HIGHER MOMENTS OF BANACH SPACE VALUED RANDOM VARIABLES

SVANTE JANSON AND STEN KAIJSER

Abstract. We define the k:th moment of a Banach space valued random variable as the expectation of its k:th tensor power; thus the moment (if it exists) is an element of a tensor power of the original Banach space.

We study both the projective and injective tensor products, and their relation. Moreover, in order to be general and flexible, we study three different types of expectations: Bochner integrals, Pettis integrals and Dunford integrals.

One of the problems studied is whether two random variables with the same injective moments (of a given order) necessarily have the same projective moments; this is of interest in applications. We show that this holds if the Banach space has the approximation property, but not in general.

Several sections are devoted to results in special Banach spaces, including Hilbert spaces, $C(K)$ and $D[0,1]$. The latter space is non-separable, which complicates the arguments, and we prove various preliminary results on e.g. measurability in $D[0,1]$ that we need.

One of the main motivations of this paper is the application to Zolotarev metrics and their use in the contraction method. This is sketched in an appendix.

1. Introduction

Let X be a random variable with values in a Banach space B. To avoid measurability problems, we assume for most of this section for simplicity that B is separable and X Borel measurable; see Section 3 for measurability in the general case. Moreover, for definiteness, we consider real Banach spaces only; the complex case is similar.

If $\mathbb{E}\,\|X\| < \infty$, then the mean $\mathbb{E}\,X$ exists as an element of B (e.g. as a Bochner integral $\int X \,\mathrm{d}\,\mathbb{P}$, see Section 5). Suppose now that we want to define the k:th moments of X for some $k \geqslant 2$, assuming for simplicity $\mathbb{E}\,\|X\|^k < \infty$. If B is finite-dimensional, then the second moment of X is a matrix (the covariance matrix, if X is centred), and higher moments are described by higher-dimensional arrays of joint moments of the components. In general, it is natural to define the k:th moment of X using tensor products, see Section 4

Date: 22 August, 2012; corrected 25 August 2012.

2010 *Mathematics Subject Classification.* 60B11; 46G10.

SJ partly supported by the Knut and Alice Wallenberg Foundation.

for details: $X^{\otimes k}$ is a random element of the projective tensor product $B^{\hat{\otimes} k}$, and we define the projective k:th moment of X as the expectation $\mathbb{E}\, X^{\otimes k} \in B^{\hat{\otimes} k}$ (when this expectation exists, e.g. if $\mathbb{E}\,\|X\|^k < \infty$); we denote this moment by $\mathbb{E}\, X^{\hat{\otimes} k}$. In particular, the second moment is $\mathbb{E}\, X^{\hat{\otimes} 2} = \mathbb{E}(X \otimes X) \in B \mathbin{\hat{\otimes}} B$.

An alternative is to consider the injective tensor product $B^{\check{\otimes} k}$ and the injective k:th moment $\mathbb{E}\, X^{\check{\otimes} k} \in B^{\check{\otimes} k}$.

Another alternative is to consider *weak moments*, i.e., joint moments of the real-valued random variables $x^*(X)$ for $x^* \in B^*$ (the dual space). The weak k:th moment thus can be defined as the function

$$(x_1^*, \dots, x_k^*) \mapsto \mathbb{E}\big(x_1^*(X) \cdots x_k^*(X)\big) \in \mathbb{R}, \tag{1.1}$$

assuming that this expectation always exists (which holds, for example, if $\mathbb{E}\,\|X\|^k < \infty$). Note that the weak k:th moment is a k-linear form on B^*.

The purpose of the present paper is to study these moments and their relations in detail, thus providing a platform for further work using moments of Banach space valued random variables. In particular, we shall give sufficient, and sometimes necessary, conditions for the existence of moments in various situations.

One example of our results on relations between the different moments is that, at least in the separable case, the weak k:th moment is equivalent to the injective moment. (See Theorem 6.10 for a precise statement.)

We study also the problem of moment equality: if Y is a second random variable with values in B, we may ask whether X and Y have the same k:th moments, for a given k and a given type of moment. (Assume for example that $\mathbb{E}\,\|X\|^k, \mathbb{E}\,\|Y\|^k < \infty$ so that the moments exist.) This problem, for the second moment, appears for example in connection with the central limit theorem for Banach space valued random variables, see e.g. [42, Chapter 10] where weak moments are used. (In particular, a B-valued random variable X is said to be *pregaussian* if $\mathbb{E}\, X = 0$ and there exists a Gaussian B-valued random variable Y such that X and Y have the same weak second moments.) This problem is also important when proving convergence in distribution of some Banach space valued random variables using a Zolotarev metric, see Appendix B. (The Zolotarev metrics are, for example, often used when applying the contraction method for sequences of random variables with a suitable recursive structure, see e.g. Neininger and Sulzbach [49]. This applications is one of the main motivations of the present paper.)

As an example of results obtained in later sections, let us consider this problem of moment equality further. In particular, we want to compare the property that X and Y have the same k:th moment for the different types of moments. (For simplicity, we assume as above that $\mathbb{E}\,\|X\|^k, \mathbb{E}\,\|Y\|^k < \infty$, so the moments exist.)

Since the dual space of $B^{\hat{\otimes} k}$ is the space of bounded k-linear forms on B (see Section 4), it follows (by Section 5) that the k:th projective moments

$\mathbb{E}\, X^{\hat{\otimes} k}$ and $\mathbb{E}\, Y^{\hat{\otimes} k}$ are equal if and only if

$$\mathbb{E}\,\alpha(X,\dots,X) = \mathbb{E}\,\alpha(Y,\dots,Y), \qquad \alpha \in L(B^k;\mathbb{R}), \tag{1.2}$$

where $L(B^k;\mathbb{R})$ denotes the space of bounded k-linear forms on B. (See Corollary 6.17.)

Moreover, by the definition of weak moments, X and Y have the same weak k:th moments if and only if

$$\mathbb{E}\big(x_1^*(X)\cdots x_k^*(X)\big) = \mathbb{E}\big(x_1^*(Y)\cdots x_k^*(Y)\big), \qquad x_1^*,\dots,x_k^* \in B^*. \tag{1.3}$$

We shall see (Corollary 6.12) that this holds if and only if the injective moments $\mathbb{E}\, X^{\check{\otimes} k}$ and $\mathbb{E}\, Y^{\check{\otimes} k}$ are equal.

For any $x_1^*,\dots,x_k^* \in B^*$, the mapping $(x_1,\dots,x_k) \mapsto x_1^*(x_1)\cdots x_k^*(x_k)$ is a bounded k-linear form on B, and thus (1.3) is a special case of (1.2). Consequently, (1.2) $\implies$ (1.3), i.e., if X and Y have the same projective moments, then they have the same weak (and injective) moments. Does the converse hold? (This question is of practical importance in applications of the contraction method, see Appendix B.) We show that this problem is non-trivial, and deeply connected to the *approximation property* of Banach spaces. (See Section 8 for definitions and proofs.) In particular, we have the following results.

Theorem 1.1. *If B is a separable Banach space with the approximation property and X and Y are random variables in B such that $\mathbb{E}\,\|X\|^k, \mathbb{E}\,\|Y\|^k < \infty$, then* (1.2) *and* (1.3) *are equivalent, i.e., X and Y have the same projective k:th moments if and only if they have the same weak k:th moments.*

Theorem 1.2. *There exists a separable Banach space B and bounded random variables X and Y in B such that, for $k = 2$,* (1.3) *holds but not* (1.2).

In Theorem 1.2, we may further require B to be reflexive.

Note that all classical Banach spaces have the approximation property (for example, ℓ^p, c_0, $L^p(\mu)$, $C(K)$, and any Banach space with a basis), and that counterexamples are notoriously difficult to find. In fact, the approximation property was formulated and studied by Grothendieck [32, 33], but it took almost 20 years until a Banach space without the approximation property was found by Enflo [26]. Hence, it is unlikely that such spaces will appear in applications, and Theorem 1.1 ought to apply to any separable Banach space B that will be used in practice. (Note, however, that the non-separable Banach space $B(H)$ of bounded operators in a Hilbert space lacks the approximation property, see Szankowski [61].)

In applications, B is often a function space, for example $C[0,1]$. In this case, we can weaken the condition for equality of moments further, by considering only point evaluations in (1.3). This yields the following result, stated here more generally for $C(K)$ where K is a compact metric space.

Theorem 1.3. *Let $B = C(K)$ where K is a compact metric space, and let X and Y be random variables in $C(K)$ such that $\mathbb{E}\,\|X\|^k$, $\mathbb{E}\,\|Y\|^k < \infty$. Then* (1.2) *is equivalent to* (1.3)*, and further to*

$$\mathbb{E}\big(X(t_1)\cdots X(t_k)\big) = \mathbb{E}\big(Y(t_1)\cdots Y(t_k)\big), \qquad t_1,\dots,t_k \in K. \tag{1.4}$$

In this case it is thus enough to study joint moments of $X(t)$, and the k:th moment of X is described by the real-valued function $\mathbb{E}\big(X(t_1)\cdots X(t_k)\big)$ on K^k. We shall further see that if $k = 2$, the integrability conditions can be weakened to $\sup_{t\in K} |X(t)|^2 < \infty$ and $\sup_{t\in K} |Y(t)|^2 < \infty$ (Theorem 11.23).

Remark 1.4. A standard polarisation argument, as in the proof of Theorem 18.8, shows that (1.3) is equivalent to

$$\mathbb{E}\big(x^*(X)^k\big) = \mathbb{E}\big(x^*(Y)^k\big), \qquad x^* \in B^*. \tag{1.5}$$

In other words, the weak k:th moments of X and Y are equal if and only $x^*(X)$ and $x^*(Y)$ have the same k:th moments for every $x^* \in B^*$. Hence we can use (1.5) instead of (1.3) in the results above. In contrast, in (1.4), it is essential to allow different $t_1,\dots,t_k$ and joint moments.

We have so far, for simplicity, assumed that B is separable. The same definitions apply in the non-separable case, but there are then technical complications (concerning measurability) that complicate both statements and proofs, and our results are less complete in this case.

The space $D[0,1]$ is a non-separable Banach space that is important in applications. We treat $D[0,1]$ in detail in Sections 13–16; by special arguments we obtain essentially the same results as in the separable case. In particular, we shall see (Theorem 16.13) that Theorem 1.3 holds also for $B = D[0,1]$, with the usual measurability condition in $D[0,1]$.

For the reader's convenience (and our own), we have in Section 2 collected some notation used in this paper, and in Sections 3–5 preliminaries on measurability, tensor products and integration in Banach spaces. (There are no new results in these sections.)

The main definitions of the moments are given in Section 6, together with various results giving sufficient, and sometimes necessary, conditions for their existence. (Some simple examples are given in Section 7.) We give simple sufficient conditions that are enough for many applications, but we also give more precise results. We try to be as precise as possible, and therefore we use three different types of integrability (Bochner, Pettis and Dunford, see Section 5) in the definition of the projective and injective moments, leading to six different cases that we treat in detail. (The multitude of cases may be bewildering, and contributes to the length of the paper. The reader is recommended to concentrate on separable Banach spaces and Bochner integrals at the first reading; this is enough for many applications. However, for e.g. applications to $D[0,1]$, this is not enough which is one motivation for considering also Pettis and Dunford integrals.)

The approximation property is defined and used in Section 8.

Sections 9–16 study special Banach spaces: Hilbert spaces in Section 9; $L^p(\mu)$ in Section 10; $C(K)$ (where K is a compact space) in Section 11 (with emphasis on the separable case, when K is metrizable); $c_0(S)$ in Section 12 (with emphasis on the non-separable case); and finally, as said above, $D[0,1]$ in Sections 13–16, where Sections 13–15 contain various preliminary results on e.g. the dual space and maximal ideal space as well as measurability and separability of random variables in $D[0,1]$. In these sections we give many results on existence of moments of the different types for random variables in these spaces.

In the final two sections we consider the collection of moments of all orders. Section 17 shows that, under certain conditions, the moments determine the distribution, and Section 18 treats the problem whether convergence of the moments for a sequence of random variables implies convergence in distribution. We give both positive and negative results.

The appendices discuss two well-known constructions related to moments.

Appendix A describes the construction of a Hilbert space (the *reproducing Hilbert space*) connected to a B-valued random variable; we show that this is closely related to the injective second moment.

Appendix B describes the Zolotarev metrics and their connection to projective moments; as said above this is a major motivation for the present paper.

Throughout we usually try to give as general results as possible. We also give various counterexamples showing limitations of the results, especially in the non-separable case; these (and many technical remarks) can be skipped at the first reading. Some open problems are stated explicitly or implicitly. (As is often the case with integration in Banach spaces, cf. e.g. [29], the non-separable case is much more complicated than the separable case and several open problems remain.)

For completeness we include some known results that we need, with or without proof; we try to give references in both cases, but omit them for some results or arguments that we regard as standard. (The absence of a reference thus does not imply that the result is new.) We believe that many other results are new.

Acknowledgement. We thank Joe Diestel, Ralph Neininger, Oleg Reinov, Viggo Stoltenberg-Hansen and Henning Sulzbach for helpful comments.

2. Notations

We will use the following standard notations, usually without comment.

$L(B_1,\dots,B_k;B')$ is the space of bounded k-linear maps $B_1\times\dots\times B_k\to B'$. In particular, with $B'=\mathbb{R}$, we have the space of bounded k-linear forms. When $B_1=\dots=B_k=B$ we also write $L(B^k;B')$.

B^* denotes the dual space of the Banach space B, i.e., the space $L(B;\mathbb{R})$ of bounded linear functionals $B\to\mathbb{R}$. If $x\in B$ and $x^*\in B^*$, we use the

notations $x^*(x)$ and $\langle x^*, x\rangle$, or $\langle x, x^*\rangle$, as synonyms. We write $x^* \perp B_1$, for a subset $B_1 \subseteq B$, if $\langle x^*, x\rangle = 0$ for every $x \in B_1$.

We use several standard Banach spaces in our results and examples; for convenience we recall their definitions here.

For any set S and $p \in [1,\infty)$, $\ell^p(S)$ is the Banach space of all functions $S \to \mathbb{R}$ such that the norm $\|f\|_{\ell^p(S)} := (\sum_{s\in S} |f(s)|^p)^{1/p}$ is finite. (We are mainly interested in the cases $p = 1, 2$.) Further, $\ell^\infty(S)$ is the Banach space of all bounded functions $S \to \mathbb{R}$, with the norm $\|f\|_{\ell^\infty(S)} := \sup_{s\in S} |f(s)|$. We use sometimes the notation f_s for $f(s)$, in particular when $S = \mathbb{N}$.

We define the *support* of f as $\operatorname{supp}(f) := \{s : f(s) \neq 0\}$, and note that if $f \in \ell^p(S)$, with $p < \infty$, then $\operatorname{supp}(f)$ is countable even if S is uncountable.

$c_0(S)$ is the space of all function $f : S \to \mathbb{R}$ such that $\{s : |f(s)| > \varepsilon\}$ is finite for each $\varepsilon > 0$; this is a closed subspace of ℓ^∞, and is thus a Banach space with the inherited norm $\|f\|_{c_0(S)} := \|f\|_{\ell^\infty(S)} := \sup_{s\in S} |f(s)|$. Note that every element of $c_0(S)$ has countable support.

e_s denotes the function $e_s(t) := \mathbf{1}\{t = s\}$ that is 1 at s and 0 everywhere else (defined for t in some set S, which will be clear from the context).

Let $c_{00}(S)$ be the space of all functions $f : S \to \mathbb{R}$ with $\operatorname{supp}(f)$ finite; this is the linear span of $\{e_s : s \in S\}$. Then $c_{00}(S) \subseteq c_0(S) \subseteq \ell^\infty(S)$, and $c_{00}(S)$ is dense in $c_0(S)$. Hence $c_0(S)$ is the closed linear span in $\ell^\infty(S)$ of the functions $\{e_s : s \in S\}$. It follows that

$$c_{00}(S)^* = c_0(S)^* = \ell^1(S),$$

with the standard pairing $\langle f, g\rangle = \sum_{s\in S} f(s)g(s)$.

When $S = \mathbb{N}$, we write just ℓ^p and c_0.

$L^p(S)$, where $S = (S, \mathcal{S}, \mu)$ is a measure space and $p \in [1,\infty)$, is the space of all measurable functions $f : S \to \mathbb{R}$ such that $\int_S |f|^p < \infty$ (as usual identifying functions that are equal a.e.).

$C(K)$, where K is a compact topological space, is the space of all continuous functions $f : K \to \mathbb{R}$, with the norm $\sup_K |f|$. (We are particulary interested in $C[0,1]$.)

For a compact set K, $M(K)$ is the Banach space of all signed Borel measures on K. By the Riesz representation theorem, the dual $C(K)^*$ can be identified with the subspace $M_{\mathrm{r}}(K)$ of $M(K)$ consisting of regular measures; see e.g. [12, Theorem 7.3.5], [22, Theorem IV.6.3] or [13, Theorem III.5.7]. (If e.g. K is compact and metrizable, then every signed measure is regular, so $C(K)^* = M(K)$, see [12, Propositions 7.1.12, 7.2.3, 7.3.3].)

$D[0,1]$ denotes the linear space of functions $[0,1] \to \mathbb{R}$ that are right-continuous with left limits, see e.g. [6, Chapter 3]. The norm is $\sup_{[0,1]} |f|$. See further Section 13.

δ_s denotes the Dirac measure at s, as well as the corresponding point evaluation $f \mapsto f(s)$ seen as a linear functional on a suitable function space.

$(\Omega, \mathcal{F}, \mathbb{P})$ denotes the underlying probability space where our random variables are defined; ω denotes an element of Ω. We assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is complete.

We let $\bar{\mathbb{E}}$ denote the upper integral of a, possibly non-measurable, real-valued function on Ω:

$$\bar{\mathbb{E}}Y := \inf\bigl\{\mathbb{E}\,Z : Z \geqslant Y \text{ and } Z \text{ is measurable}\bigr\}. \tag{2.1}$$

(If Y is measurable, then $\bar{\mathbb{E}}Y = \mathbb{E}\,Y$.) In particular, if X is B-valued, then $\bar{\mathbb{E}}\|X\| < \infty$ if and only if there exists a positive random variable Z with $\|X\| \leqslant Z$ and $\mathbb{E}\,Z < \infty$.

The exponent k is, unless otherwise stated, an arbitrary fixed integer $\geqslant 1$, but the case $k = 1$ is often trivial. For applications, $k = 2$ is the most important, and the reader is adviced to primarily think of that case.

3. Measurability

A B-valued random variable is a function $X : \Omega \to B$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. (As said above, we assume that the probability space is complete.) We further want X to be measurable, and there are several possibilities to consider; we will use the following definitions.

Definition 3.1. Let $X : \Omega \to B$ be a function on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a Banach space B.

(i) X is *Borel measurable* if X is measurable with respect to the Borel σ-field $\mathcal{B}$ on B, i.e., the σ-field generated by the open sets.

(ii) X is *weakly measurable* if X is measurable with respect to the σ-field $\mathcal{B}_w$ on B generated by the continuous linear functionals, i.e., if $\langle x^*, X\rangle$ is measurable for every $x^* \in B^*$.

(iii) X is *a.s. separably valued* if there exists a separable subspace $B_1 \subseteq B$ such that $X \in B_1$ a.s.

(iv) X is *weakly a.s. separably valued* if there exists a separable subspace $B_1 \subseteq B$ such that if $x^* \in B^*$ and $x^* \perp B_1$, then $x^*(X) = 0$ a.s.

(v) X is *Bochner measurable* if X is Borel measurable and a.s. separably valued.

Remark 3.2. X is Bochner measurable if and only if X is Borel measurable and *tight*, i.e., for every $\varepsilon > 0$, there exists a compact subset $K \subset B$ such that $\mathbb{P}(X \in K) > 1 - \varepsilon$, see [6, Theorem 1.3]. (This is also called *Radon*.) Some authors use the name *strongly measurable*. Moreover, X is Bochner measurable if and only if there exists a sequence X_n of measurable simple functions $\Omega \to B$ such that $X_n \to X$ a.s., see [22, III.2.10 and III.6.10–14]. (This is often taken as the definition of Bochner measurable.) See further [42, Chapter 2.1].

The name *scalarly measurable* is sometimes used for weakly measurable.

Some authors (e.g. [22]) let "measurable" mean what we call Bochner measurable, so care should be taken to avoid misunderstandings.

