{nk}\| > \varepsilon\}} \right), \quad (66)$$

where the constant $c(\mathbb{B}, \delta) > 0$ depends only of the type 2 constant of the space \mathbb{B} and of δ .

Proof. We fix an arbitrary $n \geq 1$ and prove (66) for the n^{th} line of the array. In view of the property (36) of $\zeta_{2+\delta}$, the problem reduces to proving that, given any mean zero random element X in \mathbb{B} such that $\mathbb{E} \|X\|^2 < \infty$, one can construct, possibly on another probability space than the one supporting X , a mean zero Gaussian random element Y in \mathbb{B} with the same covariance operator as X , such that

$$\zeta_{2+\delta}(X, Y) \leq c(\mathbb{B}, \delta) \left(\varepsilon^\delta \mathbb{E} \|X\|^2 + (1 + \mathbb{E} \|X\|^2) \mathbb{E} \|X\|^2 \mathbf{1}_{\{\|X\| > \varepsilon\}} \right). \quad (67)$$

To this aim, choosing Y as in Corollary 20, we have to estimate $|\mathbb{E} f(X) - \mathbb{E} f(Y)|$ for $f \in C_b^{(2+\delta)}(\mathbb{B})$ such that $\|f\|_{2+\delta} \leq 1$. By Taylor formula at the order 1 with integral remainder,

$$f(X) = f(0) + f'(0).X + \int_0^1 (1-t) f''(tX).(X, X) dt. \quad (68)$$

To exploit fully the membership of f in $C_b^{(2+\delta)}(\mathbb{B})$, we rephrase this formula as

$$f(X) = f(0) + f'(0).X + \frac{1}{2} f''(0).(X, X) + \int_0^1 (1-t) (f''(tX) - f''(0)).(X, X) dt.$$

Applying the same treatment to $f(Y)$ and using $\mathbb{E}(f'(0).X) = f'(0).(EX) = 0$ and similarly $\mathbb{E}(f'(0).Y) = 0$, together with (c) in Corollary 20 applied with the bilinear symmetric operator $T = f''(0)$, we are left with

$$|\mathbb{E} f(X) - \mathbb{E} f(Y)| \leq R(X) + R(Y),$$

where

$$R(Z) := \int_0^1 (1-t) \mathbb{E} \|(f''(tZ) - f''(0)).(Z, Z)\| dt, \quad Z = X, Y.$$

By the δ -Hölder continuity of f'' and $\|f\|_{2+\delta} \leq 1$, $\|f''(tZ) - f''(0)\| \leq \|tZ\|^\delta \leq \|Z\|^\delta$ whence

$$\|(f''(tZ) - f''(0)).(Z, Z)\| \leq \|Z\|^\delta \|(Z, Z)\| \leq \|Z\|^{2+\delta} \quad Z = X, Y. \quad (69)$$

As Y is Gaussian, $\mathbb{E} \|Y\|^{2+\delta} < \infty$, which gives a first estimate of $R(Y)$, by integration with respect to t in (69):

$$R(Y) \leq \frac{1}{2} \mathbb{E} \|Y\|^{2+\delta}. \quad (70)$$

Concerning $R(X)$, only the finiteness of $\mathbb{E}\|X\|^2$ is available, so we use (69) only on the event $\{X \leq \varepsilon\}$ where $\|(f''(tX) - f''(0)).(X, X)\| \leq \varepsilon^\delta \|X\|^2$. This gives

$$\begin{aligned} & \int_0^1 (1-t) \mathbb{E}(\|(f''(tX) - f''(0)).(X, X)\| \mathbf{1}_{\{\|X\| \leq \varepsilon\}}) dt \\ & \leq \int_0^1 (1-t) \varepsilon^\delta \mathbb{E}\|X\|^2 \mathbf{1}_{\{\|X\| \leq \varepsilon\}} dt \leq \frac{1}{2} \varepsilon^\delta \mathbb{E}\|X\|^2. \end{aligned} \quad (71)$$

On $\{\|X\| > \varepsilon\}$, we simply use the fact that $\|f''\| \leq 1$, so $\|f''(tX) - f''(0)\| \leq 2$. This gives

$$\begin{aligned} & \int_0^1 (1-t) \mathbb{E}(\|(f''(tX) - f''(0)).(X, X)\| \mathbf{1}_{\{\|X\| > \varepsilon\}}) dt \\ & \leq \int_0^1 2(1-t) \mathbb{E}\|X\|^2 \mathbf{1}_{\{\|X\| > \varepsilon\}} dt = \mathbb{E}\|X\|^2 \mathbf{1}_{\{\|X\| > \varepsilon\}}. \end{aligned} \quad (72)$$

Our next step is to control the bound (70) in terms of the distribution of X only. Since Y is Gaussian, there is for every $r > 0$ a constant κ_r depending on r only, such that $(\mathbb{E}\|Y\|^r)^{1/r} \leq \kappa_r (\mathbb{E}\|Y\|^2)^{1/2}$. One possible value is obtained using the inequality $P(\|Y\| > t) \leq 4 \exp(-t^2/(8c^2))$ where $c^2 = \mathbb{E}\|Y\|^2$, see e.g. (3.5) in [17], which gives $\kappa_r = 2^{3/2+2/r} \Gamma(r/2+1)^{1/r}$. In particular,

$$\mathbb{E}\|Y\|^{2+\delta} \leq \kappa_{2+\delta}^{2+\delta} \left(\mathbb{E}\|Y\|^2\right)^{1+\delta/2}.$$

Next, recalling (62), we note that $Y = Y'_\varepsilon + Y''_\varepsilon$, where

$$Y'_\varepsilon := \int_B x \mathbf{1}_{\{\|x\| \leq \varepsilon\}} dW(x) \quad \text{and} \quad Y''_\varepsilon := \int_B x \mathbf{1}_{\{\|x\| > \varepsilon\}} dW(x)$$

are independent Gaussian random elements in B by Proposition 19 and Corollary 20. Since B is of type 2, it follows by (39) that

$$\mathbb{E}\|Y\|^{2+\delta} \leq \kappa_{2+\delta}^{2+\delta} K^{2+\delta} \left(\mathbb{E}\|Y'_\varepsilon\|^2 + \mathbb{E}\|Y''_\varepsilon\|^2\right)^{1+\delta/2}.$$

By Proposition 19 (b) and the convexity inequality $(a+b)^r \leq 2^{r-1}(a^r + b^r)$, $a, b \geq 0$, $r \geq 1$, we obtain with a constant $\gamma_\delta := 2^{\delta/2} \kappa_{2+\delta}^{2+\delta} K^{2+\delta} C^{1+\delta/2}$, C being as in (61),

$$\begin{aligned} \mathbb{E}\|Y\|^{2+\delta} & \leq \gamma_\delta \left(\mathbb{E}\|X\|^{2+\delta} \mathbf{1}_{\{\|X\| \leq \varepsilon\}} + \left(\mathbb{E}\|X\|^2 \mathbf{1}_{\{\|X\| > \varepsilon\}}\right)^{1+\delta/2} \right) \\ & \leq \gamma_\delta \left(\varepsilon^\delta \mathbb{E}\|X\|^2 + \left(\mathbb{E}\|X\|^2\right)^{\delta/2} \mathbb{E}\|X\|^2 \mathbf{1}_{\{\|X\| > \varepsilon\}} \right) \\ & \leq \gamma_\delta \left(\varepsilon^\delta \mathbb{E}\|X\|^2 + \left(1 + \mathbb{E}\|X\|^2\right) \mathbb{E}\|X\|^2 \mathbf{1}_{\{\|X\| > \varepsilon\}} \right). \end{aligned} \quad (73)$$

Now gathering (70), (71), (72) and (73) gives (67) with $c(B, \delta) = (1 + \gamma_\delta/2)$.

To conclude, choose a probability space $(\Omega'_n, \mathcal{F}'_n, P'_n)$ rich enough to support a sequence of independent white noises $(W_{nk})_{1 \leq k \leq r_n}$ where the variance of W_{nk} is

the distribution of X_{nk} . Define on this probability space the corresponding sequence $(Y_{nk})_{1 \leq k \leq r_n}$ of Gaussian random elements in \mathbb{B} by $Y_{nk} := \int_{\mathbb{B}} \text{Id}_{\mathbb{B}} dW_{nk}$, $1 \leq k \leq r_n$. Each pair X_{nk}, Y_{nk} satisfies (67). Bounding $(1 + \mathbb{E}\|X_{nk}\|^2)$ by $1 + M_n$ and summing over $k = 1, \dots, r_n$, we obtain (66). \square

Hence, in p -smooth Banach space \mathbb{B} where $p > 2$, the proof of convergence in distribution of the sequence $\sum_{k=1}^n X_{nk}, n \in \mathbb{N}$, to a \mathbb{B} -valued Gaussian random variable Y_Q , is reduced by Theorem 21 to the proof of convergence in distribution of the Gaussian sequence $\sum_{k=1}^n Y_{nk}$ to Y_Q . The later is controlled by convergence of covariance operators. In any finite dimensional space this is not a problem. In any separable Hilbert space as well as in Banach space of type 2 with approximation property, the convergence $\sum_{k=1}^n Y_{nk} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y_Q$ is obtained from convergence of covariances in nuclear norm (see Chevet [8]).