If B is separable, there is no problem. Then every X is trivially a.s. separably valued; moreover, it is easy to see that $\mathcal{B}_w = \mathcal{B}$. Hence every weakly measurable X is measurable, and the notions in (i), (ii) and (v)

are equivalent. This extends immediately to a.s. separably valued random variables in an arbitrary Banach space:

Theorem 3.3 (Pettis, [22, III.6.10–11])**.** *If X is a.s. separably valued, then X is Borel measurable (and thus Bochner measurable) if and only if X is weakly measurable.* □

Remark 3.4. There is a converse, assuming some set theory hypotheses. (See further Remark 9.13.) By [45] (see also [5, Appendix III]), if the cardinality of B is not real-measurable then every Borel measurable random variable in B is a.s. separably valued. It follows, in particular, that if the continuum hypothesis holds and furthermore there is no inaccessible cardinal, then every Borel measurable random variable in any Banach space is a.s. separably valued. (These hypotheses are both consistent with the usual ZFC set theory; see further Remark 9.13.) Hence, it can safely be assumed that every Borel measurable random variable that will appear in an application is a.s. separably valued. In other words, if we want to study random variables that are not a.s. separably valued, then we cannot use Borel measurability.

In view of this, we prefer to use separable Banach spaces, or at least a.s. separably valued random variables, whenever possible. Many standard Banach spaces are separable, for example $L^p[0,1]$ ($1 \leqslant p < \infty$) and $C[0,1]$. However, the non-separable space $D[0,1]$ (see Section 13) is important for applications, and it is easily seen (Theorem 15.11 below) that a D-valued random variable X is a.s. separably valued if and only if there exists a (non-random) countable set $A \subset [0,1]$ such that X a.s. has all its points of discontinuity in A. In practical applications this means that a random variable $X \in D[0,1]$ is a.s. separably valued if it has jumps at deterministic places, but not if there are jumps at random places.

Example 3.5 (See [6, Section 15])**.** Let $U \sim \mathrm{U}(0,1)$ be a uniformly distributed random variable, and let X be the random element of $D[0,1]$ given by $X = \mathbf{1}_{[U,1]}$, i.e. $X(t) = \mathbf{1}\{U \leqslant t\}$. (This is the empirical distribution function of U, seen as a sample of size 1.) Since the functions $\mathbf{1}_{[u,1]}$, $u \in [0,1]$, all have distance 1 to each other, they form a discrete subset of $D[0,1]$ and it follows that X is not a.s. separably valued. In particular, X is not Bochner measurable.

Moreover, if A is a non-measurable subset of $[0,1]$, then the set $\{\mathbf{1}_{[u,1]} : u \in A\}$ is a closed subset of $D[0,1]$, and thus a Borel set. If we take the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where U and thus X are defined to be $[0,1]$ with Lebesgue measure, and $U(\omega) = \omega$, then $X^{-1}(A) = A \notin \mathcal{F}$ and, consequently, X is not Borel measurable in $D[0,1]$. (If we assume the continuum hypothesis, X cannot be Borel measurable for any probability space $(\Omega, \mathcal{F}, \mathbb{P})$, see Remark 3.4.)

On the other hand, $X(t)$ is measurable for each t, and by Pestman [51], see Corollary 13.2 and Theorem 15.5 below, X is weakly measurable.

Furthermore, it follows from [51] (or Corollary 13.2) also that if $x^* \in D^*$ and $x^* \perp C[0,1]$, then $x^*(\mathbf{1}_{[u,1]}) = 0$ for all but countably many u; thus $x^*(X) = 0$ a.s. which shows that X is weakly a.s. separably valued. (Cf. Theorem 15.13.)

Cf. also [62, Example 3-2-2] which studies essentially the same example but as an element of $L^\infty[0,1]$.

Remark 3.6. If X is Borel measurable, then $\|X\|$ is measurable, since $x \mapsto \|x\|$ is continuous. However, if X only is weakly measurable, then $\|X\|$ is not always measurable without additional hypotheses. (For this reason, we will sometimes use the upper integral $\bar{\mathbb{E}}\|X\|$.) If X is weakly measurable and a.s. separably valued, then $\|X\|$ is measurable, e.g. by Theorem 3.3. (In particular, there is no problem when B is separable.) Furthermore, if B has the property that there exists a countable norm-determining set of linear functionals, then every weakly measurable X in B has $\|X\|$ measurable; $D[0,1]$ is an example of such a space.

Remark 3.7. Several other forms of measurability may be considered, for example using the *Baire* σ-field (generated by the continuous functions $B \to \mathbb{R}$) [62] or the σ-field generated by the closed (or open) balls in B [20, 21], [6]. Note further that, in general, $\mathcal{B}_w$ is not the same as the σ-field generated by the weak topology. (In fact, $\mathcal{B}_w$ equals the Baire σ-field for the weak topology [23], [62].) When B is separable, all these coincide with the Borel σ-field.

See also [23], [24], [42] and [62], where further possibilities are discussed.

4. Tensor products of Banach spaces

We give a summary of the definitions and some properties of the two main tensor products of Banach spaces. We refer to e.g. Blei [8] or Ryan [57] for further details. We consider the general case of the tensor product of k different spaces. The tensor products we consider (both algebraic and completed) are associative in a natural way; for example, $B_1 \otimes B_2 \otimes B_3 = (B_1 \otimes B_2) \otimes B_3 = B_1 \otimes (B_2 \otimes B_3)$, and the general case may be reduced to tensor products of two spaces. (Many authors, including [57], thus consider only this case.)

4.1. **Algebraic tensor products.** The *algebraic tensor product* of a finite sequence of vector spaces $B_1, \dots, B_k$ (over an arbitrary field) can be defined in an abstract way as a vector space $B_1 \otimes \cdots \otimes B_k$ with a k-linear map $B_1 \times \cdots \times B_k \to B_1 \otimes \cdots \otimes B_k$, written $(x_1, \dots x_k) \to x_1 \otimes \cdots \otimes x_k$, such that if $\alpha : B_1 \times \cdots \times B_k \to A$ is any k-linear map, then there is a unique linear map $\tilde{\alpha} : B_1 \otimes \cdots \otimes B_k \to A$ such that

$$\alpha(x_1, \dots, x_k) = \tilde{\alpha}(x_1 \otimes \cdots \otimes x_k). \tag{4.1}$$

(All such spaces are naturally isomorphic, so the tensor product is uniquely defined, up to trivial isomorphisms.)

Several concrete constructions can also be given. One useful construction is to let $B_i^\sharp$ be the algebraic dual of B_i and define $B_1\otimes\cdots\otimes B_k$ as a subspace of the linear space of all k-linear forms on $B_1^\sharp\times\cdots\times B_k^\sharp$; more precisely we define $x_1\otimes\cdots\otimes x_k$ as the k-linear form on $B_1^\sharp\times\cdots\times B_k^\sharp$ defined by

$$x_1\otimes\cdots\otimes x_k(x_1^*,\dots,x_k^*)=x_1^*(x_1)\cdots x_k^*(x_k), \tag{4.2}$$

and then define $B_1\otimes\cdots\otimes B_k$ as the linear span of all $x_1\otimes\cdots\otimes x_k$.

We can modify this construction by replacing $B_i^\sharp$ by any subspace that separates the points of B_i. In particular, when each B_i is a Banach space, we can regard $B_1\otimes\cdots\otimes B_k$ as a subspace of the space of k-linear forms on $B_1^*\times\cdots\times B_k^*$.

An element of $B_1\otimes\cdots\otimes B_k$ of the form $x_1\otimes\cdots\otimes x_k$ is called an *elementary tensor*. Note that not every element of $B_1\otimes\cdots\otimes B_k$ is an elementary tensor, but every element is a finite linear combination of elementary tensors (in a non-unique way).

A sequence of linear operators $T_i:A_i\to B_i$ defines a unique linear map $T_1\otimes\cdots\otimes T_k:A_1\otimes\cdots\otimes A_k\to B_1\otimes\cdots\otimes B_k$ such that $T_1\otimes\cdots\otimes T_k(x_1\otimes\cdots\otimes x_k)=T_1x_1\otimes\cdots\otimes T_kx_k$ for elementary tensors.

4.2. **Completed tensor products.** When $B_1,\dots,B_k$ are Banach spaces, we can define several different (non-equivalent) norms on $B_1\otimes\cdots\otimes B_k$. For each norm we then can take the completion of $B_1\otimes\cdots\otimes B_k$, obtaining a Banach space. We consider two cases (the two main cases) of completed tensor products.

The *projective tensor norm* is defined on $B_1\otimes\cdots\otimes B_k$ by

$$\|u\|_\pi:=\inf\Bigl(\sum_j\|x_{1j}\|_{B_1}\cdots\|x_{kj}\|_{B_k}:u=\sum_j x_{1j}\otimes\cdots\otimes x_{kj}\Bigr), \tag{4.3}$$

taking the infimum over all ways of writing u as a finite sum of elementary tensors. The corresponding completed tensor product is called the *projective tensor product* and is written $B_1\mathbin{\widehat\otimes}\cdots\mathbin{\widehat\otimes}B_k$. It is easily seen that every $u\in B_1\mathbin{\widehat\otimes}\cdots\mathbin{\widehat\otimes}B_k$ can be written (non-uniquely) as an absolutely convergent infinite sum of elementary tensors: for any $\varepsilon>0$,

$$u=\sum_{j=1}^\infty x_{1j}\otimes\cdots\otimes x_{kj}\quad\text{with}\quad\sum_{j=1}^\infty\|x_{1j}\|_{B_1}\cdots\|x_{kj}\|_{B_k}\leqslant(1+\varepsilon)\|u\|_\pi<\infty; \tag{4.4}$$

equivalently,

$$u=\sum_{j=1}^\infty\lambda_jx_{1j}\otimes\cdots\otimes x_{kj}\quad\text{with}\quad\|x_{ij}\|_{B_i}\leqslant1,\ \lambda_j\geqslant0\text{ and}\ \sum_{j=1}^\infty\lambda_j\leqslant(1+\varepsilon)\|u\|_\pi<\infty. \tag{4.5}$$

The *injective tensor norm* is defined on $B_1 \otimes \cdots \otimes B_k$ by using (4.2) to regard $B_1 \otimes \cdots \otimes B_k$ as a subspace of $L(B_1^*, \dots, B_k^*; \mathbb{R})$, the bounded k-linear forms on $B_1^* \times \cdots \times B_k^*$, and taking the induced norm, i.e.,

$$\|u\|_\varepsilon := \sup\bigl(|u(x_1^*, \dots, x_k^*)| : \|x_1^*\|_{B_1^*}, \dots, \|x_k^*\|_{B_k^*} \leqslant 1\bigr). \tag{4.6}$$

The corresponding completed tensor product is called the *injective tensor product* and is written $B_1 \mathbin{\check{\otimes}} \cdots \mathbin{\check{\otimes}} B_k$; note that this is simply the closure of $B_1 \otimes \cdots \otimes B_k$ in $L(B_1^*, \dots, B_k^*; \mathbb{R})$ and thus can be regarded as a closed subspace of $L(B_1^*, \dots, B_k^*; \mathbb{R})$.

For elementary tensors we have

$$\|x_1 \otimes \cdots \otimes x_k\|_\pi = \|x_1 \otimes \cdots \otimes x_k\|_\varepsilon = \|x_1\|_{B_1} \cdots \|x_k\|_{B_k}, \tag{4.7}$$

so the two norms coincide. (Any such norm on $B_1 \otimes \cdots \otimes B_k$ is called a *crossnorm*.) Moreover, it is easily seen that for any $u \in B_1 \otimes \cdots \otimes B_k$, $\|u\|_\varepsilon \leqslant \|u\|_\pi$; hence the identity map $u \mapsto u$ extends to a canonical bounded linear map (of norm 1)

$$\iota : B_1 \mathbin{\hat{\otimes}} \cdots \mathbin{\hat{\otimes}} B_k \to B_1 \mathbin{\check{\otimes}} \cdots \mathbin{\check{\otimes}} B_k \tag{4.8}$$

with $\|\iota\| \leqslant 1$. Unfortunately, this map is not always injective; we return to this problem in Section 8, where we shall see that this is the source of the difference between (1.2) and (1.3).

Consider for simplicity the case $k = 2$. As said above, $B_1 \mathbin{\check{\otimes}} B_2$ can be seen (isometrically) as a subspace of $L(B_1^*, B_2^*; \mathbb{R})$. Moreover, an elementary tensor $x_1 \otimes x_2$ defines a bounded linear operator $B_1^* \to B_2$ by $x_1^* \mapsto x_1^*(x_1)x_2$, and this extends by linearity to a mapping $B_1 \otimes B_2 \to L(B_1^*; B_2)$, which is an isometry for the injective tensor norm (4.6). Hence, this mapping extends to an isometric embedding of $B_1 \mathbin{\check{\otimes}} B_2$ as a subspace of the space $L(B_1^*; B_2)$ of bounded linear operators $B_1^* \to B_2$. Explicitly, $u \in B_1 \mathbin{\check{\otimes}} B_2$ corresponds to the operator $T_u : B_1^* \to B_2$ given by

$$\langle T_u x^*, y^* \rangle = \langle x^* \otimes y^*, u \rangle, \qquad x^* \in B_1^*,\ y^* \in B_2^*. \tag{4.9}$$

(By symmetry, there is also an embedding into $L(B_2^*; B_1)$. If $u \in B_1 \mathbin{\check{\otimes}} B_2$ corresponds to the operator $T_u : B_1^* \to B_2$, it also corresponds to $T_u^* : B_2^* \to B_1$.)

When the mapping ι in (4.8) (with $k = 2$) is injective, we may regard $B_1 \mathbin{\hat{\otimes}} B_2$ as a subspace of $B_1 \mathbin{\check{\otimes}} B_2$ (with a generally different, larger, norm); hence $B_1 \mathbin{\hat{\otimes}} B_2$ may be regarded as a certain space of operators $B_1^* \to B_2$ in this case too.

Example 4.1. Suppose that $B_1, \dots, B_k$ are finite-dimensional, and let $\{e_{ij}\}_j$, $j \in \mathcal{J}_i$, be a basis of B_i for $i = 1, \dots, k$. It is easily seen that the elementary tensors $e_J := e_{1j_1} \otimes \cdots \otimes e_{kj_k}$, with $J := (j_i)_i \in \prod_{i=1}^k \mathcal{J}_i$, form a basis of $B_1 \otimes \cdots \otimes B_k$.

Hence $\dim(B_1 \otimes \cdots \otimes B_k) = \prod_{i=1}^k \dim(B_i) < \infty$. In particular, the tensor product is complete for any norm, and thus $B_1 \mathbin{\hat{\otimes}} \cdots \mathbin{\hat{\otimes}} B_k = B_1 \mathbin{\check{\otimes}} \cdots \mathbin{\check{\otimes}} B_k = B_1 \otimes \cdots \otimes B_k$ as sets (although the norms generally differ).

Using the basis $(e_J)_J$, any element of the tensor product may be written as $\sum_J a_J e_J$, where the coordinates a_J are indexed by $J \in \prod_{i=1}^k \mathcal{J}_i$, so it is natural to consider the coordinates as a k-dimensional array. (A matrix in the case $k = 2$.)

Example 4.2. Let H be a Hilbert space. The tensor products $H\hat{\otimes}H$ and $H\check{\otimes}H$ can be identified with the spaces of trace class operators and compact operators on H, respectively, see Theorem 9.2; moreover, $H\hat{\otimes}H \cong (H\check{\otimes}H)^*$.

In this case, another interesting choice of norm on $H \otimes H$ is the Hilbertian tensor norm, given by the inner product $\langle x_1 \otimes x_2, y_1 \otimes y_2\rangle = \langle x_1, y_1\rangle\langle x_2, y_2\rangle$; the corresponding completed tensor product $H \otimes_2 H$ can be identified with the space of Hilbert–Schmidt operators on H. (We will not use this tensor product in the present paper.)

If $T_i : A_i \to B_i$ are bounded linear operators, then $T_1 \otimes \cdots \otimes T_k$ is bounded for both the projective norms and the injective norms, and extends thus to bounded linear maps $T_1\hat{\otimes}\cdots\hat{\otimes}T_k : A_1\hat{\otimes}\cdots\hat{\otimes}A_k \to B_1\hat{\otimes}\cdots\hat{\otimes}B_k$ and $T_1\check{\otimes}\cdots\check{\otimes}T_k : A_1\check{\otimes}\cdots\check{\otimes}A_k \to B_1\check{\otimes}\cdots\check{\otimes}B_k$. We note the following lemma, which we for simplicity state for the case $k = 2$, although the result extends to general k.

Lemma 4.3. *If $T : A_1 \to B_1$ and $U : A_2 \to B_2$ are injective linear operators between Banach spaces, then $T\check{\otimes}U : A_1\check{\otimes}A_2 \to B_1\check{\otimes}B_2$ is injective.*

Proof. Consider first the case when $A_1 = B_1$ and $T = I$, the identity operator. We can regard $A_1\check{\otimes}A_2$ and $A_1\check{\otimes}B_2$ as subspaces of $L(A_1^*; A_2)$ and $L(A_1^*; B_2)$, and then $I\check{\otimes}U$ is the mapping $L(A_1^*; A_2) \to L(A_1^*; B_2)$ given by $S \mapsto US$, which is injective when U is.

In general we factorize $T\check{\otimes}U = (T\check{\otimes}I)(I\check{\otimes}U)$ and note that both factors $T\check{\otimes}I : A_1\check{\otimes}B_2 \to A_2\check{\otimes}B_2$ and $I \otimes U : A_1\check{\otimes}B_1 \to A_1\check{\otimes}B_2$ are injective by the first case and symmetry. □

Remark 4.4. The same proof shows that if T and U are isometric embeddings, then so is $T\check{\otimes}U$. (In other words, the injective tensor product is injective [57, Section 6.1].)

Remark 4.5. Lemma 4.3 is in general *not* true for the projective tensor product, see Example 8.7 and Remark 8.8.

The projective tensor product has instead the dual property that if T and U are quotient mappings (i.e., onto), then so is $T\hat{\otimes}U$.

When considering the dual space of a completed tensor product, note that $B_1 \otimes \cdots \otimes B_k$ is a dense subspace (by definition); hence a continuous linear functional is the same as a linear functional on $B_1 \otimes \cdots \otimes B_k$ that is bounded for the chosen tensor norm. Furthermore, each such linear functional can by (4.1) be identified with a k-linear form on $B_1 \times \cdots \times B_k$. For the projective tensor product, the definition (4.3) of the norm implies that all bounded k-linear forms yield bounded linear functionals.

Theorem 4.6 ([57, Theorem 2.9])**.** *The dual* $(B_1\mathbin{\widehat{\otimes}}\cdots\mathbin{\widehat{\otimes}}B_k)^*$ *consists of all bounded* k*-linear forms* $B_1\times\cdots\times B_k\to\mathbb{R}$*, with the natural pairing* $\langle\alpha, x_1\otimes\cdots\otimes x_k\rangle=\alpha(x_1,\dots,x_k)$*.* □

For $B_1\mathbin{\check{\otimes}}\cdots\mathbin{\check{\otimes}}B_k$, the dual space consists of a subset of all bounded k-linear forms; these forms are called *integral*, and can be described as follows. Let K_i be the closed unit ball of B_i^* with the weak* topology; thus K_i is a compact space.

Theorem 4.7 ([57, Proposition 3.14])**.** *If* $\chi\in(B_1\mathbin{\check{\otimes}}\cdots\mathbin{\check{\otimes}}B_k)^*$*, then there exists a (non-unique) signed measure* $\mu\in M(K_1\times\cdots\times K_k)$*, with* $\|\mu\|=\|\chi\|$*, such that*

$$\chi(x_1\otimes\cdots\otimes x_k)=\int_{K_1\times\cdots\times K_k}x_1^*(x_1)\cdots x_k^*(x_k)\,\mathrm{d}\mu(x_1^*,\dots,x_k^*).\tag{4.10}$$

Conversely, (4.10) *defines a bounded linear functional for every signed measure* $\mu\in M(K_1\times\cdots\times K_k)$*.*

Proof. We have defined an embedding of $B_1\otimes\cdots\otimes B_k$ in $L(B_1^*,\dots,B_k^*;\mathbb{R})$ by (4.2), and taking the restriction of the operators to $K_1\times\cdots\times K_k$ gives a linear map into $C(K_1\times\cdots\times K_k)$, which by the definition (4.6) of the injective tensor norm is an isometry. Hence we can regard $B_1\mathbin{\check{\otimes}}\cdots\mathbin{\check{\otimes}}B_k$ as a subspace of $C(K_1\times\cdots\times K_k)$, and the result follows by the Hahn–Banach theorem together with the Riesz representation theorem. □

An elementary tensor $x_1^*\otimes\cdots\otimes x_k^*\in B_1^*\otimes\cdots\otimes B_k^*$ defines a k-linear form on $B_1\times\cdots\times B_k$ by $(x_1,\dots,x_k)\mapsto x_1^*(x_1)\cdots x_k^*(x_k)$, and thus a linear functional on $B_1\otimes\cdots\otimes B_k$ by

$$\langle x_1^*\otimes\cdots\otimes x_k^*, x_1\otimes\cdots\otimes x_k\rangle=x_1^*(x_1)\cdots x_k^*(x_k).\tag{4.11}$$

By the definitions above, see (4.6) and (4.2), this linear functional extends to a linear functional on $B_1\mathbin{\check{\otimes}}\cdots\mathbin{\check{\otimes}}B_k$ of norm $\|x_1^*\|\cdots\|x_k^*\|$, and it follows by (4.3), linearity and continuity that every tensor $u^*\in B_1^*\mathbin{\widehat{\otimes}}\cdots\mathbin{\widehat{\otimes}}B_k^*$ defines a linear functional on $B_1\mathbin{\check{\otimes}}\cdots\mathbin{\check{\otimes}}B_k$ with norm at most $\|u^*\|_\pi$, i.e., an integral form on $B_1\times\cdots\times B_k$. In fact, by (4.4), the forms obtained in this way are exactly the integral forms such that there is a representation (4.10) with a discrete measure μ. These forms are called *nuclear forms* on $B_1\times\cdots\times B_k$. By (4.5), every nuclear form can be written as

$$\chi=\sum_{j=1}^\infty\lambda_j x_{1j}^*\otimes\cdots\otimes x_{kj}^*\tag{4.12}$$

with $\lambda_j\geqslant 0$, $\sum_{j=1}^\infty|\lambda_j|<\infty$ and $\|x_{ij}^*\|\leqslant 1$.

Remark 4.8. Let $\mathcal{I}(B_1,\dots,B_k)$ and $\mathcal{N}(B_1,\dots,B_k)$ be the spaces of integral and nuclear forms, respectively, on $B_1\times\cdots\times B_k$. Both $\mathcal{I}(B_1,\dots,B_k)$ and $\mathcal{N}(B_1,\dots,B_k)$ are Banach spaces (with the natural norms suggested by their definitions, see [57]), and there is an inclusion $\mathcal{N}(B_1,\dots,B_k)\subset$

$\mathcal{I}(B_1,\dots,B_k)$ with the inclusion mapping having norm at most 1. Furthermore, there is a quotient mapping $B_1^*\widehat{\otimes}\cdots\widehat{\otimes}B_k^* \to \mathcal{N}(B_1,\dots,B_k)$. In general, none of these maps is an isomorphism, but there are important cases when one or both are.

Remark 4.9. When $k=2$ (bilinear forms), there are corresponding notions for linear operators.

We say that an operator $T : B_1 \to B_2$ is *integral* if the corresponding bilinear form on $B_1 \times B_2^*$ given by $(x,y^*) \mapsto \langle Tx, y^*\rangle$ is integral. Conversely, a bilinear form $\alpha : B_1 \times B_2 \to \mathbb{R}$ is integral if and only if the corresponding operator $T : B_1 \to B_2^*$ given by $\langle Tx, y\rangle = \alpha(x,y)$ is integral. (When B_2 is reflexive, these are obviously equivalent. In general, see [57, Proposition 3.22].)

Similarly, an operator $T : B_1 \to B_2$ is *nuclear* if $\langle Tx, y^*\rangle = \langle u, x\otimes y^*\rangle$ for some $u \in B_1^*\widehat{\otimes}B_2$. When B_2 is reflexive, this says precisely that corresponding bilinear form on $B_1 \times B_2^*$ given by $(x,y^*) \mapsto \langle Tx, y^*\rangle$ is nuclear.

5. Vector-valued integration

We summarize the definitions of the main types of vector-valued integrals, see e.g. [22], [18], [62] and [57] for details. In order to conform to the rest of this paper, we use probabilistic language and consider the expectation of a B-valued random variable X. (The definitions and results extend to integrals of Banach space valued functions defined on arbitrary measure spaces with only notational changes.)

The *Bochner integral* is a straightforward generalization of the Lebesgue integral to Banach space valued functions. We have the following characterization [22, III.2.22 and III.6.9], cf. Theorem 3.3.

Theorem 5.1. *A random variable $X : (\Omega,\mathcal{F},\mathbb{P}) \to B$ is Bochner integrable if and only if X is Bochner measurable and $\mathbb{E}\,\|X\| < \infty$.* □

As discussed in Section 3, Bochner measurable variables are not enough for all applications. For a more general integral, suppose only that $x^*(X)$ is integrable for every $x^* \in B^*$. (In particular, X is weakly measurable.) Then the linear map

$$T_X : x^* \mapsto x^*(X) \tag{5.1}$$

maps B^* into $L^1(\mathbb{P})$ and by the closed graph theorem, this map is bounded. Hence $x^* \mapsto \mathbb{E}\,x^*(X) = \int x^*(X)\,\mathrm{d}\,\mathbb{P}$ is a bounded linear map $B^* \to \mathbb{R}$, i.e., an element of B^{**}. This element is called the *Dunford integral* of X. We can write the definition as

$$\langle \mathbb{E}\,X, x^*\rangle = \mathbb{E}\langle x^*, X\rangle, \tag{5.2}$$

noting that in general $\mathbb{E}\,X \in B^{**}$.

In the special case that the Dunford integral $\mathbb{E}\,X \in B$, and moreover (see Remark 5.3) $\mathbb{E}(X\mathbf{1}_E) \in B$ for every event E (i.e., every measurable set

$E \subseteq \Omega$), we say that $\mathbb{E}\, X$ is the *Pettis integral* of X. Note that the Pettis integral by definition is an element of B. If X is Bochner integrable then it is Pettis integrable (and Dunford integrable) and the integrals coincide, but the converse does not hold. The examples below show that a Pettis integrable function may fail to be Bochner integrable because of either integrability or measurability problems. (Talagrand [62, Chapter 5] and [59] give characterisations of Pettis integrability, but they are not always easy to apply; it seems that there is no simple necessary and sufficient condition. See further e.g. [24] and [36].)

Remark 5.2. If B is reflexive, i.e. $B^{**} = B$, then every Dunford integrable function is trivially Pettis integrable (and conversely). However, we will not find much use of this, since we will take integrals in tensor products $B^{\hat{\otimes} k}$ or $B^{\check{\otimes} k}$, see Section 6, and such tensor products are typically not reflexive, even if B is reflexive, see for example Theorem 9.2, Example 7.2 and, more generally, [57, Section 4.2].

Remark 5.3. Suppose that X is Dunford integrable. Then ξX is Dunford integrable for every bounded random variable ξ, i.e., for every $\xi \in L^\infty(\mathbb{P})$. Moreover, the adjoint of the bounded linear map $T_X : B^* \to L^1(\mathbb{P})$ given by (5.1) is the map $T_X^* : L^\infty(\mathbb{P}) \to B^{**}$ given by

$$T_X^* \xi = \mathbb{E}(\xi X). \tag{5.3}$$

By definition, X is Pettis integrable if $\mathbb{E}(X\mathbf{1}_E) \in B$, i.e., if $T_X^*(\mathbf{1}_E) \in B$, for every measurable set E; since the simple functions are dense in $L^\infty(\mathbb{P})$, this is equivalent to $T_X^*(\xi) = \mathbb{E}(\xi X) \in B$ for every $\xi \in L^\infty(\mathbb{P})$. Hence, X is Pettis integrable if and only if $T_X^* : L^\infty(\mathbb{P}) \to B$.

Remark 5.4. It follows that if X is Dunford integrable, then X is Pettis integrable if and only if $T_X : B^* \to L^1(\mathbb{P})$ is weak*-weak continuous. (If X is bounded, then T_X is always *sequentially* weak*-weak continuous by dominated convergence; this is not enough, as shown by Examples 5.19 and 11.28.)

Remark 5.5. It follows easily from (5.2) that if X is Dunford integrable and $\xi \in L^\infty(\mathbb{P})$, then $\mathbb{E}(\xi X) = \mathbb{E}(\xi_1 X)$ where $\xi_1 := \mathbb{E}(\xi \mid \mathcal{F}_X)$, where $\mathcal{F}_X$ is the sub-σ-field of $\mathcal{F}$ generated by all $\langle x^*, X\rangle$, $x^* \in B^*$. Hence, X is Pettis integrable if and only if $\mathbb{E}(\xi_1 X) \in B$ for every $\mathcal{F}_X$-measurable ξ_1. Each such ξ_1 is a Borel function of a countable family $(\langle x_i^*, X\rangle)_i$. It follows that the question whether X is Pettis integrable or not depends only on the distribution of X (or equivalently, the joint distribution of $\langle x^*, X\rangle$, $x^* \in B^*$) and not on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Remark 5.6. If X is a B-valued random variable and B is a closed subspace of another Banach space B_1, then X can also be seen as a B_1-valued random variable. It is easily verified (using the Hahn–Banach theorem) that X is Bochner, Dunford or Pettis integrable as a B-valued random variable if and

only if it is so as a B_1-valued random variable, and the expectations $\mathbb{E}\,X$ in B and B_1 coincide in all cases.

By the definition (and discussion) above, X is Dunford integrable if and only if $x^* \mapsto x^*(X)$ is a bounded linear operator $B^* \to L^1(\mathbb{P})$. If X is Pettis integrable, then furthermore this operator is weakly compact, i.e., it maps the unit ball into a relatively weakly compact subset of $L^1(\mathbb{P})$, see [62], [36] or [57, Proposition 3.7]. A subset of $L^1(\mathbb{P})$ is relatively weakly compact if and only if it is uniformly integrable, see [22, Corollary IV.8.11 and Theorem V.6.1], where we recall that a family $\{\xi_\alpha\}$ of random variables is uniformly integrable if

$$\begin{cases} \sup_\alpha \mathbb{E}\,|\xi_\alpha| < \infty \quad \text{and} \\ \sup_\alpha \mathbb{E}\bigl(\mathbf{1}_E|\xi_\alpha|\bigr) \to 0 \quad \text{as } \mathbb{P}(E) \to 0, \end{cases} \tag{5.4}$$

see e.g. [34, Section 5.4] or [38, Lemma 4.10]. This yields the following necessary condition.

Theorem 5.7. *If X is Pettis integrable, then the family $\{x^*(X) : x^* \in B^*, \|x^*\| \leqslant 1\}$ of (real-valued) random variables is uniformly integrable.* □

The converse does not hold, see Examples 5.19 and 11.28, but Huff [36] has shown that it holds if X is weakly a.s. separably valued. (In particular, the converse holds when B is separable.)

Theorem 5.8 (Huff [36])**.** *If $\{x^*(X) : x^* \in B^*, \|x^*\| \leqslant 1\}$ is uniformly integrable and X is weakly a.s. separably valued, then X is Pettis integrable.* □

Remark 5.9. Actually, Huff [36] uses a condition that he calls *separable-like*; the definition given in [36] is somewhat stronger than weak a.s. separability, but it seems likely that he really intended what we call weak a.s. separability, and the proof in [36] uses only weak a.s. separability. See also Stefánsson [59] where weakly a.s. separably valued is called *determined by a separable subspace* and said to be the same as Huff's separable-like. (Stefánsson [59, Theorem 2.8] has extended Theorem 5.8 by weakening the condition of weak a.s. separability, replacing separable by weakly compactly generated, but we shall not use his results, which seem more difficult to apply in our situation.)

Corollary 5.10. *If X is weakly measurable and weakly a.s. separably valued, and further $\bar{\mathbb{E}}\|X\| < \infty$, then X is Pettis integrable.*

Proof. There exists a measurable real-valued Z with $\|X\| \leqslant Z$ and $\mathbb{E}\,Z < \infty$. If $\|x^*\| \leqslant 1$, then $|x^*(X)| \leqslant \|X\| \leqslant Z$, and thus $\{x^*(X) : x^* \in B^*, \|x^*\| \leqslant 1\}$ is uniformly integrable. □

Remark 5.11. Let $1 < p \leqslant \infty$ and assume that $x^*(X) \in L^p(\mathbb{P})$ for every $x^* \in B^*$, i.e., that the map $T_X : B^* \to L^1(\mathbb{P})$ above maps B^* into $L^p(\mathbb{P})$. Then $T_X : B^* \to L^p(\mathbb{P})$ is bounded by the closed graph theorem and thus

the adjoint T_X^*, defined in Remark 5.3 as an operator $L^\infty(\mathbb{P}) \to B^{**}$, extends to $L^q(P)$, where $q \in [1,\infty)$ is the conjugate exponent given by $1/p + 1/q = 1$. Furthermore, (5.3) holds for every $\xi \in L^q(\mathbb{P})$; note that $x^*(\xi X) = \xi x^*(X) \in L^1(\mathbb{P})$ by Hölder's inequality so ξX is Dunford integrable.

If furthermore X is Pettis integrable, then $T_X^* : L^\infty(\mathbb{P}) \to B$ by Remark 5.3, and by continuity this extends to $T_X^* : L^q(\mathbb{P}) \to B$.

Note that if C is the norm of $T_X : B^* \to L^p(\mathbb{P})$, then for every $x^* \in B^*$ with $\|x^*\| \leqslant 1$, $\mathbb{E}\,|x^*(X)|^p = \mathbb{E}\,|T_X(x^*)|^p \leqslant C^p$, which implies that $\{x^*(X) : \|x^*\| \leqslant 1\}$ is uniformly integrable [34, Theorem 5.4.2]. Hence, by Theorem 5.8, if furthermore X is weakly a.s. separably valued, in particular if B is separable, then X is Pettis integrable.

Example 5.12. The standard Brownian motion (the Wiener process) $W(t)$ is a random variable with values in $C[0,1]$. W is easily seen to be Bochner measurable (cf. Corollary 11.15) and $\mathbb{E}\,\|W\| < \infty$; thus W is Bochner integrable (and thus Pettis integrable). The operator $T_X : C[0,1]^* = M[0,1] \to L^1(\mathbb{P})$ is given by

$$T_X(\mu) := \langle \mu, W\rangle = \int_0^1 W(t)\,\mathrm{d}\mu(t). \tag{5.5}$$

T_X obviously maps $C[0,1]^* \to L^p(\mathbb{P})$ for any $p < \infty$, since $\int_0^1 W\,\mathrm{d}\mu$ is Gaussian. (In fact, $\mathbb{E}\,\|W\|^p < \infty$ when $p < \infty$.) The adjoint map $T_X^* : L^q(\mathbb{P}) \to B$, which by Remark 5.11 is defined for every $q > 1$, maps a random variable $\xi \in L^q(\mathbb{P})$ to the function in $C[0,1]$ given by

$$T_X^*\xi(t) = \langle \delta_t, T_X^*\xi\rangle = \langle T_X\delta_t, \xi\rangle = \langle W(t), \xi\rangle = \mathbb{E}\big(\xi W(t)\big). \tag{5.6}$$

Hence, (5.3) says that

$$\mathbb{E}(\xi W)(t) = \mathbb{E}(\xi W(t)). \tag{5.7}$$

In particular, $\mathbb{E}\,W = 0$, as is obvious by symmetry.

We also state another sufficient condition for Pettis integrability that only applies in special Banach spaces.

Theorem 5.13 (Diestel and Uhl [18, Theorem II.3.7])**.** *Suppose that X is Dunford integrable and a.s. separably valued, and that B does not contain a subspace isomorphic to c_0. Then X is Pettis integrable.* □

Example 5.14 below shows, using Remark 5.6, that the condition on B in Theorem 5.13 also is necessary. (Hence, in some sense, Example 5.14 is the canonical example of a Dunford integrable variable that is not Pettis integrable, at least in separable spaces where there is no measurability problem.)

Example 5.14. Let N be a positive integer-valued random variable, and consider the random variable $X := a_N e_N$ in c_0, where $e_n \in c_0$ is the n:th vector in the standard basis and a_n are some real numbers. Let $p_n := \mathbb{P}(N = n)$. It is easily seen that then

(i) X is Bochner integrable if and only if $\sum_{n=1}^\infty |p_n a_n| < \infty$.
(ii) X is Dunford integrable if and only if $\sup |p_n a_n| < \infty$.
(iii) X is Pettis integrable if and only if $p_n a_n \to 0$ as $n \to \infty$.

The integral $\mathbb{E}\, X$ equals $(p_n a_n)_{n=1}^\infty$ in all cases where it is defined. (Thus, when X is Dunford integrable, $\mathbb{E}\, X \in \ell^\infty = c_0^{**}$ but in general $\mathbb{E}\, X$ does not belong to c_0.)

Example 5.15. Let $X := a_N e_N$ as in Example 5.14, but now regarded as an element of $B = \ell^2$. It is easily seen that then

(i) X is Bochner integrable if and only if $\sum_{n=1}^\infty |p_n a_n| < \infty$.
(ii) X is Pettis or Dunford integrable if and only if $\sum_{n=1}^\infty |p_n a_n|^2 < \infty$.

(There is no difference between Dunford and Pettis since B is reflexive.)

The integral $\mathbb{E}\, X$ equals $(p_n a_n)_{n=1}^\infty$ in all cases where it is defined.

Example 5.16. Let again $X := a_N e_N$ as in Example 5.14, but now regarded as an element of $B = \ell^1$. In this case,

(i) X is Bochner integrable $\iff$ X is Pettis integrable $\iff$ X is Dunford integrable $\iff$ $\sum_{n=1}^\infty |p_n a_n| < \infty$.

In fact, let $x^* = (\operatorname{sign}(a_n))_1^\infty \in \ell^\infty = (\ell^1)^*$. Then $\langle x^*, X\rangle = |a_N|$, so if X is Dunford integrable, then $\mathbb{E}\,|a_N| < \infty$, which implies Bochner integrability. The integral $\mathbb{E}\, X$ equals $(p_n a_n)_{n=1}^\infty$.

Example 5.17. Examples 5.14–5.16 are examples of the following general fact: Let N be a positive integer-valued random variable, with $\mathbb{P}(N = n) = p_n$, let $(x_n)_{n=1}^\infty$ be a sequence in a Banach space B, and let $X := x_N$. It is then easy to see the following characterizations, see e.g. [57, Proposition 3.12 and Appendix B] and [17, Chapter IV and p. 44].

(i) X is Bochner integrable if and only if $\sum_{n=1}^\infty p_n x_n$ converges absolutely.
(ii) X is Pettis integrable if and only if $\sum_{n=1}^\infty p_n x_n$ converges unconditionally.
(iii) X is Dunford integrable if and only if the series $\sum_{n=1}^\infty p_n x_n$ is weakly unconditionally Cauchy.

(This example is perhaps clearer if we do not restrict ourselves to probability measures, and regard $\sum_{n=1}^\infty x_n$ as the integral of the function $n \mapsto x_n$ defined on $\mathbb{N}$ equipped with counting measure. The sum converges as a Bochner integral, Pettis integral or Dunford integral if and only if it is absolutely summable, unconditionally summable or weakly unconditionally Cauchy, respectively.)

Example 5.18. Let $B = D[0,1]$, and let, as in Example 3.5, X be the random element of $D[0,1]$ given by $X = \mathbf{1}_{[U,1]}$, where $U \sim \mathrm{U}(0,1)$. Then X is not a.s. separably valued and thus not Bochner measurable; thus $\mathbb{E}\, X$ does not exist as a Bochner integral. (X is also not Borel measurable, at least typically, see Example 3.5.)

On the other hand, it follows from Theorem 15.5 that X is weakly measurable, and since X is bounded, it is Dunford integrable. It is easily verified, using Corollary 13.2 and Fubini's theorem, that if $x^* \in D^*$, then

$$\mathbb{E}\langle x^*, X\rangle = \langle x^*, \mathbf{t}\rangle,$$

where $\mathbf{t}$ denotes the identity function $t \mapsto t$. Hence the Dunford integral $\mathbb{E}\, X = \mathbf{t} \in D[0,1]$. It is similarly seen that if E is any event, then the Dunford integral $\mathbb{E}(X\mathbf{1}_E)$ is the function $t \mapsto \mathbb{P}(\{U \leqslant t\} \cap E)$, which is continuous and thus belongs to D. Hence, X is Pettis integrable.

By Remark 5.6, X is also Pettis integrable as a random variable in $L^\infty[0,1]$, since the hyperplane $\{f \in D[0,1] : f(1-) = f(1)\}$ in $D[0,1]$ also can be seen as a subspace of $L^\infty[0,1]$. (Cf. [62, Example 4-2-4a)].)

To find examples of bounded Dunford integrable random variables that are not Pettis integrable is more difficult and technical. Note that by Theorem 5.8, such random variables cannot be weakly a.s. separably valued. We give one example from Fremlin and Talagrand [29], omitting the (quite complicated) details. See Example 11.28 for another example.