Recall an operator $u \in \mathcal{L}(\mathbb{B})$ is said to be nuclear if it admits the representation

$$u(x) = \sum_{k=1}^{\infty} f_k(x)y_k,$$

where $f_k \in \mathbb{B}^*, y_k \in \mathbb{B}$, and

$$\sum_{k=1}^{\infty} \|f_k\| \cdot \|y_k\| < \infty.$$

The greatest lower bound of the sum $\sum_{k=1}^{\infty} \|f_k\| \cdot \|y_k\|$ taken over all possible representations of u is called the nuclear norm of u and is denoted by $v_1(u)$.

Theorem 22 *Let the Banach space \mathbb{B} be p -smooth for some $p > 2$ and has approximation property. For each $n \geq 1$ suppose that X_{n1}, \dots, X_{nr_n} is a sequence of mean zero independent \mathbb{B} -valued random elements such that $\sup_{n \in \mathbb{N}} \sum_{k=1}^{r_n} \mathbb{E}\|X_{nk}\|^2 < \infty$. Let $Q_{nj} := \text{cov}(X_{nj}), j = 1, \dots, r_n, n \geq 1$. If there is a linear bounded operator $Q \in L(\mathbb{B}^*, \mathbb{B})$ such that*

$$\lim_{n \rightarrow \infty} v_1\left(\sum_{j=1}^{r_n} Q_{nj} - Q\right) = 0, \quad (74)$$

and for each $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{r_n} \mathbb{E}\|X_{nk}\|^2 \mathbf{1}\{\|X_{nk}\| > \varepsilon\} = 0, \quad (75)$$

then Q is pregaussian and

$$\sum_{k=1}^{r_n} X_{nk} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y_Q, \quad (76)$$

where Y_Q is a mean zero Gaussian random element in \mathbb{B} with covariance Q .

Proof. Let the Gaussian triangular array $(Y_{nk}, k = 1, \dots, r_n; n \geq 1)$ be as constructed in Theorem 21. Since

$$\zeta_p\left(\sum_{k=1}^{r_n} X_{nk}, Y_Q\right) \leq \zeta_p\left(\sum_{k=1}^{r_n} X_{nk}, \sum_{k=1}^{r_n} Y_{nk}\right) + \zeta_p\left(\sum_{k=1}^{r_n} Y_{nk}, Y_Q\right),$$

it is enough by Theorem 21 to prove

$$\lim_{n \rightarrow \infty} \zeta_p \left(\sum_{k=1}^{m_n} Y_{nk}, Y_Q \right) = 0. \quad (77)$$

This is equivalent to the weak convergence of Gaussian distributions and as it is proved in Chevet [8], the convergence $\sum_{k=1}^{r_n} Y_{nk} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y_Q$ follows from (74). \square

To check convergence in nuclear norm might be a quite complex task. In some concrete Banach spaces, the direct proof of convergence in distribution of Gaussian random variables is easier to achieve. As an illustration consider now the case of L_p spaces. In what follows, $(\mathbb{S}, \mathcal{S}, \mu)$ is a measurable space where the measure μ is σ -finite. We denote by p a real in $(2, \infty)$ and by $q = p/(p-1)$ its conjugate exponent. We assume moreover that the space $L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R})$ is separable. We denote respectively by $\mathcal{S} \otimes \mathcal{S}$ and $\mu \otimes \mu$ the product σ -field and product measure on the Cartesian product \mathbb{S}^2 . We will use the abbreviations:

$$L_p(\mathbb{S}) := L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R}), \quad L_p(\mathbb{S}^2) := L_p(\mathbb{S}^2, \mathcal{S} \otimes \mathcal{S}, \mu \otimes \mu; \mathbb{R}).$$

For real valued functions u, v defined μ almost everywhere on \mathbb{S} , $u \otimes v$ denotes the function defined $\mu \otimes \mu$ almost everywhere on \mathbb{S}^2 by $(u \otimes v)(s, t) := u(s)v(t)$. This notation is extended in an obvious way to random elements in $L_p(\mathbb{S})$.

Theorem 23 (CLT in L_p , $p > 2$) *Let $(X_{nk}, k = 1, \dots, r_n; n \in \mathbb{N})$ be a triangular array of mean zero independent random elements in the separable space $L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R})$. Assume that the following conditions are satisfied.*

(a) $\sup_{n \in \mathbb{N}} \sum_{k=1}^{r_n} \mathbb{E} \|X_{nk}\|_{L_p}^2 < \infty.$

(b) For any $\varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{r_n} \mathbb{E} \|X_{nk}\|_{L_p}^2 \mathbf{1}_{\{\|X_{nk}\|_{L_p} > \varepsilon\}} = 0.$$

(c) There is a mean zero Gaussian random element Y in $L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R})$ such that

$$\sum_{k=1}^{r_n} \mathbb{E}(X_{nk} \otimes X_{nk}) \xrightarrow[n \rightarrow \infty]{} \Gamma := \mathbb{E}(Y \otimes Y) \quad \text{in } L_p(\mathbb{S}^2, \mathcal{S} \otimes \mathcal{S}, \mu \otimes \mu; \mathbb{R}).$$

(d) Denoting by σ_n and σ the non-negative elements of $L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R})$ defined by $\sigma_n^2(s) := \sum_{k=1}^{r_n} \mathbb{E} X_{nk}(s)^2$ and $\sigma^2(s) := \mathbb{E} Y(s)^2$, μ -a.e on \mathbb{S} ,

$$\int_{\mathbb{S}} \sigma_n^p d\mu \xrightarrow[n \rightarrow \infty]{} \int_{\mathbb{S}} \sigma^p d\mu.$$

Then

$$\sum_{k=1}^{r_n} X_{nk} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y \quad \text{in the space } L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R}).$$

The proof requires the preliminaries gathered from Lemma 24 to Lemma 27.

Lemma 24 Assume that $f, g \in L_q(\mathbb{S}^2)$ satisfy for all $A, B \in \mathcal{S}$ of finite μ -measure,

$$\int_{A \times B} f d(\mu \otimes \mu) = \int_{A \times B} g d(\mu \otimes \mu). \quad (78)$$

Then $f = g$, $(\mu \otimes \mu)$ -a.e. on \mathbb{S}^2 .

Proof. Let us remark first that the integrals in (78) are well defined because $\mathbf{1}_A$ and $\mathbf{1}_B$ are in $L_q(\mathbb{S})$ since $\mu(A)$ and $\mu(B)$ are finite. We first prove the lemma in the special case where $\mu(\mathbb{S}) < \infty$ and then extend the result to the general case by using the σ -finiteness of μ . To simplify the writing, we denote by $C_n \uparrow C$ the fact that the sequence of sets $(C_n)_{n \geq 1}$ increases to the set C , that is $C_n \subset C_{n+1}$ for every $n \geq 1$ and $\cup_{n \geq 1} C_n = C$.

Case where $\mu(\mathbb{S}) < \infty$. Let us introduce the class \mathcal{L} of sets $C \in \mathcal{S} \otimes \mathcal{S}$ such that f and g are $\mu \otimes \mu$ integrable on C and $\int_C f d(\mu \otimes \mu) = \int_C g d(\mu \otimes \mu)$, together with the class $\mathcal{R} := \{A \times B, A \in \mathcal{S}, B \in \mathcal{S}\}$. As $\mu(\mathbb{S})$ is finite, the same holds for $\mu(A)$ and $\mu(B)$ and (78) gives the inclusion $\mathcal{R} \subset \mathcal{L}$. Clearly, \mathcal{R} is a π -system, i.e. closed under the formation of finite intersections. The class \mathcal{L} satisfies the three following properties.

- (λ_1) \mathbb{S}^2 belongs to \mathcal{L} . Indeed $\mathbf{1}_{\mathbb{S}^2} \in L_q(\mathbb{S}^2)$ because $\mu(\mathbb{S}) < \infty$.
- (λ_2) $C, C' \in \mathcal{L}$ and $C \subset C'$ imply $C' \setminus C \in \mathcal{L}$. This follows easily by writing for $h = f, g$, $\int_{C'} h d(\mu \otimes \mu) = \int_C h d(\mu \otimes \mu) + \int_{C' \setminus C} h d(\mu \otimes \mu)$ and using the membership of C, C' in \mathcal{L} .
- (λ_3) $\{C_n, n \geq 1\} \subset \mathcal{L}$ and $C_n \uparrow C$ imply $C \in \mathcal{L}$. Indeed the equality $\int_{C_n} f d(\mu \otimes \mu) = \int_{C_n} g d(\mu \otimes \mu)$ gives $\int_{C_n} (f^+ + g^-) d(\mu \otimes \mu) = \int_{C_n} (g^+ + f^-) d(\mu \otimes \mu)$ and by B. Levi's monotone convergence theorem, we obtain $\int_C (f^+ + g^-) d(\mu \otimes \mu) = \int_C (g^+ + f^-) d(\mu \otimes \mu)$ that is $\int_C f d(\mu \otimes \mu) = \int_C g d(\mu \otimes \mu)$, so $C \in \mathcal{L}$.

Hence \mathcal{L} is a λ -system. As it contains the π -system \mathcal{R} , by Dynkin's π - λ theorem, see e.g. [5], it contains also the σ -field generated by \mathcal{R} , that is the product $\mathcal{S} \otimes \mathcal{S}$. As \mathcal{L} was defined as a subset of \mathcal{L} it follows that $\mathcal{L} = \mathcal{S} \otimes \mathcal{S}$. In other words

$$\forall C \in \mathcal{S} \otimes \mathcal{S}, \quad \int_C f d(\mu \otimes \mu) = \int_C g d(\mu \otimes \mu). \quad (79)$$

Now with $C = \{f > g\}$, (79) gives $\int_C (f - g) d(\mu \otimes \mu) = 0$. As $f - g$ is positive on C this implies $\mu \otimes \mu(\{f > g\}) = 0$. Similarly one check that $\mu \otimes \mu(\{f < g\}) = 0$, so finally $\mu \otimes \mu(\{f \neq g\}) = 0$, that is $f = g$ $\mu \otimes \mu$ -a.e. on \mathbb{S}^2 .

Case where $\mu(\mathbb{S}) = \infty$. By σ -finiteness of μ , there is a sequence $(\mathbb{S}_n)_{n \geq 1}$ in \mathcal{S} , such that $\mathbb{S}_n \uparrow \mathbb{S}$ and $\mu(\mathbb{S}_n) < \infty$ for each $n \geq 1$. Let us equip \mathbb{S}_n with the σ -field $\mathcal{S}_n := \{A \in \mathcal{S}; A \subset \mathbb{S}_n\} = \{A' \cap \mathbb{S}_n; A' \in \mathcal{S}\}$. Then we can apply the previous case to each measured space $(\mathbb{S}_n, \mathcal{S}_n, \mu)$, $n \geq 1$, which gives $f = g$, $(\mu \otimes \mu)$ -a.e. on \mathbb{S}_n^2 . As $\mathbb{S}^2 = \cup_{n \geq 1} \mathbb{S}_n \times \mathbb{S}_n$, this gives $f = g$, $(\mu \otimes \mu)$ -a.e. on \mathbb{S}^2 . \square

Proposition 25 If X and X' are mean zero random elements in $L_p(\mathbb{S})$ with finite strong second moment and the same covariance operator, then

$$E(X(s)X(t)) = E(X'(s)X'(t)) \quad \text{for } \mu \otimes \mu \text{ almost every } (s, t) \in \mathbb{S}^2.$$

Proof. If X and X' have the same covariance operator, then for all $u, v \in L_q(\mathbb{S}^2)$,

$$\mathbb{E}(\langle X, u \rangle \langle X, v \rangle) = \mathbb{E}(\langle X', u \rangle \langle X', v \rangle).$$

Hölder inequality and Fubini arguments legitimate the rephrasing of this equality as

$$\int_{\mathbb{S}^2} \mathbb{E}(X(s)X(t))u(s)v(t) d\mu \otimes \mu(s, t) = \int_{\mathbb{S}^2} \mathbb{E}(X'(s)X'(t))u(s)v(t) d\mu \otimes \mu(s, t).$$

As for any $A, B \in \mathcal{S}$ such that $\mu(A), \mu(B) < \infty$, the functions $u = \mathbf{1}_A$ and $v = \mathbf{1}_B$ are in $L_q(\mathbb{S})$, Lemma 24 gives the expected conclusion. \square

In what follows we use for notational convenience the indexation by infinite subsets of $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ to denote subsequences. So any (infinite) subsequence of $(u_n)_{n \geq 1}$ can be denoted as $(u_n)_{n \in I}$ with I infinite subset of \mathbb{N}^* and the convergence of this subsequence will be denoted by $\xrightarrow[n \rightarrow \infty, n \in I]$ or $\lim_{n \rightarrow \infty, n \in I}$.

Lemma 26 *Let ξ be a Gaussian random element in $L_p(\mathbb{S}) = L_p(\mathbb{S}, \mathcal{S}, \mu; \mathbb{R})$ having a representation*

$$\xi = \int_{L_p(\mathbb{S})} \text{Id} dW,$$

where W is a white noise. Then for μ -almost every $s \in \mathbb{S}$, $\xi(s)$ is a mean zero Gaussian random variable.

Proof. By construction of the $L_p(\mathbb{S})$ valued stochastic integral with respect to W , there is a sequence of $L_p(\mathbb{S})$ valued simple functions $f_n = \sum_{i=1}^{j_n} h_{ni} \mathbf{1}_{A_{ni}}$ where the h_{ni} are in $L_p(\mathbb{S})$, and for each n , the A_{ni} , $1 \leq i \leq j_n$ are disjoint, such that with $\xi_n := \int_{L_p(\mathbb{S})} f_n dW$, $\mathbb{E} \|\xi_n - \xi\|_p^2 \rightarrow 0$. Let us fix a representant still denoted h_{ni} in each class of functions h_{ni} . Then

$$\xi_n(s) = \sum_{i=1}^{j_n} h_{ni}(s) W(A_{ni})$$

is a Gaussian mean zero random variable as a linear combination of the independent Gaussian mean zero random variables $W(A_{ni})$. Now the conclusion of the Lemma follows if we prove that for μ -almost every $s \in \mathbb{S}$, $\mathbb{E} |\xi_n(s) - \xi(s)|^2 \rightarrow 0$.

Our first step in this way is to prove that the convergence $\mathbb{E} \|\xi_n - \xi\|_p^2 \rightarrow 0$ implies $\mathbb{E} \|\xi_n - \xi\|_p^p \rightarrow 0$ in our Gaussian setting. To this aim we use the following estimates.

$$\begin{aligned} \mathbb{E} \|\xi_n - \xi\|_p^p &\leq \mathbb{E} \left(\|\xi_n - \xi\|_p (\|\xi_n\|_p + \|\xi\|_p)^{p-1} \right) \\ &\leq \left(\mathbb{E} \|\xi_n - \xi\|_p^2 \right)^{1/2} \left(\mathbb{E} (\|\xi_n\|_p + \|\xi\|_p)^{2p-2} \right)^{1/2} \\ &\leq \left(\mathbb{E} \|\xi_n - \xi\|_p^2 \right)^{1/2} \left(2^{2p-3} \mathbb{E} \|\xi_n\|_p^{2p-2} + 2^{2p-3} \mathbb{E} \|\xi\|_p^{2p-2} \right)^{1/2} \end{aligned}$$

As the Gaussian random elements ξ_n and ξ have strong moments of any order, we just have to bound $\mathbb{E} \|\xi_n\|_p^{2p-2}$ uniformly in n . Using (3.5) in [17] we get for $r \geq 2$,

$$\mathbb{E} \|\xi_n\|_p^r = \int_0^\infty r t^{r-1} \mathbf{P}'(\|\xi_n\|_p > t) dt \leq \int_0^\infty 4 r t^{r-1} \exp\left(\frac{-t^2}{8 \mathbb{E} \|\xi_n\|_p^2}\right) dt.$$

Now the convergence to zero of $E \|\xi_n - \xi\|_p^2$ implies the convergence of $E \|\xi_n\|_p^2$ to $E \|\xi\|_p^2$ so there is some n_0 such that $E \|\xi_n\|_p^2 \leq 2E \|\xi\|_p^2$ for every $n \geq n_0$. Hence $\sup_{n \geq n_0} E \|\xi_n\|_p^r \leq \int_0^\infty 4rt^{r-1} \exp(-t^2/(16E \|\xi\|_p^2)) dt < \infty$.