Example 5.19 (Fremlin and Talagrand [29])**.** Let $\Omega = \{0,1\}^\infty$ with the infinite product measure $\mu = \left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1\right)^\infty$ (this is the Haar measure if we regard Ω as the compact group $\mathbb{Z}_2^\infty$), and let $X : \Omega \to \ell^\infty$ be the inclusion. Then X is a random variable with values in ℓ^∞ such that the coordinates X_n are i.i.d. Be(1/2). It is shown in [29] and [62, Chapter 13] (by slightly different arguments) that X is *not* weakly measurable on Ω (with the product σ-field = the Borel σ-field), but that the measure μ can be extended to a larger σ-field making X weakly measurable. More precisely, it is easily seen that $(\ell^\infty)^* = \ell^1 \oplus c_0^\perp$, and if $x^* \in \ell^1 \subset (\ell^\infty)^*$ is given by $(a_n)_1^\infty \in \ell^1$, then $x^*(X) = \sum_1^\infty a_n X_n$ where X_n as said above are i.i.d. Be(1/2); clearly $x^*(X)$ is measurable in this case. The extension of μ constructed in [29] and [62] is such that $x^*(X)$ is a.s. constant if $x^* \in c_0^\perp$. Hence $x^*(X)$ is measurable in this case too, and by linearity for every $x^* \in (\ell^\infty)^*$, so X is weakly measurable. Moreover, the extension is such that in the particular case that $x^* \in c_0^\perp$ is a multiplicative linear functional, $x^*(X) = 1$ a.s.

Consequently, using this extension of μ, X is bounded and weakly measurable, and thus Dunford integrable. However, if X had a Pettis integral $y = \mathbb{E}\, X \in \ell^\infty$, then $y_n = \mathbb{E}\, X_n = \frac{1}{2}$ for each n, since $X_n \sim \mathrm{Be}(\frac{1}{2})$, and thus $y = (\frac{1}{2}, \frac{1}{2}, \dots)$. However, if x^* is a multiplicative linear functional in $c_0^\perp$, then $x^*(y) = \mathbb{E}\, x^*(X) = 1$, a contradiction. Consequently, X is not Pettis integrable.

By Theorem 5.8, X is not weakly a.s. separably valued.

To summarize, we have defined three types of integrals. The Bochner integral is the most convenient, when it exists, but the requirement of Bochner measurablility is too strong for many applications in non-separable spaces. The Pettis integral is more general, and will be our main tool in such cases; it also requires only a weaker integrability condition. The Dunford integral

is even more general, but in general it is an element the bidual B^{**} instead of B, which makes it less useful.

Note also that all three integrals are linear and behave as expected under bounded linear operators: If $T : B \to B_1$ and X is integrable in one of these senses, then $TX \in B_1$ is integrable in the same sense, and $\mathbb{E}(TX) = T(\mathbb{E}\,X)$ ($\mathbb{E}(TX) = T^{**}(\mathbb{E}\,X)$ for Dunford integrals).

Remark 5.20. It can be shown that the space of Bochner integrable B-valued random variables equals $L^1(\mathbb{P})\hat{\otimes}B$ [57, Example 2.19], while $L^1(\mathbb{P})\check{\otimes}B$ is the completion of the space of Bochner measurable Pettis integrable B-valued random variables [57, Proposition 3.13]. (If B is separable, this is by Theorem 3.3 just the completion of the space of Pettis integrable random variables; typically the latter space is not complete, so it is necessary to take the completion.)

Remark 5.21. There are several further definitions of integrals of Banach space valued functions, see for example [7], [35], [28], although only some of these definitions work on a general probability space as required here (for example the Birkhoff integral that lies between the Bochner and Pettis integrals). These integrals too could be used to define moments as in the next section, but we do not know of any properties of them that make them more useful for our purposes than the three integrals defined above, so we do not consider them.

6. Moments

If B is a Banach space and X is a B-valued random variable, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and further k is a positive integer, we define the *projective k:th moment* of X as the expectation $\mathbb{E}\,X^{\otimes k} = \int X^{\otimes k}\,\mathrm{d}\,\mathbb{P}$ whenever this expectation (integral) exists in the projective tensor product $B^{\hat{\otimes}k}$. The expectation can here be taken in any of the three senses defined in Section 5; hence we talk about the moment existing in Bochner sense, Pettis sense or Dunford sense. Note that if the k:th moment exists in Bochner or Pettis sense, then it is an element of $B^{\hat{\otimes}k}$, but if it exists in Dunford sense, then it is an element of $(B^{\hat{\otimes}k})^{**}$ and may be outside $B^{\hat{\otimes}k}$.

Similarly, we can regard $X^{\otimes k}$ as an element of the injective tensor product $B^{\check{\otimes}k}$ and take the expectation in that space. We thus define the *injective k:th moment* of X as the expectation $\mathbb{E}\,X^{\otimes k} = \int X^{\otimes k}\,\mathrm{d}\,\mathbb{P}$ whenever this expectation (integral) exists in the injective tensor product $B^{\check{\otimes}k}$. Again, this can exist in Bochner sense, Pettis sense or Dunford sense; in the first two cases it is an element of $B^{\check{\otimes}k}$, but in the third case it is an element of $(B^{\check{\otimes}k})^{**}$.

Example 6.1. The first moment (projective or injective; there is no difference when $k = 1$) is the expectation $\mathbb{E}\,X$, which is an element of B when it exists in Bochner or Pettis sense, and an element of B^{**} when it exists in Dunford sense. The examples in Section 5 show some cases.

We may for clarity or emphasis denote $X^{\otimes k}$ by $X^{\hat{\otimes} k}$ when we regard it as an element of $B^{\hat{\otimes} k}$ and by $X^{\check{\otimes} k}$ when we regard it as an element of $B^{\check{\otimes} k}$. In particular, we distinguish between the projective and injective moments by writing them as $\mathbb{E}\, X^{\hat{\otimes} k} \in B^{\hat{\otimes} k}$ and $\mathbb{E}\, X^{\check{\otimes} k} \in B^{\check{\otimes} k}$. They are related by the following simple result.

Theorem 6.2. *If the projective k:th moment $\mathbb{E}\, X^{\hat{\otimes} k}$ exists in one of the senses above, then the injective k:th moment exists too, in the same sense, and is given by*

$$\mathbb{E}\, X^{\check{\otimes} k} = \iota(\mathbb{E}\, X^{\hat{\otimes} k}) \in B^{\check{\otimes} k} \tag{6.1}$$

in the Bochner or Pettis case and

$$\mathbb{E}\, X^{\check{\otimes} k} = \iota^{**}(\mathbb{E}\, X^{\hat{\otimes} k}) \in (B^{\check{\otimes} k})^{**} \tag{6.2}$$

in the Dunford case.

Proof. The identity map on $B^{\otimes k}$ extends to the continuous linear map $\iota : B^{\hat{\otimes} k} \to B^{\check{\otimes} k}$. □

We next consider the problem of deciding when these moments exist, in the different senses. There are six different cases to consider. We shall see (Theorem 6.7 and the examples in Section 7) that the conditions for existence differ for five of them, for both measurability and integrability reasons. This multiplicity of cases may be bewildering, but in many applications there is no problem. If B is separable (or X is a.s. separably valued) there is no problem with measurability (Theorem 3.3) and if further $\mathbb{E}\, \|X\|^k < \infty$, then both moments exist in the strongest sense (Theorem 6.7), and thus in all senses.

We begin by considering conditions for the existence of moments in the strongest sense, i.e., as Bochner integrals.

Lemma 6.3. *If X is Bochner measurable in B, then $X^{\otimes k}$ is Bochner measurable in $B^{\hat{\otimes} k}$, for every $k \geqslant 1$. In particular, then $\alpha(X, \dots, X)$ is measurable for any bounded k-linear form $\alpha \in L(B^k; \mathbb{R})$.*

Proof. The (non-linear) mapping $x \mapsto x^{\otimes k}$ is continuous $B \to B^{\hat{\otimes} k}$. The final claim follows by Theorem 4.6. □

Remark 6.4. There is no general corresponding result for weakly measurable X in the projective tensor product $B^{\hat{\otimes} k}$, see Example 7.4. If B is separable, or if X is a.s. separably valued, there is no problem since then X is Bochner measurable by Theorem 3.3, but weak a.s. separability is not enough by Example 7.4. Nevertheless, $D[0,1]$ is an example of a non-separable space where $X^{\otimes k}$ is weakly measurable in the projective tensor product for every weakly measurable X, see Corollary 15.9.

Lemma 6.5. *The following are equivalent, for any $k \geqslant 1$:*

(i) *$X^{\otimes k}$ is a.s. separably valued in $B^{\hat{\otimes} k}$.*

(ii) $X^{\otimes k}$ *is a.s. separably valued in* $B^{\check{\otimes} k}$.
(iii) X *is a.s. separably valued in* B.

Proof. (iii) $\implies$ (i) as in the proof of Lemma 6.3 and (i) $\implies$ (ii) since $\iota : B^{\hat{\otimes} k} \to B^{\check{\otimes} k}$ is continuous. Hence it remains to prove (ii) $\implies$ (iii).

Let A be a separable subspace of $B^{\check{\otimes} k}$ such that $X^{\otimes k} \in A$ a.s. Then there exists a countable family of elementary tensors $F = \{e_{i1} \otimes \cdots \otimes e_{ik}\}$ such that A is included in the closed linear span of F. Let $B_1 \subseteq B$ be the closed linear span of $\{e_{ij}\}_{ij}$. Then B_1 is separable, and if $x^* \perp B_1$, then $x^*(e_{ij}) = 0$ for all i and j, and thus $x^* \otimes \cdots \otimes x^* \perp F$; by linearity and continuity, this extends to the closed linear span of F, and thus $x^* \otimes \cdots \otimes x^* \perp A$.

Hence, if $x^{\otimes k} \in A$, then $\langle x^*, x\rangle^k = \langle (x^*)^{\otimes k}, x^{\otimes k}\rangle = 0$ and thus $x^* \perp x$ for every $x^* \perp B_1$, which implies that $x \in B_1$. Consequently, $X \in B_1$ a.s. $\square$

Remark 6.6. Lemma 6.5 includes a partial converse to Lemma 6.3, considering only the a.s. separably valued condition. There is no complete converse to Lemma 6.3 since $X^{\otimes 2}$ may be Bochner measurable even if X is not; this happens even in the one-dimensional case $B = \mathbb{R}$, as shown by the trivial example when X is a non-measurable function such that $X = \pm 1$ everywhere; then $X^{\otimes 2} = X^2 = 1$ is measurable. We will thus usually assume that X is at least weakly measurable.

It is now easy to characterise when the moments exist as Bochner integrals.

Theorem 6.7. *Suppose that X is weakly measurable. Then the following are equivalent.*

(i) *The projective moment* $\mathbb{E}\, X^{\hat{\otimes} k}$ *exists in Bochner sense.*
(ii) *The injective moment* $\mathbb{E}\, X^{\check{\otimes} k}$ *exists in Bochner sense.*
(iii) $\mathbb{E}\,\|X\|^k < \infty$ *and X is a.s. separably valued.*

Proof. By Theorem 5.1 and Lemma 6.5, (i) and (ii) both imply that X is a.s. separably valued: Hence it suffices to consider the case when X is a.s. separably valued, so by Theorem 3.3 X is Bochner measurable. By Lemma 6.3, $X^{\otimes k}$ is Bochner measurable, and the result follows by Theorem 5.1, since $\|X^{\otimes k}\|_{B^{\hat{\otimes} k}} = \|X^{\otimes k}\|_{B^{\check{\otimes} k}} = \|X\|^k$. $\square$

We turn to Dunford integrals. We first give a simple result on the existence of the weak k:th moment (1.1). Cf. Remark 5.11.

Lemma 6.8. *Suppose that X is weakly measurable. Then the following are equivalent.*

(i) *The weak k:th moment* $\mathbb{E}\big(x_1^*(X)\cdots x_k^*(X)\big)$ *exists for every* $x_1^*, \dots, x_k^* \in B^*$.
(ii) $\mathbb{E}\,|x^*(X)|^k < \infty$ *for every* $x^* \in B^*$.
(iii) $\sup\{\mathbb{E}\,|x^*(X)|^k : \|x^*\| \leqslant 1\} < \infty$.
(iv) $T_X : x^* \mapsto x^*(X)$ *is a bounded operator* $B^* \to L^k(\mathbb{P})$.

In this case, $\mathbb{E}\bigl|x_1^*(X)\cdots x_k^*(X)\bigr|$ *is bounded for* $x_1^*,\dots,x_k^*$ *in the unit ball of* B^*.

Proof. (i) $\iff$ (ii): If (i) holds, then (ii) follows by choosing $x_1^* = \cdots = x_k^* = x^*$. The converse follows by Hölder's inequality.

(iii) $\iff$ (iv): By definition.

(ii) $\iff$ (iv): If (ii) holds, then $T_X : B^* \to L^k(\mathbb{P})$ is bounded by the closed graph theorem. The converse is trivial.

The final claim follows by (iii) and Hölder's inequality. □

In all cases we know, the existence of the weak k:th moment is equivalent to the existence of the injective k:th moment in Dunford sense. We have, however, failed to prove this in full generality and suspect that there are counterexamples. We thus give a theorem with some technical sufficient conditions, and leave it as an open problem whether the theorem holds more generally. (This is, at least for bounded X, equivalent to whether X weakly measurable implies $X^{\otimes k}$ weakly measurable in $B^{\check{\otimes} k}$; cf. Remark 6.4 which shows that this does not hold for the projective tensor product.)

We first state a lemma.

Lemma 6.9. *Suppose that* X *is weakly measurable and a.s. separably valued. Then* $\langle x^*, X\rangle$ *is jointly measurable on* $B^* \times \Omega$, *where* B^* *is given the Borel* σ*-field for the weak* topology.*

Proof. X is Bochner measurable by Theorem 3.3 and thus there is a sequence X_n of measurable simple random variables in B such that $X_n \to X$ a.s., see Remark 3.2. Then each $\langle x^*, X_n\rangle$ is jointly measurable on $B^* \times \Omega$, and thus $\langle x^*, X\rangle$ is jointly measurable. □

Theorem 6.10. (i) *If the injective* k*:th moment* $\mathbb{E}\, X^{\check{\otimes} k}$ *exists in Dunford sense, then the weak* k*:th moment* $\mathbb{E}\bigl(x_1^*(X)\cdots x_k^*(X)\bigr)$ *exists for every* $x_1^*,\dots,x_k^* \in B^*$, *and*

$$\mathbb{E}\bigl(x_1^*(X)\cdots x_k^*(X)\bigr) = \langle \mathbb{E}\, X^{\check{\otimes} k}, x_1^* \otimes \cdots \otimes x_k^*\rangle. \tag{6.3}$$

Furthermore, $\sup\{\mathbb{E}\,|x^*(X)|^k : \|x^*\| \leqslant 1\} < \infty$.

(ii) *Suppose that* X *is weakly measurable, and that one of the following additional condition holds.*
   (a) B *is separable.*
   (b) X *is a.s. separably valued.*
   (c) *Every integral* k*-linear form* $B^k \to \mathbb{R}$ *is nuclear.*

   Then the injective moment $\mathbb{E}\, X^{\check{\otimes} k}$ *exists in Dunford sense if and only if the weak* k*:th moment exists, i.e., if and only if* $\mathbb{E}\,|x^*(X)|^k < \infty$ *for every* $x^* \in B^*$.

Proof. (i): Directly from the definition of the Dunford integral, since $x_1^* \otimes \cdots \otimes x_k^*$ is a continuous linear functional on $B^{\check{\otimes} k}$. The final claim follows by Lemma 6.8.

(ii): By (i) and Lemma 6.8, it remains to show that if the weak k:th moment exists, then $\chi(X^{\otimes k})$ is an integrable random variable, and in particular measurable, for every $\chi \in (B^{\check{\otimes} k})^*$. By Theorem 4.7, χ is represented by a signed measure $\mu \in M(K^k)$, where K is the closed unit ball of B^*, and

$$\chi(X^{\otimes k}) = \int_{K^k} x_1^*(X) \cdots x_k^*(X) \, \mathrm{d}\mu(x_1^*, \dots, x_k^*). \tag{6.4}$$

(a) obviously is a special case of (b). If (b) holds, then $\langle x^*, X\rangle$ is jointly measurable on $B^* \times \Omega$ by Lemma 6.9. Hence, $x_1^*(X) \cdots x_k^*(X)$ is jointly measurable on $K^k \times \Omega$, and thus we can take expectations in (6.4) and apply Fubini's theorem, yielding

$$\mathbb{E}\, \chi(X^{\otimes k}) = \int_{K^k} \mathbb{E}\big(x_1^*(X) \cdots x_k^*(X)\big) \, \mathrm{d}\mu(x_1^*, \dots, x_k^*); \tag{6.5}$$

note that $\mathbb{E}\big|x_1^*(X) \cdots x_k^*(X)\big|$ is bounded on K^k by Lemma 6.8, so the double integral is absolutely convergent.

If (c) holds, then the integral form χ is nuclear and thus, by (4.12),

$$\chi = \sum_{n=1}^{\infty} \lambda_n x_{1n}^* \otimes \cdots \otimes x_{kn}^* \tag{6.6}$$

where $\lambda_n \geqslant 0$, $\sum_{n=1}^{\infty} \lambda_n < \infty$ and each $x_{in}^* \in K$. Consequently, using (4.11),

$$\big\langle \chi, X^{\otimes k}\big\rangle = \sum_{n=1}^{\infty} \lambda_n x_{1n}^*(X) \cdots x_{kn}^*(X), \tag{6.7}$$

which is integrable by Lemma 6.8. Taking expectations in (6.7), we see that (6.5) holds in this case too, now for the finite discrete measure

$$\mu := \sum_{n=1}^{\infty} \lambda_n \delta_{(x_{1n}^*, \cdots, x_{kn}^*)}. \tag{6.8}$$

□

Remark 6.11. We do not know any characterization of the Banach spaces B such that every integral k-linear form is nuclear. For $k = 2$, this can be translated to operators $B \to B^*$; a sufficient condition then is that B^* has the Radon–Nikodým property and the approximation property, see [57, Theorem 5.34].

Corollary 6.12. *Suppose that X and Y are weakly measurable B-valued random variables, and that one of the following conditions holds:*

(a) *B is separable.*
(b) *X and Y are a.s. separably valued.*
(c) *Every integral k-linear form $B^k \to \mathbb{R}$ is nuclear.*

Then (1.3) *holds if and only $\mathbb{E}\, X^{\check{\otimes} k} = \mathbb{E}\, Y^{\check{\otimes} k}$, with the injective k:th moments existing in Dunford sense. (In other words, the injective moment is determined by the weak moment.)*

Proof. If $\mathbb{E}\, X^{\check{\otimes}k} = \mathbb{E}\, Y^{\check{\otimes}k}$, then the weak moments are equal by (6.3).

Conversely, suppose that the weak moments exist and are equal. By Theorem 6.10, the injective moments $\mathbb{E}\, X^{\check{\otimes}k}$ and $\mathbb{E}\, Y^{\check{\otimes}k}$ exist in Dunford sense. Moreover, the proof of Theorem 6.10 shows that for any $\chi \in (B^{\check{\otimes}k})^*$, (6.5) holds for some signed measure μ, which shows that if the weak moments are equal, then

$$\langle \chi, \mathbb{E}\, X^{\check{\otimes}k} \rangle = \mathbb{E}\langle \chi, X^{\check{\otimes}k} \rangle = \mathbb{E}\langle \chi, Y^{\check{\otimes}k} \rangle = \langle \chi, \mathbb{E}\, Y^{\check{\otimes}k} \rangle$$

for every χ and thus $\mathbb{E}\, X^{\check{\otimes}k} = \mathbb{E}\, Y^{\check{\otimes}k}$ in $(B^{\check{\otimes}k})^{**}$. □

Recall that $B^{\check{\otimes}k}$ can be regarded as a subspace of $L((B^*)^k; \mathbb{R})$, the bounded k-linear forms on B^*. This leads to the following interpretation of the injective k:th moment when it exists in Pettis (or Bochner) sense, and thus is an element of $B^{\check{\otimes}k}$, which again shows that the injective k:th moment is essentially the same as the weak moment defined by (1.1).

Theorem 6.13. *If X is a B-valued random variable such that the injective k:th moment $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Pettis sense, then $\mathbb{E}\, X^{\check{\otimes}k} \in B^{\check{\otimes}k}$ is the k-linear form on B^**

$$(x_1^*, \dots, x_k^*) \mapsto \mathbb{E}\big(x_1^*(X) \cdots x_k^*(X)\big).$$

Proof. When $\mathbb{E}\, X^{\check{\otimes}k} \in B^{\check{\otimes}k}$ is regarded as a k-linear form, its value at $(x_1^*, \dots, x_k^*) \in (B^*)^k$ equals $\langle \mathbb{E}\, X^{\check{\otimes}k}, x_1^* \otimes \cdots \otimes x_k^* \rangle$, and the result follows by (6.3). □

This yields a simpler version of Corollary 6.12, assuming that the moments exist in Pettis sense.