Finally, since

$$E \|\xi_n - \xi\|_p^p = \int_{\mathbb{S}} E |\xi_n(s) - \xi(s)|^p d\mu(s) \xrightarrow{n \rightarrow \infty} 0,$$

we can extract a subsequence $(E |\xi_n(s) - \xi(s)|^p)_{n \in I}$ which converges to zero μ -almost everywhere on \mathbb{S} . So there is some measurable subset \mathbb{S}' such that $\mu(\mathbb{S} \setminus \mathbb{S}') = 0$ and for every $s \in \mathbb{S}'$ $\lim_{n \rightarrow \infty, n \in I} E |\xi_n(s) - \xi(s)|^p = 0$. As $p > 2$, this implies $\lim_{n \rightarrow \infty, n \in I} E |\xi_n(s) - \xi(s)|^2 = 0$. So for every s in \mathbb{S}' , $\xi(s)$ is limit in quadratic mean of the sequence of mean zero Gaussian random variables $(\xi_n(s))_{n \in I}$ hence $\xi(s)$ is a mean zero Gaussian random variable. \square

Lemma 27 *Let X be random element in $L_p(\mathbb{S})$ such that $E \|X\|_{L_p}^2 < \infty$. Let ξ be a Gaussian random element of the form $\xi = \int_{L_p(\mathbb{S})} \text{Id} dW$, where W is a white noise with variance P_X .*

- i) For μ almost every $s \in \mathbb{S}$, $\sigma^2(s) := EX(s)^2 = E\xi(s)^2$.
- ii) Moreover $\sigma \in L_p(\mathbb{S}^2)$.

Proof. To prove i), we recall that the proof of Lemma 26, provides a measurable subset \mathbb{S}' such that $\mu(\mathbb{S} \setminus \mathbb{S}') = 0$ and an infinite subset I of \mathbb{N}^* such that for every $s \in \mathbb{S}'$ $(E |\xi_n(s) - \xi(s)|^p)_{n \in I}$ converges to zero and $(E \xi_n(s)^2)_{n \in I}$ converges to $E \xi(s)^2$. So it suffices to prove that one can extract a subsequence $(E \xi_n(s)^2)_{n \in J}$ for some infinite subset J of I , converging to $EX(s)^2$ for μ almost every $s \in \mathbb{S}'$. Moreover it is enough to prove i) in the case where $\mu(\mathbb{S}) < \infty$. Indeed when $\mu(\mathbb{S}) = \infty$, by σ -finiteness of μ , there is a sequence $(\mathbb{S}_n)_{n \geq 1}$ in \mathcal{S} , such that $\mathbb{S}_n \uparrow \mathbb{S}$ and $\mu(\mathbb{S}_n) < \infty$ for each $n \geq 1$ and the same holds with $(\mathbb{S}'_n)_{n \geq 1}$ and \mathbb{S}' , where $\mathbb{S}'_n := \mathbb{S} \cap \mathbb{S}'_n$. Then clearly if $EX(s)^2 = E\xi(s)^2$ μ -a.e. in \mathbb{S}'_n , the same equality holds μ -a.e. on \mathbb{S}' . So let us assume from now on, that $\mu(\mathbb{S})$ is finite.

Now we note that ξ_n was defined as $\xi_n := \int_{L_p(\mathbb{S})} f_n dW$, with

$$f_n \xrightarrow{n \rightarrow \infty} \text{Id}_{L_p(\mathbb{S})}, \quad \text{in the space } L^2(L_p(\mathbb{S}), \mathcal{B}_{L_p(\mathbb{S})}, P_X; L_p(\mathbb{S})).$$

This convergence means that

$$\int_{L_p(\mathbb{S})} \left\| f_n(x) - \text{Id}_{L_p(\mathbb{S})}(x) \right\|_{L_p(\mathbb{S})}^2 dP_X \xrightarrow{n \rightarrow \infty} 0,$$

which can be reformulated as

$$E \|f_n(X) - X\|_{L_p(\mathbb{S})}^2 = E \left(\int_{\mathbb{S}} |f_n(X(s)) - X(s)|^p d\mu(s) \right)^{2/p} \xrightarrow{n \rightarrow \infty} 0. \quad (80)$$

Since $\mu(\mathbb{S})$ is finite, $\mu/\mu(\mathbb{S})$ is a probability, whence as $p > 2$, for any $g \in L_p(\mathbb{S})$,

$$\left(\int_{\mathbb{S}} |g|^p d\mu \right)^{1/p} \geq \mu(\mathbb{S})^{1/p-1/2} \left(\int_{\mathbb{S}} |g|^2 d\mu \right)^{1/2}.$$

This enables us to deduce from (80) that

$$\mathbb{E} \int_{\mathbb{S}} |f_n(X(s)) - X(s)|^2 d\mu(s) = \int_{\mathbb{S}'} \mathbb{E} |f_n(X(s)) - X(s)|^2 d\mu(s) \xrightarrow{n \rightarrow \infty, n \in I} 0$$

Then there is a measurable subset \mathbb{S}'' of \mathbb{S}' such that $\mu(\mathbb{S}' \setminus \mathbb{S}'') = 0$ together with a subsequence $(\mathbb{E} |f_n(X(s)) - X(s)|^2)_{n \in J}$, with $J \subset I$, converging to zero for every $s \in \mathbb{S}''$. Now we have for every $s \in \mathbb{S}''$, $\lim_{n \rightarrow \infty, n \in J} \mathbb{E} f_n(X(s))^2 = \mathbb{E} X(s)^2$. As $f_n(X) = \sum_{i=1}^{j_n} x_{ni} \mathbf{1}_{A_{ni}}(X)$, with the A_{ni} pairwise disjoint, $\mathbb{E} f_n(X(s))^2 = \sum_{i=1}^{j_n} x_{ni}(s)^2 P_X(A_{ni})$. On the other hand, $\xi_n(s) = I_W(f_n)(s) = \sum_{i=1}^{j_n} x_{ni}(s) W(A_{ni})$ where the $W(A_{ni})$ are independent centered Gaussian random variables with respective variances $P_X(A_{ni})$, $\mathbb{E} \xi_n(s)^2 = \sum_{i=1}^{j_n} x_{ni}(s)^2 P_X(A_{ni})$. So $\mathbb{E} f_n(X(s))^2 = \mathbb{E} \xi_n(s)^2$ μ -a.e. on \mathbb{S}'' . Finally,

$$\mathbb{E} X(s)^2 = \lim_{n \rightarrow \infty, n \in J} \mathbb{E} f_n(X(s))^2 = \lim_{n \rightarrow \infty, n \in J} \mathbb{E} \xi_n(s)^2 = \mathbb{E} \xi(s)^2, \quad \mu\text{-a.e. on } \mathbb{S}'',$$

which completes the proof of i).

To check ii), by combining i) and Lemma 26, one see that for μ almost every $s \in \mathbb{S}$, $\sigma^2(s) = \mathbb{E} \xi(s)^2$ and $\xi(s)$ is a mean zero Gaussian random variable. For every such s , $\sigma(s) = (\mathbb{E} \xi(s)^2)^{1/2} \leq (\mathbb{E} |\xi(s)|^p)^{1/p}$ since $p > 2$, whence $\sigma(s)^p \leq \mathbb{E} |\xi(s)|^p$. Therefore

$$\int_{\mathbb{S}} \sigma^p d\mu \leq \int_{\mathbb{S}} \mathbb{E} |\xi|^p d\mu = \mathbb{E} \|\xi\|_{L_p}^p < \infty,$$

since the Gaussian random element ξ in $L_p(\mathbb{S})$ has finite moments of any order. \square

Proof of Th. 23. Putting $2 + \delta = \min(p, 3)$ and applying Th.21, we deduce from (a) and (b) that $\lim_{n \rightarrow \infty} \zeta_{2+\delta}(\sum_{k=1}^{r_n} X_{nk}, \sum_{k=1}^{r_n} Y_{nk}) = 0$, where the Y_{nk} are chosen as in the proof of Th.21, that is $Y_{nk} = \int_{L_p(\mathbb{S})} \text{Id}_{L_p(\mathbb{S})} dW_{nk}$, where the W_{nk} are independent white noises with respective variances $P_{X_{nk}}$. So it remains to prove that $\zeta_{2+\delta}(\sum_{k=1}^{r_n} Y_{nk}, Y)$ converges to zero, which is equivalent to $\sum_{k=1}^{r_n} Y_{nk} \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y$ in the space $L_p(\mathbb{S})$. This last convergence will be established by proving that

- (i) For every $u \in L_q(\mathbb{S}, \mathcal{S}, \mu, \mathbb{R})$, $\langle \sum_{k=1}^{r_n} Y_{nk}, u \rangle$ converges in distribution to $\langle Y, u \rangle$.
- (ii) The sequence $(\sum_{k=1}^{r_n} Y_{nk})_{n \geq 1}$ is tight in $L_p(\mathbb{S})$.

To prove (i), we remark that $\langle \sum_{k=1}^{r_n} Y_{nk}, u \rangle$ and $\langle Y, u \rangle$ are mean zero Gaussian random variables, so the announced convergence in distribution will follow from convergence of their variances. To prove this last one, we note that

$$\text{var}\langle Y_{nk}, u \rangle = \mathbb{E}\langle Y_{nk}, u \rangle^2 = \int_{\mathbb{S}^2} u(s)u(t) \mathbb{E} Y_{nk}(s) Y_{nk}(t) d\mu \otimes \mu(s, t),$$

so using the independence of the Gaussian random variables $\langle Y_{nk}, u \rangle$, $1 \leq k \leq r_n$, we just have to prove that

$$\int_{\mathbb{S}^2} u(s)u(t) \sum_{k=1}^{r_n} \mathbb{E} Y_{nk}(s) Y_{nk}(t) d\mu \otimes \mu(s, t) \xrightarrow{n \rightarrow \infty} \int_{\mathbb{S}^2} u(s)u(t) \mathbb{E} Y(s) Y(t) d\mu \otimes \mu(s, t).$$

By Proposition 25, we can replace the X_{nk} 's by the Y_{nk} 's in the above convergence which then appears as an obvious consequence of Assumption (c) since $u \otimes u$ belongs to $L_q(\mathbb{S}^2)$.

To prove (ii), according to Cremers and Kadelka [9, Th.2], it suffices to prove that with $Y_n := \sum_{k=1}^n Y_{nk}$,

$$\limsup_{n \rightarrow \infty} \int_{\mathbb{S}} \mathbb{E} |Y_n|^p \, d\mu \leq \int_{\mathbb{S}} \mathbb{E} |Y|^p \, d\mu.$$

In fact, we will prove that

$$\lim_{n \rightarrow \infty} \int_{\mathbb{S}} \mathbb{E} |Y_n|^p \, d\mu = \int_{\mathbb{S}} \mathbb{E} |Y|^p \, d\mu. \quad (81)$$

By Lemma 26 and Lemma 27 i), for μ almost every $s \in \mathbb{S}$, $Y_n(s)$ and $Y(s)$ are centered Gaussian random variables with respective variances $\sigma_n^2(s) = \sum_{k=1}^n \mathbb{E} X_{nk}(s)^2$ and $\sigma^2(s)$. This implies that

$$\mathbb{E} |Y_n(s)|^p = \sigma_n(s)^p m_p, \quad \mathbb{E} |Y(s)|^p = \sigma(s)^p m_p$$

where $m_p := (2\pi)^{-1/2} \int_{-\infty}^{\infty} |z|^p \exp(-z^2/2) \, dz$. This way, (81) is reduced to Assumption (d) and the proof is complete. \square

5 Asymptotic normality of weighted sums

Let $(X_j, j \in \mathbb{Z})$ be a set of \mathbb{B} -valued random elements. Assume that $\mathbb{E}(X_j) = 0$ and $\mathbb{E} \|X_j\|^2 < \infty$ for any $j \in \mathbb{Z}$. Consider the weighted sums

$$Z_n := \sum_{k=0}^{\infty} a_{n,k} X_k, \quad n \in \mathbb{N}, \quad (82)$$

whenever they are correctly defined, where $\{(a_{n,k}, k \geq 0), n \in \mathbb{N}\} \subset \mathbb{R}$. We assume that for each $n \in \mathbb{N}$, $\sum_k a_{nk}^2 < \infty$.

Theorem 28 *Assume that the Banach space \mathbb{B} is p -smooth with some $p > 2$. Let $(X_k, k \in \mathbb{Z})$ be i.i.d. \mathbb{B} -valued random elements and $Q := \text{cov}(X_i)$. Assume that*

- (i) $c_n := \sup_{k \geq 0} |a_{nk}| \rightarrow 0$ as $n \rightarrow \infty$;
- (ii) $b_n^2 := \sum_{k \geq 0} a_{nk}^2 \rightarrow 1$ as $n \rightarrow \infty$.

Then, for each n , the series $\sum_k a_{nk} X_k$ converges a.s., and

$$\sum_{k=0}^{\infty} a_{nk} X_k \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y_Q. \quad (83)$$

Proof. Since the space B is of type 2, we get by (40) for any $1 \leq l \leq m$,

$$\begin{aligned} \mathbb{E} \left\| \sum_{k=l}^m a_{nk} X_k \right\| &\leq K \left(\sum_{k=l}^m \mathbb{E} \|a_{nk} X_k\|^2 \right)^{1/2} \\ &\leq K \left(\sum_{k=l}^m a_{nk}^2 \right)^{1/2} (\mathbb{E} \|X\|^2)^{1/2}. \end{aligned}$$

Due to Condition (ii), this implies that the series $\sum_k a_{nk} X_k$ satisfies Cauchy's criterion in the space $L^1(\Omega, \mathcal{F}, P; B)$, hence converges in L^1 and in probability. By independence of its terms, it converges also a.s., according to Ito-Nisio theorem.

Without loss of generality we assume that $p = 2 + \delta$ with some $\delta \in (0, 1)$. Let $m \geq 1$. We apply Theorem 22 for $X_{nk} = a_{nk} X_k$, $k = 0, \dots, m$. In this case the sum $\sum_{k=0}^m Y_{nk}$ has the same distribution as $A_{nm} Y_Q$, where $A_{nm}^2 = \sum_{k=0}^m a_{nk}^2$, $A_{nm} \geq 0$. Hence, by (66) in Th.21, with $c := c(B, \delta)$,

$$\begin{aligned} \zeta_p \left(\sum_{k=0}^m a_{nk} X_k, A_{nm} Y_Q \right) &\leq c \left[\varepsilon^\delta A_{nm}^2 \mathbb{E} \|X_1\|^2 + \right. \\ &\quad \left. (1 + A_{nm}^2 \mathbb{E} \|X_1\|^2) \sum_{k=0}^m a_{nk}^2 \mathbb{E} \|X_1\|^2 \mathbf{1}_{\{|a_{nk}| \|X_1\| > \varepsilon\}} \right]. \end{aligned}$$

This gives the following bound uniform in $m \geq 0$:

$$\begin{aligned} \zeta_p \left(\sum_{k=0}^m a_{nk} X_k, A_{nm} Y_Q \right) &\leq c b_n^2 \left[\varepsilon^\delta \mathbb{E} \|X_1\|^2 + \right. \\ &\quad \left. (1 + b_n^2 \mathbb{E} \|X_1\|^2) \mathbb{E} \|X_1\|^2 \mathbf{1}_{\{c_n \|X_1\| > \varepsilon\}} \right]. \quad (84) \end{aligned}$$

Now we estimate

$$\zeta_p(A_{nm} Y_Q, Y_Q) = \sup \left\{ |\mathbb{E} f(A_{nm} Y_Q) - \mathbb{E} f(Y_Q)| : f \in C_b^{(p)}(B), \|f\|_p \leq 1 \right\}.$$

As the random elements $A_{nm} Y_Q$ and Y_Q are defined on *the same* probability space, the writing $\mathbb{E} (f(A_{nm} Y_Q) - f(Y_Q))$ makes sense and we obtain

$$\begin{aligned} |\mathbb{E} f(A_{nm} Y_Q) - \mathbb{E} f(Y_Q)| &= |\mathbb{E} (f(A_{nm} Y_Q) - f(Y_Q))| \leq \mathbb{E} |f(A_{nm} Y_Q) - f(Y_Q)| \\ &\leq \mathbb{E} \sup_{x \in B} \|f'(x)\| \|A_{nm} Y_Q - Y_Q\| \\ &\leq |1 - A_{nm}| \mathbb{E} \|Y_Q\|, \end{aligned}$$

because $\|f\|_p \leq 1$ implies $\sup_{x \in B} \|f'(x)\| \leq 1$. Therefore

$$\zeta_p(A_{nm} Y_Q, Y_Q) \leq |1 - A_{nm}| \mathbb{E} \|Y_Q\|, \quad m \geq 0. \quad (85)$$

Next, by the regularity property of ζ_p , see (b) p.16 and the independence of the X_k 's,

$$\zeta_p \left(\sum_{k=0}^m a_{nk} X_k, \sum_{k=0}^{\infty} a_{nk} X_k \right) \leq \zeta_p \left(0, \sum_{k=m+1}^{\infty} a_{nk} X_k \right).$$