Corollary 6.14. *Suppose that X and Y are B-valued random variables such that the injective k:th moments $\mathbb{E}\, X^{\check{\otimes}k}$ and $\mathbb{E}\, Y^{\check{\otimes}k}$ exist in Pettis sense. Then* (1.3) *holds if and only $\mathbb{E}\, X^{\check{\otimes}k} = \mathbb{E}\, Y^{\check{\otimes}k}$. (In other words, the injective moment is determined by the weak moment.)* □

For the projective k:th moment in Dunford sense, there is a general similar equivalence, now using arbitrary bounded k-linear forms on B. We have no simple necessary and sufficient condition for the existence, but we give a sufficient condition which is necessary in at least some cases (Example 7.1), but not in others (Example 7.3).

Theorem 6.15. *The following are equivalent.*

(i) *The projective moment $\mathbb{E}\, X^{\hat{\otimes}k}$ exists in Dunford sense.*
(ii) *The moment $\mathbb{E}\, \alpha(X, \dots, X)$ exists for every bounded k-linear form $\alpha : B^k \to \mathbb{R}$.*

In this case,

$$\mathbb{E}\, \alpha(X, \dots, X) = \langle \mathbb{E}\, X^{\hat{\otimes}k}, \alpha \rangle \tag{6.9}$$

for every $\alpha \in L(B^k; \mathbb{R})$.

Moreover, if $\bar{\mathbb{E}}\|X\|^k < \infty$ *and* $\alpha(X,\dots,X)$ *is measurable for every* $\alpha \in L(B^k;\mathbb{R})$*, then* (i) *and* (ii) *hold.*

Proof. The equivalence of (i) and (ii) is an immediate consequence of the definition of the Dunford integral and Theorem 4.6, and so is (6.9).

Furthermore, $|\alpha(X,\dots,X)| \leqslant \|\alpha\|\|X\|^k$. Hence, if $\bar{\mathbb{E}}\|X\|^k < \infty$, then (ii) holds provided $\alpha(X,\dots,X)$ is measurable. □

Remark 6.16. Although we use multilinear forms α in (1.2) and Theorem 6.15, we are only interested in the values $\alpha(x,\dots,x)$ on the diagonal. This restriction to the diagonal defines a function $\tilde{\alpha}(x) : B \to \mathbb{R}$, which is a quadratic form for $k = 2$, a cubic form for $k = 3$, etc., and (1.2) can be expressed as $\mathbb{E}\,\tilde{\alpha}(X) = \mathbb{E}\,\tilde{\alpha}(Y)$ for all such forms $\tilde{\alpha}$.

Note also that it suffices to consider *symmetric* multilinear forms α, since we always may replace α by its symmetrization.

Corollary 6.17. *Suppose that* X *and* Y *are* B*-valued random variables. Then* (1.2) *holds, with finite and well-defined expectations for every* α*, if and only if the projective* k*:th moments* $\mathbb{E}\,X^{\hat{\otimes}k}$ *and* $\mathbb{E}\,Y^{\hat{\otimes}k}$ *exist in Dunford sense and* $\mathbb{E}\,X^{\hat{\otimes}k} = \mathbb{E}\,Y^{\hat{\otimes}k}$*.*

Proof. An immediate consequence of Theorem 6.15, together with Theorem 4.6. □

The problem whether (1.2) and (1.3) are equivalent is thus reduced to the problem whether $\mathbb{E}\,X^{\check{\otimes}k} = \mathbb{E}\,Y^{\check{\otimes}k}$ implies $\mathbb{E}\,X^{\hat{\otimes}k} = \mathbb{E}\,Y^{\hat{\otimes}k}$, at least if we assume that the projective moments exist in Dunford sense, and that one of the additional assumptions in Corollary 6.12 or 6.14 holds. By (6.2), it then is sufficient that ι^{**} is injective. However, in applications we prefer not to use the bidual (recall that tensor products typically are not reflexive, even if B is reflexive, see Remark 5.2); we thus prefer to use moments in Pettis or Bochner sense. For these moments, the question whether $\mathbb{E}\,X^{\hat{\otimes}k} = \mathbb{E}\,Y^{\hat{\otimes}k} \iff \mathbb{E}\,X^{\check{\otimes}k} = \mathbb{E}\,Y^{\check{\otimes}k}$ (for arbitrary X and Y in a given Banach space B) is by (6.1) almost equivalent to whether $\iota : B^{\hat{\otimes}k} \to B^{\check{\otimes}k}$ is injective; this will be studied in Section 8.

Remark 6.18. We write "almost", because $\mathbb{E}\,X^{\hat{\otimes}k}$ is a symmetric tensor, so we are really only interested in whether ι is injective on the subspace of symmetric tensors in $B^{\hat{\otimes}k}$ (i.e., on the symmetric tensor product). We conjecture that ι is injective on this subspace if and only it is injective on the full tensor product $B^{\hat{\otimes}k}$, but as far as we know this question has not been investigated and we leave it as an open problem.

We turn to considering conditions for the existence of moments in Pettis sense. We only give a result for the injective moments, corresponding to Theorem 6.10, since we do not know any corresponding result for projective moments.

Theorem 6.19. (i) *If the injective k:th moment $\mathbb{E}\, X^{\check{\otimes} k}$ exists in Pettis sense, then $\{|x^*(X)|^k : x^* \in B^*,\ \|x^*\| \leqslant 1\}$ is uniformly integrable.*

(ii) *Suppose that X is weakly measurable, and that one of the following additional condition holds.*
   (a) *B is separable.*
   (b) *X is a.s. separably valued.*
   (c) *Every integral k-linear form $B^k \to \mathbb{R}$ is nuclear, and $X^{\otimes k}$ is weakly a.s. separably valued in $B^{\check{\otimes} k}$.*

   Then the injective moment $\mathbb{E}\, X^{\check{\otimes} k}$ exists in Pettis sense if and only if $\{|x^*(X)|^k : x^* \in B^*,\ \|x^*\| \leqslant 1\}$ is uniformly integrable.

Proof. (i): By Theorem 5.7, since $\langle (x^*)^{\otimes k}, X^{\otimes k}\rangle = x^*(X)^k$ and $(x^*)^{\otimes k} \in (B^{\check{\otimes} k})^*$ with norm $\|x^*\|_{B^*}^k \leqslant 1$ when $\|x^*\|_{B^*} \leqslant 1$.

(ii): We shall modify the proof of Theorem 6.10. Let $\chi \in (B^{\check{\otimes} k})^*$ with $\|\chi\| \leqslant 1$. As in the proof of Theorem 6.10, by Theorem 4.7 there exists a signed measure $\mu \in M(K^k)$, where K is the closed unit ball of B^*, such that (6.4) holds; further $\|\mu\| = \|\chi\| \leqslant 1$. Taking absolute values, and replacing μ by $|\mu|$, we obtain

$$\bigl|\chi(X^{\otimes k})\bigr| \leqslant \int_{K^k} \bigl|x_1^*(X)\cdots x_k^*(X)\bigr|\,\mathrm{d}\mu(x_1^*,\dots,x_k^*). \tag{6.10}$$

If (a) or (b) holds, then $\langle x^*, X\rangle$ is jointly measurable on $B^*\times\Omega$ by Lemma 6.9; thus we can take expectations in (6.10) and apply Fubini's theorem, yielding

$$\mathbb{E}\bigl|\chi(X^{\otimes k})\bigr| \leqslant \int_{K^k} \mathbb{E}\bigl|x_1^*(X)\cdots x_k^*(X)\bigr|\,\mathrm{d}\mu(x_1^*,\dots,x_k^*). \tag{6.11}$$

If (c) holds, then the integral form χ is nuclear and thus (6.6) holds, with $\lambda_n \geqslant 0$, $\sum_{n=1}^\infty \lambda_n < \infty$ and each $x_{in}^* \in K$. We do not know whether the nuclear norm of χ always equals the integral norm $\|\chi\|$ when (c) holds, but, at least, the open mapping theorem implies that there exists a constant C (possibly depending on B) such that we can choose a representation (6.6) with $\sum_n \lambda_n \leqslant C$. Then (6.7) holds, and taking absolute values and expectations we see that (6.11) holds in this case too, for the measure (6.8), which satisfies $\|\mu\| = \sum_n \lambda_n \leqslant C$.

Consequently, in both cases (6.11) holds, with $\|\mu\| \leqslant C$ (where $C = 1$ in cases (a)–(b)). By the arithmetic-geometric inequality, (6.11) implies

$$\begin{aligned}
\mathbb{E}\bigl|\chi(X^{\otimes k})\bigr| &\leqslant \int_{K^k} \mathbb{E}\Bigl(\frac{1}{k}\sum_{i=1}^k |x_i^*(X)|^k\Bigr)\,\mathrm{d}\mu(x_1^*,\dots,x_k^*)\\
&= \frac{1}{k}\sum_{i=1}^k \int_{K^k} \mathbb{E}\,|x_i^*(X)|^k\,\mathrm{d}\mu(x_1^*,\dots,x_k^*)\\
&\leqslant C\sup\bigl\{\mathbb{E}\,|x^*(X)|^k : x^*\in K\bigr\},
\end{aligned} \tag{6.12}$$

and applying (6.12) to $\mathbf{1}_E X$, for an arbitrary event $E \in \mathcal{F}$, we obtain

$$\mathbb{E}\bigl(\mathbf{1}_E|\chi(X^{\otimes k})|\bigr) = \mathbb{E}\bigl|\chi(\mathbf{1}_E X^{\otimes k})\bigr| \leqslant C \sup_{x^*\in K} \mathbb{E}\bigl(\mathbf{1}_E|x^*(X)|^k\bigr). \tag{6.13}$$

This holds for any $\chi \in (B^{\check\otimes k})^*$ with $\|\chi\| \leqslant 1$. Hence, it follows from (6.12)–(6.13) and the condition (5.4) for uniform integrability that if $\{|x^*(X)|^k : x^* \in K\}$ is uniformly integrable, then also $\{\langle \chi, X^{\otimes k}\rangle : \chi \in (B^{\check\otimes k})^*, \|\chi\| \leqslant 1\}$ is uniformly integrable. Furthermore, $X^{\check\otimes k}$ is weakly a.s. separably valued, by (a)–(b) and Lemma 6.5 or assumption in (c). The result now follows by Theorem 5.8 applied to $X^{\check\otimes k}$, □

Remark 6.20. Changing the norm in B to an equivalent one will not change the tensor products $B^{\hat\otimes k}$ and $B^{\check\otimes k}$ (except for a change of norms), and the existence (in any of the three senses above) and values of the moments $\mathbb{E}\, X^{\hat\otimes k}$ and $\mathbb{E}\, X^{\check\otimes k}$ will not be affected.

Remark 6.21. The moments behave as expected under linear transformations. If X is a B-valued random variable and $T : B \to B_1$ is a bounded linear map into another Banach space B_1, then TX is a B_1-valued random variable, which is [Borel, weakly, Bochner] measurable if X is. If the moment $\mathbb{E}\, X^{\hat\otimes k}$ or $\mathbb{E}\, X^{\check\otimes k}$ exists in any of the three senses above, then $\mathbb{E}(TX)^{\hat\otimes k}$ or $\mathbb{E}(TX)^{\check\otimes k}$ exists in the same sense; moreover, for moments in Bochner or Pettis sense $\mathbb{E}(TX)^{\hat\otimes k} = T^{\hat\otimes k}(\mathbb{E}\, X^{\hat\otimes k})$, and for moments in Dunford sense $\mathbb{E}(TX)^{\hat\otimes k} = (T^{\hat\otimes k})^{**}(\mathbb{E}\, X^{\hat\otimes k})$, and similarly for injective moments.

Remark 6.22. If X is a B-valued random variable and B is a closed subspace of another Banach space B_1, then X can also be seen as a B_1-valued random variable. For the injective tensor product, then $B^{\check\otimes k}$ is a closed subspace of $B_1^{\check\otimes k}$, see Remark 4.4, and thus by Remark 5.6, the injective moment $\mathbb{E}\, X^{\check\otimes k}$ exists in $B^{\check\otimes k}$ (or $(B^{\check\otimes k})^{**}$) in any of the three senses if and only if it exists in $B_1^{\check\otimes k}$ (or $(B_1^{\check\otimes k})^{**}$) in the same sense; moreover, then the value of the moment in the two spaces coincide.

For the projective moments, the situation is more complicated since Remark 4.4 does not hold for the projective tensor product. If we consider moments in Bochner sense (and assume that X is weakly measurable), then by Theorem 6.7 $\mathbb{E}\, X^{\hat\otimes k}$ exists in $B^{\hat\otimes k}$ if and only it exists in $B_1^{\hat\otimes k}$. However, for projective moments in Pettis sense, we can in general only say that if the moment $\mathbb{E}\, X^{\hat\otimes k}$ exists in $B^{\hat\otimes k}$, then it exists in $B_1^{\hat\otimes k}$, and the values are the same (by Remark 6.21 applied to the inclusion map); we shall see in Example 7.3 that the converse does not hold. (And similarly for Dunford sense, where we have to consider the biduals.) This shows that when considering projective moments of a Banach space valued random variable, we may have to be careful to specify which Banach space we are using.

If B is a complemented subspace of B_1 there is no problem: then there is a bounded projection $P : B_1 \to B$, and it follows from Remark 6.21 that

also the projective moments exist for X as a B-valued random variable if and only if they exist for X as a B_1-valued random variable.

Remark 6.23. We may define moments also in other tensor products (not considered in the present paper) in the same way; one example is to take the Hilbertian tensor product in Example 4.2 when B is a Hilbert space. When the projective k:th moment exists, these moments too are given by mapping the projective k:th moment $\mathbb{E}\, X^{\hat{\otimes}k} \in B^{\hat{\otimes}k}$ to the chosen tensor product as in (6.1) or (6.2). This is one reason to take the projective k:th moment as the standard k:th moment, when it exists.

Remark 6.24. It is possible to define mixed moments of random variables $X_1, \dots, X_k$ with values in possibly different Banach spaces $B_1, \dots, B_k$ in the same way, by taking the expectation of $X_1 \otimes \cdots \otimes X_k$ in $B_1 \hat{\otimes} \cdots \hat{\otimes} B_k$ or $B_1 \check{\otimes} \cdots \check{\otimes} B_k$. Analoguous results hold, but are left to the reader. In most cases, we can consider $X = (X_1, \dots, X_k)$ in the direct sum $B := B_1 \oplus \cdots \oplus B_k$ and take its moments (provided they exist); they contain the mixed moments as components, and we are reduced to the case treated above of moments of a single variable. For example, with $k = 2$, $(B_1 \oplus B_2)^{\hat{\otimes}2}$ is the direct sum of $B_1 \hat{\otimes} B_1$, $B_1 \hat{\otimes} B_2$, $B_2 \hat{\otimes} B_1$ and $B_2 \hat{\otimes} B_2$, and the components of $\mathbb{E}\, X^{\otimes 2}$ in these subspaces are $\mathbb{E}\, X_1^{\otimes 2}$, $\mathbb{E}(X_1 \otimes X_2)$, $\mathbb{E}(X_2 \otimes X_1)$ and $\mathbb{E}\, X_2^{\otimes 2}$, where the two mixed moments are the same, using the natural isomorphism $B_1 \hat{\otimes} B_2 \cong B_2 \hat{\otimes} B_1$.

7. Examples

We give here some simple (counter)examples to illustrate the results above. Further examples, more important for applications, are given later.

We let N be a positive integer-valued random variable, with $p_n = \mathbb{P}(N = n)$, and let $U \sim \mathrm{U}(0,1)$; we may suppose that U is the identity function defined on $[0,1]$ with Lebesgue measure. We use standard notations from Section 2.

Example 7.1. Let $B = H$ be a separable Hilbert space and let $k = 2$.

Let $\alpha \in L(B^2; \mathbb{R})$ be the inner product in H. Then $\alpha(X, X) = \langle X, X \rangle = \|X\|^2$. Hence, if the projective moment $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense, then, by Theorem 6.15, $\mathbb{E}\, \|X\|^2 < \infty$. Furthermore, weak and Bochner measurability are equivalent by Theorem 3.3, and it follows, using also Lemma 6.3, that the projective second moment $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense if and only if it exists in Bochner sense, and consequently if and only if it exists in Pettis sense.

In this case, the projective second moment thus exists in any sense if and only if $X^{\hat{\otimes}2}$ is measurable and $\mathbb{E}\, \|X\|^2 < \infty$. In particular, the sufficient condition in Theorem 6.15 is also necessary in this case.

Example 7.2. Specialize Example 7.1 to $B = \ell^2$ and let $X = a_N e_N$ for some sequence $(a_n)_1^\infty$. We have seen that:

(i) The projective second moment $\mathbb{E}\, X^{\hat{\otimes}2}$ exists, in any of the three senses, if and only if $\mathbb{E}\,\|X\|^2 = \mathbb{E}\,|a_N|^2 = \sum_{n=1}^\infty p_n a_n^2 < \infty$.

The projective tensor product $\ell^2 \hat{\otimes} \ell^2$ can be seen as the space $\mathcal{N}(\ell^2)$ of trace class operators on ℓ^2 (see Theorem 9.2), and $\mathbb{E}\, X^{\hat{\otimes}2}$ is the diagonal operator $\sum_{n=1}^\infty p_n a_n^2 e_n \otimes e_n$.

The injective tensor product $\ell^2 \check{\otimes} \ell^2$ can, similarly, be seen as the space $\mathcal{K}(\ell^2)$ of compact operators in ℓ^2. A diagonal operator $\sum_n b_n e_n \otimes e_n$ has norm $\sup |b_n|$, and the subspace of diagonal operators in $\mathcal{K}(\ell^2)$ is isomorphic to c_0. Since $X \otimes X = a_N^2 e_N \otimes e_N$ belongs to this subspace, it follows from Example 5.14 that

(ii) $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Dunford sense $\iff \sup p_n a_n^2 < \infty$.
(iii) $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Pettis sense $\iff p_n a_n^2 \to 0$ as $n \to \infty$.
(iv) $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Bochner sense $\iff \sum_{n=1}^\infty p_n a_n^2 < \infty$.

$\mathbb{E}\, X^{\check{\otimes}2}$ is, when it exists, the diagonal operator $\sum p_n a_n^2 e_n \otimes e_n$, just as the projective second moment. As usual, in the Dunford case, $\mathbb{E}\, X^{\check{\otimes}2} \in \mathcal{K}(\ell^2)^{**} = B(\ell^2) = L(\ell^2; \ell^2)$. In fact, the diagonal operator $\sum p_n a_n^2 e_n \otimes e_n$ is compact, i.e. $\mathbb{E}\, X^{\check{\otimes}2} \in \mathcal{K}(\ell^2)$, if and only if $p_n a_n^2 \to 0$ as $n \to \infty$.

Example 7.3. Let $B = L^1[0,1]$. Then $B \hat{\otimes} B = L^1([0,1]^2)$, see e.g. [63, Theorem 46.2 and Exercise 46.5].

Let $X := a_N r_N$, where $(a_n)_1^\infty$ is some sequence of real numbers and $r_n \in L^1[0,1]$ are the Rademacher functions. Then $X \otimes X = a_N^2 r_N \otimes r_N$.

By Khintchin's inequality [8, Theorem II.1] (which applies as well to $r_n \otimes r_n \in L^1([0,1]^2)$, since these functions too can be seen as a sequence of independent symmetric ± 1 random variables), the L^1-norm and L^2-norm are equivalent on the closed linear span R_2 of $\{r_n \otimes r_n\}$ in $L^1([0,1]^2)$. Since $X \otimes X \in R_2$, the expectation $\mathbb{E}(X \otimes X)$ in $B \hat{\otimes} B = L^1([0,1]^2)$, in any of the three senses, can just as well be computed in $L^2([0,1]^2)$. However, the functions $r_n \otimes r_n$ form an orthonormal sequence in L^2, and there is thus an isomorphism between R_2 and ℓ^2, given by $r_n \otimes r_n \mapsto e_n$. Hence it follows from Example 5.15 that the projective moment $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Bochner sense if and only if $\mathbb{E}\,\|X\|^2 = \sum p_n a_n^2 < \infty$, while $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford or Pettis sense if and only if $\sum p_n^2 a_n^4 < \infty$.

For the injective moments, we use Remark 6.22. X lies in the closed linear span R_1 of $\{r_n\}$ in $L^1[0,1]$, which by Khintchin's inequality is isomorphic to ℓ^2; thus $\mathbb{E}\, X \check{\otimes} X$ may be calculated in $R_1 \check{\otimes} R_1 \cong \ell^2 \check{\otimes} \ell^2$. This brings us back to Example 7.2, and thus (ii)–(iv) in Example 7.2 hold in the present case too.

Note that if we consider X as an R_1-valued random variable, then Example 7.2(i) shows that the projective moment $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Pettis (or Dunford) sense if and only if $\sum_n p_n a_n^2 < \infty$. Hence, choosing p_n and a_n such that $p_n a_n^2 = 1/n$, we see that, although R_1 is a closed subspace of $L^1[0,1]$, $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Pettis (or Dunford) sense if we regard X as a $L^1[0,1]$-valued

random variable, but not if we regard X as an R_1-valued random variable, cf. Remark 6.22.

Example 7.4. Let $B = \ell^2[0,1]$ and let $X := a(U)e_U$ where $a : [0,1] \to (0,1]$ is some given function. Then $\langle X, f\rangle = 0$ a.s. for any $f \in \ell^2[0,1]$, since f has countable support, and thus X is weakly measurable and weakly a.s. separably valued. Trivially, $\mathbb{E}\, X = 0$ in Pettis sense. However, a subspace A such that $X \in A$ a.s. has to contain e_t for a.e. $t \in [0,1]$, and is thus not separable. Consequently, X is not a.s. separably valued, and therefore not Bochner integrable.

Furthermore, every integral bilinear form on a Hilbert space is nuclear (Theorem 9.3 below), and thus by Theorem 6.10(ii)(c), $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Dunford sense; by (6.3), $\mathbb{E}\, X^{\check{\otimes}2} = 0$. (By Theorem 9.9, $\mathbb{E}\, X^{\check{\otimes}2} = 0$ also in Pettis sense.)

Let $\alpha \in L(B^2;\mathbb{R})$ be the inner product in $B = \ell^2[0,1]$. Then $\alpha(X,X) = \langle X, X\rangle = \|X\|^2 = a(U)^2$. If we assume that a is a non-measurable function (for Lebesgue measure on $[0,1]$), then $a(U)$ is non-measurable. Hence, $\alpha(X,X)$ is non-measurable, and by Theorem 6.15, the projective moment $\mathbb{E}\, X^{\hat{\otimes}2}$ does not exist in Dunford sense (and thus not in the other, stronger, senses).

In particular, we see that although X is weakly measurable in B, $X \otimes X$ is not weakly measurable in $B\hat{\otimes}B$, since $\langle \alpha, X \otimes X\rangle = \alpha(X,X)$ is not measurable. (We use here Theorem 4.6.)

Furthermore, $X \otimes X$ is not weakly a.s. separably valued in $B\hat{\otimes}B$, by the following argument. Suppose that A is a separable subspace of $B\hat{\otimes}B$. Then, as in the proof of Lemma 6.5, there exists a countable family of elementary tensors $F = \{e_{i1} \otimes e_{i2}\}$ such that A is included in the closed linear span of F. Let $R := \bigcup_{i,j} \operatorname{supp}(e_{ij})$; then $R \subseteq [0,1]$ is countable and if $y \in A$, then $\operatorname{supp}(y) \subseteq R \times R$. Define $\beta \in (\ell^2\hat{\otimes}\ell^2)^*$ by $\beta(f,g) := \sum_{t\notin R} f(t)g(t)$. Then $\beta(e_{i1}, e_{i2}) = 0$ for all i, and thus $\beta \perp F$ and $\beta \perp A$, but $\beta(X,X) = \langle X, X\rangle = a(U)^2 \neq 0$ a.s.

Example 7.5. Let, as in Example 5.14, $B = c_0$ and $X = a_N e_N$. Thus $X^{\otimes k} = a_N^k e_N^{\otimes k}$.

Let $\Delta_k^\circ \subset c_0^{\otimes k}$ be the subspace of finite linear combinations of tensors $e_N^{\otimes k}$; we claim that the closure of Δ_k° is the same in both the projective and injective tensor products $c_0^{\hat{\otimes}k}$ and $c_0^{\check{\otimes}k}$ and that it equals $\Delta_k := \{\sum_{n=1}^\infty b_n e_n^{\otimes k} : (b_n)_1^\infty \in c_0\}$. (For the injective tensor product, this follows immediately from Theorem 12.1, but the projective case is less obvious.) Since Δ_k obviously is isomorphic to c_0, it then follows from Example 5.14 that:

(i) $\mathbb{E}\, X^{\hat{\otimes}k}$ exists in Dunford sense $\iff$ $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Dunford sense $\iff$ $\sup p_n a_n^k < \infty$.

(ii) $\mathbb{E}\, X^{\hat{\otimes}k}$ exists in Pettis sense $\iff$ $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Pettis sense $\iff$ $p_n a_n^k \to 0$ as $n \to \infty$.

(iii) $\mathbb{E}\, X^{\hat{\otimes} k}$ exists in Bochner sense $\iff$ $\mathbb{E}\, X^{\check{\otimes} k}$ exists in Bochner sense $\iff$ $\sum_{n=1}^{\infty} p_n a_n^k < \infty$.

Hence there is no difference between projective and injective moments in this example. (Cf. Example 7.2, where we see that Δ_2 is the closure of Δ_2° also in $\ell^2 \check{\otimes} \ell^2$, but not in $\ell^2 \hat{\otimes} \ell^2$.)

To verify the claim, consider for simplicity first $k = 2$, and let $u = \sum_{n=1}^{M} b_n e_n \otimes e_n \in \Delta_2^\circ$. Recalling (4.3) and (4.6), and taking $x_1^* = x_2^* = e_n$ in the latter, we have immediately

$$\|u\|_\pi \geqslant \|u\|_\varepsilon \geqslant \max_n |b_n|. \tag{7.1}$$

For $\zeta = (\zeta_1, \dots, \zeta_M)$ with $\zeta_i = \pm 1$, take $v_\zeta := \sum_{n=1}^{M} \zeta_n b_n e_n \in c_0$ and $w_\zeta := \sum_{n=1}^{M} \zeta_n e_n \in c_0$. Then, see (4.3) and (4.7),

$$\|v_\zeta \otimes w_\zeta\|_\pi = \|v_\zeta\|_{c_0} \|w_\zeta\|_{c_0} = \max_n |b_n|. \tag{7.2}$$

Taking the average of $v_\zeta \otimes w_\zeta$ over the 2^M possible choices of ζ, we obtain u, and thus

$$\|u\|_\pi \leqslant \max_n |b_n|. \tag{7.3}$$

Consequently, we have equalities in (7.1), for any $u \in \Delta^\circ$, which shows that the closure in either $c_0 \hat{\otimes} c_0$ or $c_0 \check{\otimes} c_0$ is isomorphic to c_0 and equals Δ_2.

The argument extends to arbitrary $k \geqslant 2$ by letting the possible values of ζ_n be the k:th roots of unity and considering $v_\zeta \otimes w_\zeta^{\otimes k-1}$. (These vectors are complex, but we can separate them into real and imaginary parts, possibly introducing a constant factor 2^k in the norm estimate (7.3).)

Example 7.6. We say that a random variable X in a Banach space B is *weakly Gaussian* if $x^*(X)$ is Gaussian with mean 0 (for convenience) for any $x^* \in B^*$. To exclude cases such as Example 7.4 where $x^*(X) = 0$ a.s. (and thus is Gaussian) but X does not look very Gaussian, we say that X is *Gaussian* if it is weakly Gaussian and a.s. separably valued. (By Theorem 3.3, this is equivalent to weakly Gaussian and Bochner measurable.)

If X is (weakly) Gaussian, then $x^*(X)$ is Gaussian and thus has finite moments of all orders, for any $x^* \in B^*$. Thus Theorem 6.10(ii)(b) shows that every injective moment $\mathbb{E}\, X^{\check{\otimes} k}$ exists in Dunford sense.

Moreover, if X is Gaussian, by [42, Lemma 3.1 and Corollary 3.2] (applied to a suitable separable subspace $B_1 \subseteq B$ with $X \in B_1$ a.s.), $\|X\|^k < \infty$ for every k. Hence, Theorem 6.7 shows that the projective and injective moments $\mathbb{E}\, X^{\hat{\otimes} k}$ and $\mathbb{E}\, X^{\check{\otimes} k}$ exist in Bochner sense for every $k \geqslant 1$.

The odd moments vanish by symmetry.

The even injective moments can be expressed in terms of the second moment $\Sigma := \mathbb{E}\, X^{\check{\otimes} 2}$ as follows. Consider first $k = 4$. Then, by Theorem 6.13

and Wick's theorem [37, Theorem 1.28], $\mathbb{E}\, X^{\check{\otimes}4} \in B^{\check{\otimes}4}$ is determined by

$$\begin{aligned}\langle X^{\check{\otimes}4}, x_1^* \otimes x_2^* \otimes x_3^* \otimes x_4^* \rangle &= \mathbb{E}\big(x_1^*(X)x_2^*(X)x_3^*(X)x_4^*(X)\big) \\ &= \mathbb{E}\big(x_1^*(X)x_2^*(X)\big)\,\mathbb{E}\big(x_3^*(X)x_4^*(X)\big) + \mathbb{E}\big(x_1^*(X)x_3^*(X)\big)\,\mathbb{E}\big(x_2^*(X)x_4^*(X)\big) \\ &\qquad + \mathbb{E}\big(x_1^*(X)x_4^*(X)\big)\,\mathbb{E}\big(x_2^*(X)x_3^*(X)\big) \\ &= \langle \Sigma, x_1^* \otimes x_2^* \rangle \langle \Sigma, x_3^* \otimes x_4^* \rangle + \langle \Sigma, x_1^* \otimes x_3^* \rangle \langle \Sigma, x_2^* \otimes x_4^* \rangle \\ &\qquad + \langle \Sigma, x_1^* \otimes x_4^* \rangle \langle \Sigma, x_2^* \otimes x_3^* \rangle.\end{aligned}$$

This can be written as

$$\mathbb{E}\, X^{\check{\otimes}4} = \Sigma \otimes \Sigma + \pi_{(23)}\big(\Sigma \otimes \Sigma\big) + \pi_{(24)}\big(\Sigma \otimes \Sigma\big), \tag{7.4}$$

where π_σ denotes permuting the coordinates by the premutation σ. There are 24 permutations of $\{1,2,3,4\}$, but 8 of these leave $\Sigma \otimes \Sigma$ invariant, and we may write (7.4) more symmetrically as $\mathbb{E}\, X^{\check{\otimes}4} = 3\,\mathrm{Symm}\big(\Sigma \otimes \Sigma\big)$, where Symm means the symmetrization of the tensor by taking the average over all permutations of the coordinates.

More generally, for any even $k = 2\ell$, by the same argument,

$$\mathbb{E}\, X^{\check{\otimes}2\ell} = \frac{(2\ell)!}{2^\ell \ell!}\,\mathrm{Symm}\left((\mathbb{E}\, X^{\check{\otimes}2})^{\check{\otimes}\ell}\right) \in B^{\check{\otimes}2\ell}, \tag{7.5}$$

which generalizes the standard formula $\mathbb{E}\,\xi^{2\ell} = \frac{(2\ell)!}{2^\ell \ell!}(\mathrm{Var}\,\xi)^\ell$ for a real-valued centred Gaussian variable ξ.

We conjecture that the corresponding formula for projective moments holds too. If B has the approximation property, this follows by (7.5) and Theorem 8.3 below, but we leave the general case as an open problem.

8. The approximation property

Let B_1, B_2 be Banach spaces. Recall that a finite rank operator $F : B_1 \to B_2$ is a continuous linear operator whose range has finite dimension; equivalently, it is a linear operator that can be written as a finite sum $F(x) = \sum_{i=1}^n x_i^*(x) y_i$ for some $x_i^* \in B_1^*$ and $y_i \in B_2$.

We say that a linear operator $T : B_1 \to B_2$ is *uniformly approximable by finite rank operators* if for every $\varepsilon > 0$ there exists a finite rank operator $F : B_1 \to B_2$ such that $\|T - F\| < \varepsilon$. Similarly, we say that a linear operator $T : B_1 \to B_2$ is *approximable on compacts by finite rank operators* if for every compact set $K \subset B_1$ and every $\varepsilon > 0$ there exists a finite rank operator $F : B_1 \to B_2$ such that $\sup\{\|Tx - Fx\| : x \in K\} < \varepsilon$.

If B is a Banach space, then the following properties are equivalent; see e.g. [43, Section 1.e] and [57, Chapter 4] for proofs. The Banach space B is said to have the *approximation property* when these properties hold.

(i) The identity operator $I : B \to B$ is approximable on compacts by finite rank operators.
(ii) For every Banach space B_1, every bounded operator $T : B \to B_1$ is approximable on compacts by finite rank operators.

(iii) For every Banach space B_1, every bounded operator $T : B_1 \to B$ is approximable on compacts by finite rank operators.
(iv) For every Banach space B_1, every compact operator $T : B_1 \to B$ is uniformly approximable by finite rank operators.
(v) For every pair of sequences $x_n \in B$ and $x_n^* \in B^*$, $n \geqslant 1$, such that $\sum_{n=1}^\infty \|x_n\|\|x_n^*\| < \infty$ and $\sum_{n=1}^\infty x_n^*(x)x_n = 0$ for all $x \in B$, we have $\sum_{n=1}^\infty x_n^*(x_n) = 0$.
(vi) For every pair of sequences $x_n \in B$ and $x_n^* \in B^*$, $n \geqslant 1$, such that $\sum_{n=1}^\infty \|x_n\|\|x_n^*\| < \infty$ and $\sum_{n=1}^\infty x_n^*(x)x_n = 0$ for all $x \in B$, we have $\sum_{n=1}^\infty x_n \otimes x_n^* = 0$ in $B\hat{\otimes}B^*$.
(vii) For every Banach space B_1 and every pair of sequences $x_n \in B$ and $y_n \in B_1$, $n \geqslant 1$, such that $\sum_{n=1}^\infty \|x_n\|\|y_n\| < \infty$ and $\sum_{n=1}^\infty x^*(x_n)y_n = 0$ for all $x^* \in B^*$, we have $\sum_{n=1}^\infty x_n \otimes y_n = 0$ in $B\hat{\otimes}B_1$.

Remark 8.1. The property dual to (iv), viz. that every compact operator $T : B \to B_1$ is uniformly approximable by finite rank operators, for every Banach space B_1, is *not* equivalent to the other properties; in fact, it is equivalent to the approximation property of B^*.

Moreover, it is known that if B^* has the approximation property, then B has the approximation property, but the converse does not hold. (See Example 8.5 for a concrete example. There are also counterexamples that are separable with a separable dual [43].)

We can reformulate conditions (vi) and (vii) as follows, recalling the canonical injection ι in (4.8). (This is implicit in the references above, and explicit in e.g. [61].)

Theorem 8.2. *Let B be a Banach space. If B has the approximation property, then the canonical mapping $\iota : B\hat{\otimes}B_1 \to B\check{\otimes}B_1$ is injective for every Banach space B_1.*

Conversely, if the canonical mapping $\iota : B\hat{\otimes}B^* \to B\check{\otimes}B^*$ is injective, then B has the approximation property.

Proof. Let $u \in B\hat{\otimes}B_1$; then $u = \sum_{n=1}^\infty x_n \otimes y_n$ for some $x_n \in B$ and $y_n \in B_1$ with $\sum_{n=1}^\infty \|x_n\|\|y_n\| < \infty$. We can regard $B\check{\otimes}B_1$ as space of bilinear forms on $B^* \times B_1^*$, and then, for any $x^* \in B^*$ and $y^* \in B_1^*$

$$\iota(u)(x^*, y^*) = \sum_{n=1}^\infty x^*(x_n)y^*(y_n) = y^*\Big(\sum_{n=1}^\infty x^*(x_n)y_n\Big).$$

Hence, $\iota(u) = 0$ if and only if $\sum_{n=1}^\infty x^*(x_n)y_n = 0$ for every $x^* \in B^*$.

Consequently, (vii) says precisely that, for any B_1, if $u \in B\hat{\otimes}B_1$ and $\iota(u) = 0$, then $u = 0$, i.e. that $\iota : B\hat{\otimes}B_1 \to B\check{\otimes}B_1$ is injective.

Furthermore, (vi) is the special case $B_1 = B^*$, and thus says that $\iota : B\hat{\otimes}B^* \to B\check{\otimes}B^*$ is injective. □

This can be extended to tensor products of several spaces. We state only the case of tensor powers of a single space, which is the case we need.

Theorem 8.3. *If a Banach space B has the approximation property, then the canonical mapping $\iota : B^{\hat{\otimes}k} \to B^{\check{\otimes}k}$ is injective.*

Proof. We use induction in k. The case $k = 1$ is trivial, and $k = 2$ is a consequence of Theorem 8.2. For $k \geqslant 3$ we write $\iota : B^{\hat{\otimes}k} \to B^{\check{\otimes}k}$ as the composition

$$B^{\hat{\otimes}k} = B\hat{\otimes}B^{\hat{\otimes}(k-1)} \to B\check{\otimes}B^{\hat{\otimes}(k-1)} \to B\check{\otimes}B^{\check{\otimes}(k-1)} = B^{\check{\otimes}k}$$

where the first map is injective by Theorem 8.2 and the second is injective by induction and Lemma 4.3. □

Remark 8.4. The approximation property for B is *not* equivalent to $\iota : B\hat{\otimes}B \to B\check{\otimes}B$ being injective. In fact, a counterexample by Pisier [52, Theorem 10.6] shows that there exists an infinite-dimensional Banach space B such that $B\hat{\otimes}B = B\check{\otimes}B$ (with equivalent norms); moreover, this space B lacks the approximation property.

The study of the approximation property was initiated by Grothendieck [32, 33] who found most of the results above but did not know whether any Banach spaces without the approximation property exist. The first counterexample was found by Enflo [26], who constructed a separable, reflexive Banach space B without the approximation property. A modification of the counterexample given by Davie [15, 16], see also Lindenstrauss and Tzafriri [43, Theorem 2.d.6], shows that B may be taken as a subspace of c_0 or of ℓ^p, for any $2 < p < \infty$. Another counterexample was found by Szankowski [61], who showed that the space $B(H)$ of bounded operators in an infinite-dimensional Hilbert space does not have the approximation property.

On the other hand, it is easy to see that any Banach space with a (Schauder) basis has the approximation property; this includes all classical examples of separable Banach spaces. (In fact, Enflo's counterexample [26] was also the first known separable Banach space without a basis.) There are also many non-separable Banach spaces with the approximation property. The list of Banach spaces with the approximation property includes, for example, ℓ^p $(1 \leqslant p \leqslant \infty)$, c_0, $L^p(\mu)$ $(1 \leqslant p \leqslant \infty$, μ any measure), $C(K)$ (K a compact set).

Example 8.5. The tensor products $\ell^2\hat{\otimes}\ell^2$ and $\ell^2\check{\otimes}\ell^2$ have bases and thus have the approximation property, see e.g. [57, Proposition 4.25 and Exercise 4.5]. As said in Example 4.2, see also Theorem 9.2, these spaces can be identified with the spaces of trace class operators (= nuclear operators) and compact operators in ℓ^2, respectively; moreover, $\ell^2\hat{\otimes}\ell^2 \cong (\ell^2\check{\otimes}\ell^2)^*$. However, $(\ell^2\hat{\otimes}\ell^2)^*$ can be identified with the space $B(\ell^2)$ of bounded operators in ℓ^2, which as just said does not have the approximation property [61].

We can now prove Theorems 1.1 and 1.2.

Proof of Theorem 1.1. By Theorem 6.7, the projective and injective moments $\mathbb{E}\,X^{\hat{\otimes}k}$, $\mathbb{E}\,Y^{\hat{\otimes}k}$, $\mathbb{E}\,X^{\check{\otimes}k}$, $\mathbb{E}\,Y^{\check{\otimes}k}$ exist in Bochner sense. If (1.3) holds,

then $\mathbb{E}\,X^{\check{\otimes}k} = \mathbb{E}\,Y^{\check{\otimes}k}$ by Corollary 6.14. By (6.1), this can be written $\iota(\mathbb{E}\,X^{\hat{\otimes}k}) = \iota(\mathbb{E}\,Y^{\hat{\otimes}k})$, and since Theorem 8.3 shows that ι is injective, we have $\mathbb{E}\,X^{\hat{\otimes}k} = \mathbb{E}\,Y^{\hat{\otimes}k}$. By Corollary 6.17, this yields (1.2).

The converse is trivial. □

Remark 8.6. More generally, Theorem 8.3 implies that for any Banach space B with the approximation property, if $\mathbb{E}\,X^{\hat{\otimes}k}$ and $\mathbb{E}\,Y^{\hat{\otimes}k}$ (and thus also $\mathbb{E}\,X^{\check{\otimes}k}$ and $\mathbb{E}\,Y^{\check{\otimes}k}$) exist in Pettis sense and $\mathbb{E}\,X^{\check{\otimes}k} = \mathbb{E}\,Y^{\check{\otimes}k}$, then $\mathbb{E}\,X^{\hat{\otimes}k} = \mathbb{E}\,Y^{\hat{\otimes}k}$. We do not know whether this remains valid if we only assume that the moments exist in Dunford sense. (Theorem 11.23 is a positive result in a special case.)