By Taylor formula at the order 1 with integral remainder, see (68), it is easily seen that if Z is a random element in \mathbb{B} and $\mathbb{E}Z = 0$, $\mathbb{E}\|Z\|^2 < \infty$, then $\zeta_p(0, Z) \leq \frac{1}{2} \mathbb{E}\|Z\|^2$. Therefore,

$$\zeta_p \left(\sum_{k=0}^m a_{nk} X_k, \sum_{k=0}^{\infty} a_{nk} X_k \right) \leq \frac{K^2}{2} \mathbb{E}\|X_1\|^2 \sum_{k=m+1}^{\infty} a_{nk}^2. \quad (86)$$

Finally, by triangular inequality for the distance ζ_p , gathering the estimates (84), (85) and (86) gives

$$\zeta_p \left(\sum_{k=0}^{\infty} a_{nk} X_k, Y_Q \right) \leq \frac{K^2}{2} \mathbb{E}\|X_1\|^2 \sum_{k=m+1}^{\infty} a_{nk}^2 + u_n(\varepsilon) + |1 - A_{nm}| \|Y_Q\|,$$

where $u_n(\varepsilon)$ denotes the right-hand side of (84). Using Assumption (ii), letting m tend to infinity in the above inequality gives $\zeta_p(\sum_{k=0}^{\infty} a_{nk} X_k, Y_Q) \leq u_n(\varepsilon)$, whence by (i) and (ii),

$$\limsup_{n \rightarrow \infty} \zeta_p \left(\sum_{k=0}^{\infty} a_{nk} X_k, Y_Q \right) \leq c \mathbb{E}\|X_1\|^2 \varepsilon^\delta.$$

By arbitrariness of ε , we conclude that $\lim_{n \rightarrow \infty} \zeta_p(\sum_{k=0}^{\infty} a_{nk} X_k, Y_Q) = 0$. \square

Next consider a \mathbb{B} -valued linear process $(X_k, k \in \mathbb{Z})$ defined by

$$X_k = \sum_{j=0}^{\infty} \psi_j \varepsilon_{k-j}, \quad (87)$$

where innovations $(\varepsilon_k, k \in \mathbb{Z})$ are i.i.d. \mathbb{B} -valued random variables such that $\mathbb{E}\varepsilon_0 = 0$, $Q_\varepsilon = \text{cov}(\varepsilon_0)$, $0 < \sigma^2 := \mathbb{E}\|\varepsilon_k\|^2 < \infty$ and the linear filter $(\psi_j, j \geq 0) \subset L(\mathbb{B})$ is a sequence of linear bounded operators such that $\psi_0 = \text{Id}_{\mathbb{B}}$ and

$$\sum_{j=0}^{\infty} \|\psi_j\| < \infty. \quad (88)$$

This condition ensures the a.s. convergence of the series in (87). In this case we set $\Psi = \sum_{j=0}^{\infty} \psi_j$.

Theorem 29 *Let \mathbb{B} be a p -smooth Banach space, $p > 2$. Let (X_k) be a linear process defined by (87), where (ψ_k) satisfies (88). Let $(a_{n,j}, j \in \mathbb{Z}, n \in \mathbb{N}) \subset \mathbb{R}$ satisfy conditions (i)–(ii) of Theorem 28 and*

$$(iii) \lim_{n \rightarrow \infty} \sum_{k \in \mathbb{Z}} (a_{n,k+1} - a_{n,k})^2 = 0.$$

Then,

$$\sum_{k=0}^{\infty} a_{nk} X_k \xrightarrow[n \rightarrow \infty]{\mathcal{D}} Y_{\Psi^* \tilde{Q}_\varepsilon \Psi}.$$

Proof. We have

$$Z_n := \sum_{k=0}^{\infty} a_{nk} X_k = \sum_{k=0}^{\infty} a_{nk} \sum_{j=0}^{\infty} \psi_j \varepsilon_{k-j} = \sum_{j=0}^{\infty} \psi_j \sum_{k=0}^{\infty} a_{nk} \varepsilon_{k-j} = \sum_{j=0}^{\infty} \psi_j Z_{nj},$$

where $Z_{nj} := \sum_{k=0}^{\infty} a_{nk} \varepsilon_{k-j}$. Writing $Z_n = Z'_n + Z''_n$, where

$$Z'_n := \sum_{j=0}^{\infty} \psi_j (Z_{nj} - Z_{n0}), \quad Z''_n := \sum_{j=0}^{\infty} \psi_j Z_{n0}$$

we consider each Z'_n and Z''_n separately. By Theorem 28,

$$Z''_n = \Psi(Z_{n0}) = \Psi\left(\sum_{k=0}^{\infty} a_{nk} \varepsilon_k\right) \xrightarrow[n \rightarrow \infty]{\mathcal{D}} \Psi(Y_{Q_\varepsilon}) \sim Y_{\Psi^* Q_\varepsilon} \Psi.$$

To complete the proof we show that $Z'_n \xrightarrow[n \rightarrow \infty]{P} 0$. To this aim let us assume for a moment that the following two properties hold true:

$$\sup_{n,j} \mathbb{E} \|Z_{nj}\|^2 < \infty \quad (89)$$

and

$$\lim_{n \rightarrow \infty} \|Z_{nj} - Z_{n0}\| = 0 \quad \text{in probability, for each } j \in \mathbb{N}. \quad (90)$$

Let $\varepsilon > 0$ and $J \in \mathbb{N}$. Splitting Z'_n in two sums indexed by $j \leq J$ and $j > J$, leads to

$$\mathbb{P}(\|Z'_n\| > \varepsilon) \leq \mathbb{P}\left(\sum_{j=0}^J \|\psi_j\| \cdot \|Z_{nj} - Z_{n0}\| > \frac{\varepsilon}{2}\right) + \mathbb{P}\left(\sum_{j>J} \|\psi_j\| \cdot \|Z_{nj} - Z_{n0}\| > \frac{\varepsilon}{2}\right). \quad (91)$$

Applying Markov inequality at order one gives

$$\mathbb{P}\left(\sum_{j>J} \|\psi_j\| \cdot \|Z_{nj} - Z_{n0}\| > \frac{\varepsilon}{2}\right) \leq \frac{4}{\varepsilon} \sup_{n,j} \mathbb{E} \|Z_{nj}\| \sum_{j>J} \|\psi_j\|.$$

By (89) and (88), taking $J \in \mathbb{N}$ large enough, one can make the right side of the preceding bound as small as one wish. Then the first probability on the right side of (91) is small as one wish by (90) and taking $n \in \mathbb{N}_+$ large enough. Therefore $Z'_n \xrightarrow[n \rightarrow \infty]{P} 0$ holds true subject to the forthcoming proof of (89) and (90).

For (89), as the sequence $(\varepsilon_i)_{i \in \mathbb{Z}}$ is i.i.d., it is clear that for each $n \geq 1$, all the Z_{nj} have the same distribution, so it suffices to check that $\sup_{n \geq 1} \mathbb{E} \|Z_{n0}\|^2 < \infty$. As B is of type 2, it easily follows from (39) and the equidistribution of the independent ε_k that

$$\mathbb{E} \|Z_{n0}\|^2 \leq K^2 \mathbb{E} \|\varepsilon_0\|^2 \sum_{k=0}^{\infty} a_{nk}^2.$$

Hence (89) results from Assumption (ii).

To check (90), we show that in the decomposition

$$Z_{nj} - Z_{n0} = \sum_{i=-j}^{-1} a_{n,i+j} \varepsilon_i + \sum_{i=0}^{\infty} (a_{n,i+j} - a_{n,i}) \varepsilon_i,$$

both sums converge to zero in quadratic mean. For the first one,

$$\mathbb{E} \left\| \sum_{i=-j}^{-1} a_{n,i+j} \varepsilon_i \right\|^2 \leq K^2 \sum_{i=-j}^{-1} \mathbb{E} \|a_{n,i+j} \varepsilon_i\|^2 = K^2 \sum_{i=-j}^{-1} a_{n,i+j}^2 \mathbb{E} \|\varepsilon_0\|^2 \leq K^2 J \mathbb{E} \|\varepsilon_0\|^2 c_n^2,$$

which tends to zero as n goes to infinity by Assumption (i). For the second sum,

$$\mathbb{E} \left\| \sum_{i=0}^{\infty} (a_{n,i+j} - a_{n,i}) \varepsilon_i \right\|^2 \leq K^2 \mathbb{E} \|\varepsilon_0\|^2 \sum_{i=0}^{\infty} (a_{n,i+j} - a_{n,i})^2.$$

Let us denote by d the Euclidean distance in the sequence space $\ell^2(\mathbb{N})$. Then

$$\begin{aligned} \sum_{i=0}^{\infty} (a_{n,i+j} - a_{n,i})^2 &\leq \left(\sum_{l=0}^{j-1} d \left((a_{n,i+l+1})_{i \geq 0}, (a_{n,i+l})_{i \geq 0} \right) \right)^2 \\ &\leq j^2 \max_{0 \leq l < j} d \left((a_{n,i+l+1})_{i \geq 0}, (a_{n,i+l})_{i \geq 0} \right)^2 \\ &\leq j^2 \sum_{k=0}^{\infty} (a_{n,k+1} - a_{n,k})^2, \end{aligned}$$

which tends to zero as n tends to infinity by Assumption (iii). Hence (90) is established and the proof complete. \square

Examples of summation methods $(a_{n,k}, k \in \mathbb{N}, n \in \mathbb{N})$ that satisfy conditions (i)–(iii) include the following (see [10] for more examples):

Cesàro summation corresponding to

$$a_{n,k} = (n+1)^{-1/2} \begin{cases} 1, & \text{if } 0 \leq k \leq n, \\ 0, & \text{otherwise.} \end{cases}$$

Here, $c_n = (n+1)^{-1/2}$, $b_n^2 = 1$ and $\sum_{k=0}^{\infty} (a_{n,k+1} - a_{n,k})^2 = 1/(n+1)$.