Proof of Theorem 1.2. Let B_0 be a Banach space without the approximation property and let $B := B_0 \oplus B_0^*$, with the norm $\|(x,x^*)\|_B := \max\bigl(\|x\|,\|x^*\|\bigr)$. By choosing B_0 to be separable and reflexive, we obtain B separable and reflexive too. We shall show that there exist bounded Bochner measurable random variables X and Y in B such that, for $k=2$, (1.3) holds but not (1.2).

Since B_0 does not have the approximation property, by (v) above, there exist sequences $x_n \in B_0$ and $x_n^* \in B_0^*$ with such that $\sum_{n=1}^\infty \|x_n\|\|x_n^*\| < \infty$ and $\sum_{n=1}^\infty x_n^*(x)x_n = 0$ for all $x \in B$, but $\sum_{n=1}^\infty x_n^*(x_n) = 0$. Let $a_n := \|x_n\|\,\|x_n^*\|$, so $0 < \sum_{n=1}^\infty a_n < \infty$. We may eliminate all (x_n,x_n^*) with $a_n = 0$, and we may thus assume that $a_n > 0$ for each n. Define $y_n := x_n/\|x_n\| \in B_0$, $y_n^* := x_n^*/\|x_n^*\| \in B_0^*$ and $p_n := a_n/\sum_{m=1}^\infty a_m$. Thus y_n and y_n^* are unit vectors and $\sum_{n=1}^\infty p_n = 1$. Furthermore, the properties of x_n and x_n^* translate to

$$\sum_{n=1}^\infty p_n y_n^*(z) y_n = 0 \qquad \text{for every } z \in B_0, \tag{8.1}$$

$$\sum_{n=1}^\infty p_n y_n^*(y_n) \neq 0. \tag{8.2}$$

Let N be a random positive integer with the distribution $\mathbb{P}(N=n) = p_n$, and let X and Y be the B-valued random variables

$$X := (y_N, y_N^*),$$
$$Y := (y_N, -y_N^*).$$

Note that $\|X\| = \|Y\| = 1$ a.s. Then, if α is any bounded bilinear form $B \times B \to \mathbb{R}$, then

$$\alpha(X,X) = \alpha\bigl((y_N,0),(y_N,0)\bigr) + \alpha\bigl((y_N,0),(0,y_N^*)\bigr) + \alpha\bigl((0,y_N^*),(y_N,0)\bigr) + \alpha\bigl((0,y_N^*),(0,y_N^*)\bigr)$$

and similarly for Y. Hence,

$$\mathbb{E}\,\alpha(X,X) - \mathbb{E}\,\alpha(Y,Y) = 2\,\mathbb{E}\,\alpha\bigl((y_N,0),(0,y_N^*)\bigr) + 2\,\mathbb{E}\,\alpha\bigl((0,y_N^*),(y_N,0)\bigr). \tag{8.3}$$

In particular, letting α be the bounded bilinear form $\alpha\bigl((x,x^*),(y,y^*)\bigr) := x^*(y)$, we have by (8.2)

$$\mathbb{E}\,\alpha(X,X) - \mathbb{E}\,\alpha(Y,Y) = 2\,\mathbb{E}\,y_N^*(y_N) = 2\sum_{n=1}^{\infty} p_n y_n^*(y_n) \neq 0.$$

Hence, (1.2) does not hold.

On the other hand, if $x_1^*, x_2^* \in B^*$, then $x_i^*(x,x^*) = z_i^*(x) + z_i^{**}(x^*)$ for some $z_i^* \in B_0^*$ and $z_i^{**} \in B_0^{**}$, $i = 1,2$. Hence,

$$\mathbb{E}\bigl(x_1^*(y_N,0)x_2^*(0,y_N^*)\bigr) = \mathbb{E}\bigl(z_1^*(y_N)z_2^{**}(y_N^*)\bigr) = z_2^{**}\Bigl(\mathbb{E}\bigl(z_1^*(y_N)y_N^*\bigr)\Bigr). \quad (8.4)$$

However, the continuous linear functional

$$y^* := \mathbb{E}\bigl(z_1^*(y_N)y_N^*\bigr) = \sum_{n=1}^{\infty} p_n z_1^*(y_n)y_n^* \in B_0^*$$

satisfies, for every $x \in B_0$,

$$y^*(x) = \sum_{n=1}^{\infty} p_n z_1^*(y_n)y_n^*(x) = z_1^*\Bigl(\sum_{n=1}^{\infty} p_n y_n^*(x)y_n\Bigr) = 0$$

by (8.1). Thus $y^* = 0$ and (8.4) yields

$$\mathbb{E}\bigl(x_1^*(y_N,0)x_2^*(0,y_N^*)\bigr) = z_2^{**}(y^*) = 0.$$

Interchanging x_1^* and x_2^*, we see that also $\mathbb{E}\bigl(x_1^*(0,y_N^*)x_2^*(y_N,0)\bigr) = 0$. Thus (8.3) with $\alpha(x_1,x_2) := x_1^*(x_1)x_2^*(x_2)$ yields

$$\mathbb{E}\bigl(x_1^*(X)x_2^*(X)\bigr) - \mathbb{E}\bigl(x_1^*(Y)x_2^*(Y)\bigr) = 0,$$

which shows that (1.3) holds. □

The counterexample in this proof has also other unpleasant consequences.

Example 8.7. Let the separable Banach space B and the random variables X and Y be as in the proof of Theorem 1.2. By Corollaries 6.17 and 6.14, $\mathbb{E}(X\hat{\otimes}X) \neq \mathbb{E}(Y\hat{\otimes}Y)$ in $B^{\hat{\otimes}2}$ but $\mathbb{E}(X\check{\otimes}X) = \mathbb{E}(Y\check{\otimes}Y)$ in $B^{\check{\otimes}2}$, where all moments exist (in Bochner sense) by Theorem 6.7.

We can embed B as a closed subspace of a Banach space B_1 with the approximation property, for example by the Banach–Mazur theorem which says that every separable Banach space can be embedded as a closed subspace of $C[0,1]$. Let $i : B \to B_1$ denote the embedding.

We may regard X and Y also as B_1-valued random variables. Then the injective second moment $\mathbb{E}(X\check{\otimes}X)$ in $B\check{\otimes}B$ is mapped by $i\check{\otimes}i$ to the injective second moment $\mathbb{E}(X\check{\otimes}X)$ in $B_1\check{\otimes}B_1$, and it follows that $\mathbb{E}(X\check{\otimes}X) = \mathbb{E}(Y\check{\otimes}Y)$ also in $B_1\check{\otimes}B_1$, i.e., X and Y have the same injective second moments in B_1.

It follows from Theorem 8.3 (as in the proof of Theorem 1.1) that $\mathbb{E}(X\hat{\otimes}X) = \mathbb{E}(Y\hat{\otimes}Y)$ in $B_1\hat{\otimes}B_1$, i.e., X and Y have the same projective second moments in B_1 although they have different projective second moments in $B\otimes B$. This

shows that if a Banach space valued random variable takes values in a subspace of the Banach space, we may have to be careful with in which space we calculate the projective moments. (Note that Lemma 4.3 and Remark 4.4 show that there is no such problem for injective moments.)

Since $i\hat{\otimes}i$ maps the second projective moments in $B\hat{\otimes}B$ to the second projective moments in $B_1\hat{\otimes}B_1$, we have shown that $(i\hat{\otimes}i)\,\mathbb{E}(X\hat{\otimes}X) = (i\hat{\otimes}i)\,\mathbb{E}(Y\hat{\otimes}Y)$, although $\mathbb{E}(X\hat{\otimes}X) \neq \mathbb{E}(Y\hat{\otimes}Y)$. Consequently, $i\hat{\otimes}i : B\hat{\otimes}B \to B_1\hat{\otimes}B_1$ is not injective, see Remark 4.5.

Remark 8.8. A simplified version of the counterexample in Example 8.7, without mentioning moments, is the following: let B be a subspace of a Banach space B_1 such that B_1 has the approximation property but B has not. Let $i : B \to B_1$ be the inclusion.

Then $\iota : B\hat{\otimes}B^* \to B\check{\otimes}B^*$ is not injective by Theorem 8.2. Thus the composition $(i\check{\otimes}I)\iota : B\hat{\otimes}B^* \to B\check{\otimes}B^* \to B_1\check{\otimes}B^*$ is not injective, but this equals the composition $\iota(i\hat{\otimes}I) : B\hat{\otimes}B^* \to B_1\hat{\otimes}B^* \to B_1\check{\otimes}B^*$. On the other hand, since B_1 has the approximation property; $\iota : B_1\hat{\otimes}B^* \to B_1\check{\otimes}B^*$ is injective by Theorem 8.2; hence $i\hat{\otimes}I : B\hat{\otimes}B^* \to B_1\hat{\otimes}B^*$ is not injective.

9. Hilbert spaces

Consider the case $B = H$, a Hilbert space. We shall give some special results for second moments. (We do not know whether the results extend to moments of order $k \geqslant 3$ or not, and leave this as open problems.) We begin with some well-known results.

Theorem 9.1. *A Hilbert space has the approximation property.*

Proof. Property (i) in Section 8 is easily verified using suitable orthogonal projections. □

Next, we note that $H^* = H$; hence the correspondence (4.9) yields an isometric embedding of $H\check{\otimes}H$ into $B(H) = L(H;H)$, the space of bounded linear operators on H. We identify a tensor in $H\check{\otimes}H$ and the corresponding operator without further comment; hence we regard $H\check{\otimes}H$ as a subspace of $B(H)$. Moreover, an elementary tensor $x \otimes y$ corresponds to an operator of rank 1, and every operator of rank 1 is given by an elementary tensor; hence the tensors in $H \otimes H$, which are finite sums of elementary tensors, are exactly the operators on B of finite rank. The injective tensor product $H\check{\otimes}H$ is thus the closure in $B(H)$ of the set of finite rank operators, which shows (see (iv) in Section 8) that $H\check{\otimes}H = \mathcal{K}(H)$, the space of compact operators $H \to H$.

The natural map $H\hat{\otimes}H \to \mathcal{N}(H,H)$ onto the nuclear forms is a bijection; moreover, the nuclear and integral bilinear forms on H coincide. Equivalently, the integral and nuclear operators $H \to H$ coincide, and furthermore, the set of them equals the set of *trace class operators*, which we denote by $\mathcal{N}(H)$. We can thus identify all these spaces of bilinear forms or operators

with $H\hat{\otimes}H$. (See e.g. [63, Chapter 48 and Proposition 49.6]. See further e.g. [13, Exercise IX.2.20] or [41, Chapter 31].)

With these identifications, the map $\iota : H\hat{\otimes}H \to H\check{\otimes}H$ is just the inclusion map $\mathcal{N}(H) \to \mathcal{K}(H) \subseteq B(H)$. (By Theorem 8.2 and $H^* = H$, the approximation property is equivalent to the fact that $\iota : H\hat{\otimes}H \to H\check{\otimes}H$ is injective, which we thus also see explicitly.)

We summarize these descriptions of the tensor products.

Theorem 9.2. *Under the identification* (4.9) *of tensors and operators $H^* = H \to H$, we have $H\hat{\otimes}H = \mathcal{N}(H)$, the space of trace class operators on H (which equals the space of nuclear operators), and $H\check{\otimes}H = \mathcal{K}(H)$, the space of compact operators.* □

Theorem 9.3. *Every integral bilinear form on H is nuclear. The space of these forms can be identified with $H\hat{\otimes}H$.* □

The general formulas $(B_1\hat{\otimes}B_2)^* = L(B_1, B_2; \mathbb{R})$ and $(B_1\check{\otimes}B_2)^* = \mathcal{I}(B_1, B_2)$, see Theorems 4.6–4.7, can be translated to operators as $\mathcal{N}(H)^* = B(H)$ and $\mathcal{K}(H)^* = \mathcal{N}(H)$, where the dualities are given by the trace form $\langle T, S\rangle = \operatorname{Tr}(TS^*)$.

In particular, $(H\check{\otimes}H)^{**} = \mathcal{K}(H)^{**} = B(H)$. Thus, if the second injective moment exists in Bochner or Pettis sense, it is by the correspondence (4.9) given by a compact operator in $\mathcal{K}(H)$, and if it exists in Dunford sense it is, again by (4.9), given by an operator in $B(H)$. Similarly, if the second projective moment exists in Bochner or Pettis sense, it is given by a trace class operator in $\mathcal{N}(H)$, and the second injective moment equals the same operator. In all these cases, the following theorem shows that the second moment is a positive operator. (In particular, it is self-adjoint.)

Theorem 9.4. *If the second moment $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Bochner or Pettis sense, or $\mathbb{E}\, X^{\check{\otimes}2}$ exists in any sense, then, regarding the moment as an operator in $B(H)$, it is a positive operator.*

Proof. In all cases, it follows that the injective moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Dunford sense. Thus, for $x \in H$, by (6.3),

$$\langle \mathbb{E}\, X^{\check{\otimes}2} x, x\rangle = \langle \mathbb{E}\, X^{\check{\otimes}2}, x\otimes x\rangle = \mathbb{E}\langle X, x\rangle^2 \geqslant 0. \qquad \square$$

Remark 9.5. The only case remaining is the second projective moment $\mathbb{E}\, X^{\hat{\otimes}2}$ in Dunford sense, which belongs to $\mathcal{N}(H)^{**} = B(H)^*$. We shall see in Example 9.15 (under a set theory hypothesis) that this moment is not always given by an operator on H.

However, if H is separable, or more generally if X is a.s. separably valued, and $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense, then it exists in Pettis sense too and thus $\mathbb{E}\, X^{\hat{\otimes}2} \in \mathcal{N}(H)$. This follows from Theorem 5.13, since $\mathcal{N}(H)$ does not contain a subspace isomorphic to c_0 (e.g. by [17, Theorem V.10], since $H\hat{\otimes}H$ is a separable dual space when H is separable; we omit the details); moreover, we shall prove a more general result by a different method in

Theorem 9.14, which shows that in essentially all cases (again depending on a set theory hypothesis), the Dunford and Pettis senses coincide and thus $\mathbb{E}\, X^{\hat{\otimes}2} \in \mathcal{N}(H)$.

We next characterize when the injective second moment exists, in the three different senses; we begin with two lemmas.

Lemma 9.6. *Every weakly measurable H-valued random variable is weakly a.s. separably valued.*

Example 7.4 shows that X is not necessarily a.s. separably valued.

Proof. Let $\{e_s\}_{s\in S}$ be an ON basis in the Hilbert space H, and let $\xi_s := |\langle e_s, X\rangle|^2$. Since X is weakly measurable, each ξ_s is a non-negative random variable.

Let $\mathcal{C}$ be the collection of countable subsets $A \subseteq S$. For $A \in \mathcal{C}$, let

$$\xi_A := \sum_{s\in A} \xi_s = \sum_{s\in A} |\langle e_s, X\rangle|^2 \leqslant \|X\|^2 < \infty$$

Thus every ξ_A is a finite non-negative random variable, and $A_1 \subseteq A_2 \implies \xi_{A_1} \leqslant \xi_{A_2}$.

Let $m_A := \mathbb{E} \arctan \xi_A$, and

$$m_* := \sup\{m_A : A \in \mathcal{C}\}. \tag{9.1}$$

Note that $m_* \leqslant \pi/2 < \infty$.

There exist $A_n \in \mathcal{C}$ with $m_{A_n} > m_* - 1/n$, so taking $A_* := \bigcup_{n=1}^\infty A_n$ we have $m_{A_*} \geqslant m_*$ and thus the sup in (9.1) is attained. Moreover, for any $s \notin A_*$, $A_* \cup \{s\} \in \mathcal{C}$ and thus $m_{A_*\cup\{s\}} \leqslant m_* = m_{A_*}$, i.e.,

$$\mathbb{E} \arctan \xi_{A_*\cup\{s\}} \leqslant \mathbb{E} \arctan \xi_{A_*}. \tag{9.2}$$

Since $\xi_{A_*\cup\{s\}} = \xi_{A_*} + \xi_s \geqslant \xi_{A_*}$, (9.2) implies $\arctan(\xi_{A_*} + \xi_s) = \arctan \xi_{A_*}$ a.s., and since $\xi_{A_*} < \infty$, thus $\xi_s = 0$ a.s. Consequently, if $s \notin A_*$, then

$$\langle e_s, X\rangle = 0 \quad \text{a.s.} \tag{9.3}$$

Let M be the closed linear span of $\{e_s : s \in A_*\}$. M is a separable subspace of H. If $y \in M^\perp$, then $y = \sum_{s\notin A_*} a_s e_s$ with only a countable number of $a_s \neq 0$; hence (9.3) implies $\langle y, X\rangle = 0$ a.s. □

Lemma 9.7. *If X is a weakly measurable H-valued random variable, then $X^{\otimes 2}$ is weakly a.s. separably valued in $H^{\breve{\otimes}2}$.*

Proof. By Lemma 9.6, there exists a separable subspace $M \subseteq H$ such that $x^*(X) = 0$ a.s. for every $x^* \perp M$. Let $\widetilde{M}$ be the closed subspace of $H\breve{\otimes}H$ spanned by $\{x \otimes y : x, y \in M\}$.

Let $\alpha \in (H\breve{\otimes}H)^*$; then α is a bilinear form on H which by Theorem 9.3 is nuclear; thus there exist x_n^*, y_n^* with $\sum_{n=1}^\infty \|x_n^*\|\|y_n^*\| < \infty$ and

$$\alpha(x, y) = \sum_{n=1}^\infty x_n^*(x) y_n^*(y). \tag{9.4}$$

Assume that $\alpha \perp \widetilde{M}$.

Let $P : H \to M$ be the orthogonal projection onto M and $Q := I - P$. Decompose α as $\alpha_{PP} + \alpha_{PQ} + \alpha_{QP} + \alpha_{QQ}$, where $\alpha_{PP}(x, y) := \alpha(Px, Py)$, $\alpha_{PQ}(x, y) := \alpha(Px, Qy)$, $\alpha_{QP}(x, y) := \alpha(Qx, Py)$, $\alpha_{QQ}(x, y) := \alpha(Qx, Qy)$.

For any $x, y \in H$, $\alpha(Px, Py) = \langle \alpha, Px \otimes Py \rangle = 0$, since $x \otimes y \in \widetilde{M}$. Hence $\alpha_{PP} = 0$. Consider one of the other terms, for example α_{PQ}. By (9.4),

$$\alpha_{PQ}(X, X) = \sum_{n=1}^{\infty} \langle x_n^*, PX \rangle \langle y_n^*, QX \rangle = \sum_{n=1}^{\infty} \langle Px_n^*, X \rangle \langle Qy_n^*, X \rangle. \tag{9.5}$$

However, $Qy_n^* \perp M$ and thus, by the choice of M, $\langle Qy_n^*, X \rangle = 0$ a.s., for every n. Hence (9.5) yields $\alpha_{PQ}(X, X) = 0$ a.s.

Similarly, $\alpha_{PQ}(X, X) = 0$ a.s. and $\alpha_{PQ}(X, X) = 0$ a.s., and thus

$$\langle \alpha, X \otimes X \rangle = \alpha(X, X) = \alpha_{PQ}(X, X) + \alpha_{QP}(X, X) + \alpha_{QQ}(X, X) = 0 \quad \text{a.s.}$$

This holds for every $\alpha \perp \widetilde{M}$, and thus $X \otimes X$ is weakly a.s. separably valued. □

Theorem 9.8. *Suppose that H is a Hilbert space and that X is a weakly measurable H-valued random variable.*

(i) *The injective second moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Dunford sense if and only if $\mathbb{E}\, |\langle y, X \rangle|^2 < \infty$ for every $y \in H$.*

(ii) *The injective second moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Pettis sense if and only if the random variables $|\langle y, X \rangle|^2$, for $y \in H$ with $\|y\| \leqslant 1$, are uniformly integrable.*

(iii) *The injective second moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Bochner sense if and only if X is a.s. separably valued and $\mathbb{E}\, \|X\|^2 < \infty$.*

Proof. (i): By Theorem 6.10(ii)(c), using Theorem 9.3.

(ii): By Theorem 9.3, Lemma 9.7 and Theorem 6.19(ii)(c).

(iii): A special case of Theorem 6.7. □

In the Hilbert space case, the different types of integrability can also be characterized by the value of the moment.