Abel summation corresponding to

$$a_{n,k} = \sqrt{2\lambda_n} (1 - e^{-1/\lambda_n}) e^{-k/\lambda_n}, \quad k \geq 0,$$

where $\lambda_n \rightarrow \infty$ as $n \rightarrow \infty$. In this case,

$$c_n = \sqrt{2\lambda_n} (1 - e^{-1/\lambda_n}) \sim \left(\frac{2}{\lambda_n} \right)^{1/2} \xrightarrow{n \rightarrow \infty} 0,$$

$$b_n^2 = 2\lambda_n (1 - e^{-1/\lambda_n})^2 \sum_{k=0}^{\infty} e^{-2k/\lambda_n} = \frac{2\lambda_n (1 - e^{-1/\lambda_n})^2}{1 - e^{-2/\lambda_n}} = \frac{2\lambda_n (1 - e^{-1/\lambda_n})}{1 + e^{-1/\lambda_n}} \xrightarrow{n \rightarrow \infty} 1,$$

$$\begin{aligned} \sum_{k=0}^{\infty} (a_{n,k+1} - a_{n,k})^2 &= 2\lambda_n (1 - e^{-1/\lambda_n})^2 \sum_{k=0}^{\infty} \left(e^{-(k+1)/\lambda_n} - e^{-k/\lambda_n} \right)^2 \\ &= \frac{2\lambda_n (1 - e^{-1/\lambda_n})^4}{1 - e^{-2/\lambda_n}} \sim \frac{1}{\lambda_n^2} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Borel summation corresponding to

$$a_{n,k} = \sqrt{2}(\pi\lambda_n)^{1/4} \frac{e^{-\lambda_n} \lambda_n^k}{k!}, \quad k \geq 0, \quad \lambda_n \xrightarrow{n \rightarrow \infty} \infty.$$

To check (i), recalling that $\max_{k \geq 0} P(N = k)$ where the random variable N has the Poisson distribution with parameter λ_n is $P(N = m)$ with $m \leq \lambda_n < m + 1$, we get

$$c_n = \sqrt{2}(\pi\lambda_n)^{1/4} \frac{e^{-\lambda_n} \lambda_n^m}{m!}.$$

By Stirling formula, $m! = \sqrt{2\pi m} m^{m+1/2} e^{-m} (1 + \delta_m)$ with $\lim_{m \rightarrow \infty} \delta_m = 0$, whence

$$c_n \leq \pi^{-1/4} \frac{(m+1)^{1/4} e^{-m} (m+1)^m}{m^{m+1/2} e^{-m} (1 + \delta_m)} = \frac{\left(1 + \frac{1}{m}\right)^{m+1/4}}{(\pi m)^{1/4} (1 + \delta_m)} \sim \frac{e}{(\pi m)^{1/4}} \xrightarrow{n \rightarrow \infty} 0.$$

To check (ii), we refer to [10] where it is proved by using the Bessel function of the first kind, see (2.8) and (2.9) therein.

To check (iii), we note first that

$$\sum_{k=0}^{\infty} (a_{n,k+1} - a_{n,k})^2 = \sum_{k=0}^{\infty} 2\sqrt{\pi\lambda_n} e^{-2\lambda_n} \frac{\lambda_n^{2k}}{(k!)^2} \left(\frac{\lambda_n}{k+1} - 1 \right)^2 = \int_{\mathbb{N}} f_{\lambda_n} d\mu_{\lambda_n},$$

where $f_{\lambda}(k) := (\lambda(k+1)^{-1} - 1)^2$ and μ_{λ} is the discrete measure

$$\mu_{\lambda} := \sum_{k=0}^{\infty} \left(2\sqrt{\pi\lambda} e^{-2\lambda} \frac{\lambda^{2k}}{(k!)^2} \right) \delta_k.$$

In what follows we simplify the notations by replacing λ_n ($n \rightarrow \infty$) by λ ($\lambda \rightarrow \infty$). It is easily seen that the peak of the point masses of μ_{λ} is at $k = [\lambda]$. We will use the following estimates for the left and right tails of μ_{λ} , obtained by comparison with geometric sums.

$$\sum_{k \leq j} \frac{\lambda^{2k}}{(k!)^2} < \frac{\lambda^{2j}}{(j!)^2} \frac{\lambda^2}{\lambda^2 - j^2}, \quad 0 \leq j < \lambda, \quad (92)$$

$$\sum_{k \geq j} \frac{\lambda^{2k}}{(k!)^2} < \frac{\lambda^{2j}}{(j!)^2} \frac{(j+1)^2}{(j+1)^2 - \lambda^2}, \quad j > \lambda. \quad (93)$$

Let $1/2 < \tau < 1$. We split $\int_{\mathbb{N}}$ in $\int_L + \int_C + \int_R$ with left, center and right intervals $L := \mathbb{N} \cap [0, \lambda - \lambda^{\tau}]$, $C := \mathbb{N} \cap (\lambda - \lambda^{\tau}, \lambda + \lambda^{\tau})$, $R := \mathbb{N} \cap [\lambda + \lambda^{\tau}, \infty)$.

Estimation of $\int_L f_\lambda d\mu_\lambda$. Let j be the unique integer such that $j \leq \lambda - \lambda^\tau < j+1$. As $f_\lambda(k) \leq \lambda^2$ on L and accounting (92),

$$\int_L f_\lambda d\mu_\lambda \leq 2\sqrt{\pi}\lambda^{5/2} \frac{e^{-2\lambda}\lambda^{2j}}{(j!)^2} \frac{\lambda^2}{\lambda^2 - j^2}.$$

We note that

$$\frac{\lambda^2}{\lambda^2 - j^2} \leq \frac{\lambda^2}{\lambda^2 - (\lambda - \lambda^\tau)^2} = \frac{1}{1 - (1 - \lambda^{\tau-1})^2} \sim \frac{1}{2}\lambda^{1-\tau}.$$

By Stirling formula, $((j+1)!)^2 = 2\pi(j+1)^{2(j+1)+1}e^{-2(j+1)}(1+\delta_{j+1})^2$, so as $j+1 > \lambda - \lambda^\tau$, $((j+1)!)^2 \geq 2\pi(\lambda - \lambda^\tau)^{2\lambda - 2\lambda^\tau + 1}e^{-2\lambda + 2\lambda^\tau - 2}(1+\delta_{j+1})^2$, whence

$$\frac{e^{-2\lambda}\lambda^{2j}}{(j!)^2} = \frac{e^{-2\lambda}\lambda^{2j}(j+1)^2}{((j+1)!)^2} \leq \frac{e^2}{2\pi} \frac{e^{-2\lambda^\tau}\lambda^{2\lambda - 2\lambda^\tau}(\lambda - \lambda^\tau + 1)^2}{(\lambda - \lambda^\tau)^{2\lambda - 2\lambda^\tau + 1}(1+\delta_{j+1})^2} \sim \frac{e^2}{2\pi}T_1(\lambda),$$

where $T_1(\lambda) := \lambda e^{-2\lambda^\tau}(1 - \lambda^{\tau-1})^{2\lambda^\tau - 2\lambda}$. Next, using $\ln(1-t) = -t - t^2/2 + o(t^2)$ as $t \rightarrow 0$,

$$\begin{aligned} T_1(\lambda) &= \lambda \exp\left(-2\lambda^\tau + (2\lambda^\tau - 2\lambda)\left(-\lambda^{\tau-1} - \frac{\lambda^{2\tau-2}}{2} + o(\lambda^{2\tau-2})\right)\right) \\ &= \lambda \exp\left(-\lambda^{2\tau-1} + o(\lambda^{2\tau-1})\right) \end{aligned}$$

since $\lambda^{3\tau-2} = o(\lambda^{2\tau-1})$. As $2\tau - 1 > 0$, for any $a > 0$ and $c \in (0, 1)$, $\lambda^a T_1(\lambda) = O(\exp(-c\lambda^{2\tau-1}))$, whence

$$\int_L f_\lambda d\mu_\lambda = O(\exp(-c\lambda^{2\tau-1})) = o(1).$$

Estimation of $\int_C f_\lambda d\mu_\lambda$. One easily check that for $k \in C$, $f_\lambda(k) \leq \left(\frac{\lambda^\tau}{\lambda - \lambda^\tau}\right)^2 \sim \lambda^{2\tau-2}$. As $\mu_\lambda(C) < \mu_\lambda(\mathbb{N}) \sim 1$, this gives

$$\int_C f_\lambda d\mu_\lambda = O(\lambda^{2\tau-2}) = o(1).$$

Estimation of $\int_R f_\lambda d\mu_\lambda$. For $k \geq \lambda + \lambda^\tau$, $0 < \lambda(k+1)^{-1} < 1$ whence $f_\lambda(k) < 1$ so

$$\int_R f_\lambda d\mu_\lambda \leq 2\sqrt{\pi} \sum_{k > \lambda + \lambda^\tau} \lambda^{1/2} e^{-2\lambda} \frac{\lambda^{2k}}{(k!)^2}.$$

Denoting by j the unique integer such that $j \leq \lambda + \lambda^\tau < j+1$, (93) gives

$$\sum_{k > \lambda + \lambda^\tau} \lambda^{1/2} e^{-2\lambda} \frac{\lambda^{2k}}{(k!)^2} \leq \lambda^{1/2} \frac{e^{-2\lambda}\lambda^{2j}}{(j!)^2} \frac{(j+1)^2}{(j+1)^2 - \lambda^2}.$$

By increasingness of the function $s \mapsto s/(s-a)$ on (a, ∞) , the last factor is estimated as

$$\frac{(j+1)^2}{(j+1)^2 - \lambda^2} \leq \frac{(\lambda + \lambda^\tau)^2}{(\lambda + \lambda^\tau)^2 - \lambda^2} = \frac{\lambda^2(1 + \lambda^{\tau-1})^2}{2\lambda^{1+\tau} + \lambda^{2\tau}} \sim \frac{\lambda^{1-\tau}}{2}.$$

By Stirling formula, $((j+1)!)^2 \geq 2\pi(\lambda + \lambda^\tau)^{2\lambda+2\lambda^\tau+1} e^{-2\lambda-2\lambda^\tau-2}(1 + \delta_{j+1})^2$, so

$$\frac{e^{-2\lambda} \lambda^{2j}}{(j!)^2} = \frac{\lambda^{2j} (j+1)^2}{((j+1)!)^2} \leq \frac{e^2}{2\pi} \frac{e^{2\lambda^\tau} \lambda^{2\lambda+2\lambda^\tau} (\lambda + \lambda^\tau + 1)^2}{(\lambda + \lambda^\tau)^{2\lambda+2\lambda^\tau} (\lambda + \lambda^\tau) (1 + \delta_{j+1})^2} \sim \frac{e^2}{2\pi} T_2(\lambda),$$

where $T_2(\lambda) := \lambda e^{2\lambda^\tau} (1 + \lambda^{\tau-1})^{-2\lambda-2\lambda^\tau}$. Since $\ln(1+t) \geq t - t^2/2$ for $t \geq 0$,

$$T_2(\lambda) \leq \lambda \exp\left(2\lambda^\tau - (2\lambda + 2\lambda^\tau) \left(\lambda^{\tau-1} - \frac{\lambda^{2\tau-2}}{2}\right)\right) = \lambda \exp(-\lambda^{2\tau-1} + \lambda^{3\tau-2}).$$

As $2\tau - 1 > 0$, for any $a > 0$ and $c \in (0, 1)$, $\lambda^a T_2(\lambda) = O(\exp(-c\lambda^{2\tau-1}))$, whence

$$\int_{\mathbb{R}} f_\lambda d\mu_\lambda = O(\exp(-c\lambda^{2\tau-1})) = o(1).$$

Gathering all estimates gives $\int_{\mathbb{N}} f_\lambda d\mu_\lambda = O(\lambda^{2\tau-2}) = o(1)$, concluding the check of (iii).

Funding

The research of Alfredas Račkauskas is supported by the Research Council of Lithuania, grant No. S-MIP-17-76.

Conflict of interest

The authors declare that they have no conflict of interest.

Availability of data and material

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Code availability

Not applicable.

References

1. Banna, M.: Limiting spectral distribution of Gram matrices associated with functionals of mixing processes. *Journal of Mathematical Analysis and Applications* **433**(1), 416–433 (2016)
2. Banna, M., Merlevède, F.: Limiting spectral distribution of large sample covariance matrices associated with a class of stationary processes. *J. Theoret. Probab.* **28**, 745–783 (2015)
3. Banna, M., Merlevède, F., Peligrad, M.: On the limiting spectral distribution for a large class of symmetric random matrices with correlated entries. *Stochastic Processes Appl.* **125**(7), 2700–2726 (2015)

4. Bentkus, V., Götze, F., Paulauskas, V., Račkauskas, A.: The accuracy of Gaussian approximation in Banach spaces. *Itogi nauki i tehniki* **81**, 39–139 (1991)
5. Billingsley, P.: *Convergence of probability measures*. Wiley, New York (1968)
6. Cartan, H.: *Calcul Différentiel. Formes Différentielles*. Hermann, Paris (1967)
7. Chatterjee, S.: A generalization of the Lindeberg principle. *Ann. Probab.* **34**(6), 2061–2076 (2006)
8. Chevet, S.: Compacité dans l'espace des probabilités de Radon gaussiennes sur un Banach. *C. R. Acad. Sci. Paris Sér. I Math.* **296**, 275–278 (1983)
9. Cremers, H., Kadelka, D.: On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L_p^E . *Stochastic Processes Appl.* **21**, 305–317 (1986)
10. Embrechts, P., Maejima, M.: The central limit theorem for summability methods of i.i.d. random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **68**, 191–204 (1984)
11. Fraenkel, L.E.: Formulae for high derivatives of composite functions. *Math. Proc. Camb. Phil. Soc.* **83**, 159–165 (1978)
12. Giné, E., León, J.R.: On the central limit theorem in Hilbert space. *Stochastica* **4**(1), 43–71 (1980)
13. Giné, E., Nickl, R.: *Mathematical Foundations of Infinite-Dimensional Statistical Models*. Cambridge University Press (2016)
14. Hájek, P., Johaniš, M.: *Smooth Analysis in Banach Spaces, De Gruyter Series in Nonlinear Analysis and Applications*, vol. 19. De Gruyter (2014)
15. Hoffmann-Jørgensen, J., Pisier, G.: The law of large numbers and the central limit theorem in Banach spaces. *Ann. Probab.* **4**(4), 587–599 (1976)
16. Hytönen, T., van Neerven, J., Veraar, M., Weiss, L.: *Analysis in Banach Spaces*. Springer (2016)
17. Ledoux, M., Talagrand, M.: *Probability in Banach Spaces*. Springer-Verlag, Berlin, Heidelberg (1991)
18. Lindeberg, J.W.: Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung. *Math. Zeit.* **15**, 211–225 (1922)
19. Paulauskas, V., Račkauskas, A.: *Approximation Theory in the Central Limit Theorem. Exact results in Banach spaces*. Kluwer Academic Publishers (1989)
20. Pisier, G.: Martingales with values in uniformly convex spaces. *Israel J. of Math.* **20**(3–4), 326–350 (1975)
21. Rao Ranga, R.: Relations between weak and uniform convergence of measures with applications. *Ann. Math. Statist.* **33**(2), 659–680 (1962)
22. Rosiński, J.: Central limit theorem for dependent random vectors in Banach spaces. *Lecture Notes in Math.* **939**, 157–180 (1982)
23. Sundaresan, K.: Smooth Banach spaces. *Math. Annalen* **173**, 191–199 (1967)
24. Tao, T., Vu, V.: *Acta Math. Random matrices: universality of local eigenvalue statistics* **206**(1), 127–204 (2011)
25. Tao, T., Vu, V.: Random matrices: the universality phenomenon for Wigner ensembles. In: *Proceedings of Symposia in Applied Mathematics*, vol. 72, pp. 121–172 (2014)
26. Zolotarev, V.M.: Ideal metrics in the problem of approximating the distributions of sums of independent random variables. *Theor. Probability Appl.* **22**(3), 433–449 (1977)