Theorem 9.9. *Suppose that H is a Hilbert space and that X is a weakly measurable H-valued random variable such that the injective second moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Dunford sense. Regard $\mathbb{E}\, X^{\check{\otimes}2}$ as a bounded operator in $B(H)$.*

(i) *The injective second moment exists in Pettis sense if and only if $\mathbb{E}\, X^{\check{\otimes}2} \in \mathcal{K}(H)$.*

(ii) *The injective second moment exists in Bochner sense if and only if X is a.s. separably valued and $\mathbb{E}\, X^{\check{\otimes}2} \in \mathcal{N}(H)$.*

Proof. (i): If the moments exist in Pettis sense, then $\mathbb{E}\, X^{\check{\otimes}2} \in H \check{\otimes} H = \mathcal{K}(H)$.

Conversely, suppose that $\mathbb{E}\,X^{\check{\otimes}2} \in \mathcal{K}(H)$. Let E be any event. Then

$$\mathbb{E}(\mathbf{1}_E X)^{\check{\otimes}2} + \mathbb{E}(\mathbf{1}_{E^{\mathrm{c}}} X)^{\check{\otimes}2} = \mathbb{E}(\mathbf{1}_E X^{\check{\otimes}2}) + \mathbb{E}(\mathbf{1}_{E^{\mathrm{c}}} X^{\check{\otimes}2}) = \mathbb{E}\,X^{\check{\otimes}2}.$$

Since $\mathbb{E}(\mathbf{1}_E X)^{\check{\otimes}2} \geqslant 0$ and $\mathbb{E}(\mathbf{1}_{E^{\mathrm{c}}} X)^{\check{\otimes}2} \geqslant 0$ (in operator sense) by Theorem 9.4, it follows that

$$0 \leqslant \mathbb{E}(\mathbf{1}_E X^{\check{\otimes}2}) \leqslant \mathbb{E}\,X^{\check{\otimes}2}. \tag{9.6}$$

It is easily verified that if $S, T \in B(H)$ with $0 \leqslant S \leqslant T$ and T compact, then S too is compact. (For example because $S^{1/2} = VT^{1/2}$ for some $V \in B(H)$, and thus $S = S^{1/2}(S^{1/2})^* = VTV^*$.) Hence, (9.6) implies that $\mathbb{E}(\mathbf{1}_E X^{\check{\otimes}2}) \in \mathcal{K}(H) = H\check{\otimes}H$ for every event E, which means that $X^{\check{\otimes}2}$ satisfies the definition of Pettis integrability.

(ii): If $\mathbb{E}\,X^{\check{\otimes}2}$ exists in Bochner sense, then $\mathbb{E}\,X^{\hat{\otimes}2}$ too exists in Bochner sense by Theorem 6.7. Hence, $\mathbb{E}\,X^{\check{\otimes}2} = \mathbb{E}\,X^{\hat{\otimes}2} \in H\widehat{\otimes}H = \mathcal{N}(H)$. Moreover, X is a.s. separably valued by Theorem 6.7.

Conversely, suppose that $\mathbb{E}\,X^{\check{\otimes}2} \in \mathcal{N}(H)$ and X is a.s. separably valued. Let M be separable subspace of H such that $X \in M$ a.s. and let $\{e_n\}_n$ be a (countable) ON basis in M. Then $\|X\|^2 = \sum_n \langle X, e_n\rangle^2$ a.s., and thus

$$\begin{aligned}\mathbb{E}\,\|X\|^2 &= \sum_n \mathbb{E}\langle X, e_n\rangle^2 = \sum_n \langle \mathbb{E}\,X^{\check{\otimes}2} e_n, e_n\rangle = \operatorname{Tr}\bigl(\mathbb{E}\,X^{\check{\otimes}2}\bigr)\\ &\leqslant \|\,\mathbb{E}\,X^{\check{\otimes}2}\|_{\mathcal{N}(H)} < \infty.\end{aligned}$$

Thus $\mathbb{E}\,X^{\check{\otimes}2}$ exists in Bochner sense by Theorem 6.7. □

Theorem 9.10. *Suppose that H is a Hilbert space and that X is a weakly measurable H-valued random variable such that the injective second moment $\mathbb{E}\,X^{\check{\otimes}2}$ exists in Dunford sense. Regard $\mathbb{E}\,X^{\check{\otimes}2}$ as a bounded operator in $B(H)$. If $\mathbb{E}\,X^{\check{\otimes}2} \in \mathcal{N}(H)$ and X is a.s. separably valued, then the projective second moment $\mathbb{E}\,X^{\check{\otimes}2}$ exists in Bochner sense.*

Proof. By Theorem 9.9(ii) and Theorem 6.7. □

Remark 9.11. If the projective second moment exists in Bochner or Pettis sense, it is an element of $H\widehat{\otimes}H = \mathcal{N}(H)$. The second injective moment is the same, and is then thus a trace class operator.

Theorem 9.10 gives a converse when X is a.s. separably valued. However, the converse does not hold in general; a weakly measurable random variable in a non-separable Hilbert space can have an injective moment in Pettis sense that is a trace class operator, even if the projective second moment does not exist; see Example 7.4.

For the projective second moment, all three senses coincide for H-valued random variables, provided $\dim(H)$ is not too large.

Definition 9.12. A cardinal $\mathfrak{m}$ is *real-measurable* if there exists a set S with cardinality $|S| = \mathfrak{m}$ and a probability measure μ defined on the σ-field

2^S of *all* subsets of S that is *diffuse*, i.e., such that $\mu\{s\} = 0$ for every $s \in S$. (Obviously, then such a measure μ exists for every set S with $|S| = \mathfrak{m}$.)

A cardinal $\mathfrak{m}$ is *measurable* if there exists such a measure that takes only the values 0 and 1.

Remark 9.13. If measurable cardinals exist, they have to be very large; larger than the first strongly inaccessible cardinal. It is consistent with the standard ZFC axioms for set theory to assume that there are no strongly inaccessible cardinals, and thus no measurable cardinals. Whether it also is consistent to assume the existence of measurable cardinals is not known. See [40, Chapter IX.3–4] and [39].

Real-measurable cardinals may be smaller. There exists a real-measurable cardinal that is non-measurable $\iff$ $\mathfrak{c}$ is real-measurable $\iff$ Lebesgue measure on $[0,1]$ can be extended to all subsets of $[0,1]$, see [64] and [62, Section 16.2].

If the Continuum Hypothesis holds, then $\mathfrak{c}$ is not real-measurable, see [4] and [64], and thus every real-measurable cardinal is measurable and thus extremely large. Consequently, it is consistent to assume that there are no real-measurable cardinals at all. In this case, the following theorem applies to all Hilbert spaces without qualification. (The separable case was given in Example 7.1.)

Theorem 9.14. *Let H be a Hilbert space such that $\dim H$ is a non-real-measurable cardinal. If X is an H-valued random variable, then the following are equivalent.*

(i) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Dunford sense.*
(ii) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Pettis sense.*
(iii) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Bochner sense.*
(iv) *X is a.s. separably valued, $X^{\hat{\otimes}2}$ is weakly measurable in $H^{\hat{\otimes}2}$ and $\mathbb{E}\,\|X\|^2 < \infty$.*

Proof. (iii) $\implies$ (ii) $\implies$ (i) is trivial and (iv) $\implies$ (iii) for any space by Theorem 5.1, using Lemma 6.5 and Theorem 3.3. It remains to show that (i) $\implies$ (iv).

Suppose that $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense and let $z := \mathbb{E}\, X^{\hat{\otimes}2}$. By definition, $X^{\otimes 2}$ is weakly measurable in $H^{\hat{\otimes}2}$.

Let M be a closed subspace of H and let $P_M : H \to M$ be the orthogonal projection. Then $\alpha_M(x,y) := \langle P_M x, P_M y\rangle$ is a bounded bilinear form on H; α_M can by Theorem 4.6 be regarded as a continuous linear functional on $H \hat{\otimes} H$ and

$$\langle \alpha_M, z\rangle = \mathbb{E}\langle \alpha_M, X \otimes X\rangle = \mathbb{E}\,\alpha_M(X,X) = \mathbb{E}\,\|P_M X\|^2.$$

Hence, $\|P_M X\|^2$ is measurable and its expectation is finite. Define

$$\mu_M := \mathbb{E}\,\|P_M X\|^2 < \infty.$$

In particular, $\mathbb{E}\,\|X\|^2 = \mu_H < \infty$.

If $M_1 \subseteq M_2$, then $\|P_{M_1}X\| \leqslant \|P_{M_2}X\|$ and thus $\mu_{M_1} \leqslant \mu_{M_2}$. In particular, $\mu_M \leqslant \mu_H$ for every closed subspace $M \subseteq H$, so $\{\mu_M : M \subseteq H\}$ is bounded.

Furthermore, if $M_1 \perp M_2$, then $\|P_{M_1\oplus M_2}X\|^2 = \|P_{M_1}X\|^2 + \|P_{M_2}X\|^2$ and thus

$$\mu_{M_1\oplus M_2} = \mu_{M_1} + \mu_{M_2}. \tag{9.7}$$

Let $\mathcal{Z}$ be the set of separable closed subspaces of H, and let

$$\mu_* := \sup\{\mu_M : M \in \mathcal{Z}\}. \tag{9.8}$$

Thus $\mu_* \leqslant \mu_H < \infty$. There exist $M_n \in \mathcal{Z}$ such that $\mu_{M_n} > \mu_* - 1/n$, $n \geqslant 1$. Let M_* be the closed linear hull of $\bigcup_1^\infty M_n$. Then M_* is separable so $M_* \in \mathcal{Z}$, and $M_n \subseteq M_*$ so $\mu_{M_*} \geqslant \mu_{M_n} > \mu_* - 1/n$. Consequently, $\mu_{M_*} = \mu_*$ and the supremum in (9.8) is attained.

By (9.7),

$$\mu_H = \mu_{M_*} + \mu_{M_*^\perp} = \mu_* + \mu_{M_*^\perp}. \tag{9.9}$$

Assume first $\mu_* = \mu_H$. Then (9.9) yields

$$0 = \mu_{M_*^\perp} = \mathbb{E}\,\|P_{M_*^\perp}X\|^2,$$

and thus $P_{M_*^\perp}X = 0$ a.s.; hence $X \in M_*$ a.s. so X is a.s. separably valued.

It remains to show that if $\mu_* < \mu_H$, then $\dim H$ is real-measurable, which contradicts our assumption. In this case, choose an ON basis $\{e_s\}_{s\in S}$ in $M_*^\perp$. For each subset $A \subseteq S$, let M_A be the closed subspace of H spanned by $\{e_s\}_{s\in A}$ and let $\mu(A) := \mu_{M_A} = \mathbb{E}\,\|P_{M_A}X\|^2$. Then μ is finitely additive by (9.7). Moreover, if $A_n \nearrow A$, then $\|P_{M_{A_n}}X\| \nearrow \|P_{M_A}X\|$ and thus $\mu(A_n) \nearrow \mu(A)$ by monotone convergence. Consequently, μ is a σ-additive measure defined on $(S, 2^S)$. For any $s \in S$, $M_{\{s\}} = \mathbb{R}e_s$ is one-dimensional and thus $M_* \oplus M_{\{s\}}$ is separable, so

$$\mu_* + \mu_{M_{\{s\}}} = \mu_{M_*} + \mu_{M_{\{s\}}} = \mu_{M_*\oplus M_{\{s\}}} \leqslant \mu_*.$$

Hence, $\mu\{s\} := \mu_{M_{\{s\}}} = 0$, which shows that the measure μ is diffuse. On the other hand, $M_S = M_*^\perp$ and thus by (9.9)

$$\mu(S) = \mu_{M_S} = \mu_{M_*^\perp} = \mu_H - \mu_* > 0.$$

Consequently, by normalizing μ we obtain a diffuse probability measure on $(S, 2^S)$, which shows that $|S|$ is a real-measurable cardinal. Since $\dim(H) \geqslant |S|$, $\dim(H)$ is real-measurable too. (In fact, $\dim H = |S| + \dim M_* = |S| + \aleph_0 = |S|$.) □

The assumption in Theorem 9.14 that $\dim H$ is not real-measurable is necessary as is seen by the following example. (Cf. the related Example 7.4.) Nevertheless, Theorem 9.14 shows that the projective second moments for practical applications only can be used for Bochner measurable random variables, i.e., in the a.s. separably valued case.

Example 9.15. Suppose that $\dim H$ is a real-measurable cardinal. Let μ be a diffuse probability measure on $(S, 2^S)$ for some set S with $|S| = \dim H$. Since then $H \cong \ell^2(S)$, we may assume that $H = \ell^2(S)$.

Define $X : (S, 2^S, \mu) \to H = \ell^2(S)$ by $X(s) = e_s$. Then X is bounded, and since *every* function on $(S, 2^S, \mu)$ is measurable, X is weakly measurable in H and $X \otimes X$ is weakly measurable in $H\hat{\otimes}H$. Hence $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense. (X and $X \otimes X$ are also Borel measurable, for the same reason.) Furthermore, if α is the bilinear form on H given by the inner product, $\langle \mathbb{E}\, X^{\hat{\otimes}2}, \alpha\rangle = \mathbb{E}\, \alpha(X, X) = \mathbb{E}\, \|X\|^2 = 1$ and thus $\mathbb{E}\, X^{\hat{\otimes}2} \neq 0$.

On the other hand, the argument in Example 7.4 shows that X is not a.s. separably valued. (In fact, if M is any separable subspace of H, then $X \perp M$ a.s.) By Theorem 6.7, $\mathbb{E}\, X^{\hat{\otimes}2}$ does not exist in Bochner sense.

Note further that $\langle x^*, X\rangle = 0$ a.s., for every $x^* \in H^* = H$. Thus the injective second moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Dunford (and Pettis) sense by Theorem 9.8, and (6.3) shows that $\mathbb{E}\, X^{\check{\otimes}2} = 0$. However, using Dunford senses, $\mathbb{E}\, X^{\hat{\otimes}2} \neq 0$, as shown above, and $\mathbb{E}\, X^{\check{\otimes}2} = \iota^{**}(\mathbb{E}\, X^{\hat{\otimes}2})$ by (6.2). Since $\iota : H^{\hat{\otimes}2} \to H^{\check{\otimes}2}$ is injective, this shows that $\mathbb{E}\, X^{\hat{\otimes}2} \notin H^{\hat{\otimes}2}$; hence $\mathbb{E}\, X^{\hat{\otimes}2}$ does not exist in Pettis sense.

Example 9.16. Consider Gaussian random variables in a Hilbert space H. As shown in Example 7.6, if X is Gaussian, then $\mathbb{E}\, X^{\hat{\otimes}k}$ and $\mathbb{E}\, X^{\check{\otimes}k}$ exist in Bochner sense. In particular, $\mathbb{E}\, X^{\hat{\otimes}2} = \mathbb{E}\, X^{\check{\otimes}2}$ is an element of $H\hat{\otimes}H = \mathcal{N}(H)$, i.e., a trace class operator. By Theorem 9.4, $\mathbb{E}\, X^{\hat{\otimes}2} = \mathbb{E}\, X^{\check{\otimes}2}$ is a positive trace class operator.

Conversely, if Σ is any positive trace class operator in a Hilbert space H, then there exists by the spectral theorem for compact self-adjoint operators, see e.g. [13, Corollary II.5.4], an ON set $(e_n)_{n=1}^\infty$ in H such that

$$\Sigma = \sum_{n=1}^{\infty} \lambda_n e_n \otimes e_n, \tag{9.10}$$

where $\lambda_n \geqslant 0$ and $\sum_{n=1}^\infty \lambda_n = \|\Sigma\|_{\mathcal{N}(H)} < \infty$. Let $(\xi_n)_{n=1}^\infty$ be i.i.d. standard normal variables and define $X := \sum_{n=1}^\infty \lambda^{1/2}\xi_n e_n$; it is easily seen that this sum converges a.s., that X is Gaussian, and that $\mathbb{E}\, X^{\hat{\otimes}2} = \sum_{n=1}^\infty \lambda_n e_n \otimes e_n = \Sigma$.

Consequently, the second moment of a Gaussian random variable in a Hilbert space is a positive trace class operator, and can be any such operator. Moreover, the second moment determines the higher moments by (7.5), and thus the distribution by Theorem 17.3 below.

10. $L^p(\mu)$

Let $1 \leqslant p < \infty$ and let $B = L^p(\mu) = L^p(S, \mathcal{S}, \mu)$, where μ is a σ-finite measure on a measurable space $(S, \mathcal{S})$. Note first that $L^p(\mu)$ has the approximation property, see e.g. [57, Example 4.5]. (It suffices to consider the case when μ is a probability measure, and then (i) in Section 8 is satisfied by using conditional expectations on suitable finite sub-σ-fields; we omit the

details.) Furthermore, $L^p(S,\mathcal{S},\mu)^* = L^q(S,\mathcal{S},\mu)$, where $q \in (1,\infty]$ is the conjugate exponent given by $p^{-1}+q^{-1}=1$ [22, Corollary IV.8.1,5].

If $X:(s,\omega)\mapsto X(s,\omega)\in\mathbb{R}$ is an $(\mathcal{S}\times\mathcal{F})$-measurable function on $S\times\Omega$, then $X(s)=X(s,\cdot)$ is a (real-valued) random variable for every s; moreover, for each $\omega\in\Omega$, $X(\cdot,\omega)$ is a function on S, and if further $\int_S |X(s,\omega)|^p\,\mathrm{d}\mu(s)<\infty$ a.s., then X can be regarded as a mapping $(\Omega,\mathcal{F},\mathbb{P})\to L^p(\mu)$, i.e., as a random variable in $B=L^p(\mu)$. This random variable is Bochner measurable, as stated in the following lemma; see [22, III.11.16–17].

Lemma 10.1. *If $X:(s,\omega)\mapsto X(s,\omega)\in\mathbb{R}$ is an $(\mathcal{S}\times\mathcal{F})$-measurable function on $S\times\Omega$ and $\int_S |X(s,\omega)|^p\,\mathrm{d}\mu(s)<\infty$ a.s., then X can be regarded as a Bochner measurable mapping $(\Omega,\mathcal{F},\mathbb{P})\to L^p(\mu)$, i.e., a Bochner measurable $L^p(\mu)$-valued random variable; conversely, every Bochner measurable $L^p(\mu)$-valued random variable is (a.s.) represented in this way by some $(\mu\times\mathbb{P})$-a.e. unique $(\mathcal{S}\times\mathcal{F})$-measurable X on $S\times\Omega$.* □

Theorem 10.2. *Let $1\leqslant p<\infty$, let q be the conjugate exponent given by $p^{-1}+q^{-1}=1$, and let $(S,\mathcal{S},\mu)$ be a σ-finite measure space.*

(i) *Suppose that $X:S\times\Omega\to\mathbb{R}$ is $(\mathcal{S}\times\mathcal{F})$-measurable and that $\|X\|_{L^p(\mu)}^p := \int_S |X(s,\omega)|^p\,\mathrm{d}\mu(s)<\infty$ a.s. Regard X as an $L^p(\mu)$-valued random variable and suppose further that $\mathbb{E}\,\|X\|_{L^p(\mu)}^k<\infty$. Then, $\mathbb{E}\,X^{\hat\otimes k}\in L^p(S,\mu)^{\hat\otimes k}$ and $\mathbb{E}\,X^{\check\otimes k}\in L^p(S,\mu)^{\check\otimes k}$ exist in Bochner sense. Furthermore, $\mathbb{E}\,X^{\hat\otimes k}$ and $\mathbb{E}\,X^{\check\otimes k}$ are represented by the a.e. finite function*

$$\Phi_k(s_1,\dots,s_k) := \mathbb{E}\big(X(s_1,\omega)\cdots X(s_k,\omega)\big) \tag{10.1}$$

in the sense that if $g_1,\dots,g_k\in L^q(\mu)=(L^p(\mu))^*$, then

$$\begin{aligned}\langle \mathbb{E}\,X^{\hat\otimes k}, g_1\otimes\cdots\otimes g_k\rangle &= \langle \mathbb{E}\,X^{\check\otimes k}, g_1\otimes\cdots\otimes g_k\rangle\\ &= \int_{S^k}\Phi_k(s_1,\dots,s_k)g_1(s_1)\cdots g_k(s_k)\,\mathrm{d}\mu(s_1)\cdots\,\mathrm{d}\mu(s_k).\end{aligned} \tag{10.2}$$

(ii) *Let Y be another $L^p(\mu)$-valued random variable represented by an $(\mathcal{S}\times\mathcal{F})$-measurable function $Y:S\times\Omega\to\mathbb{R}$ such that $\mathbb{E}\,\|Y\|_{L^p(\mu)}^k<\infty$, and let*

$$\Psi_k(s_1,\dots,s_k) := \mathbb{E}\big(Y(s_1,\omega)\cdots Y(s_k,\omega)\big). \tag{10.3}$$

Then the following are equivalent

(a) $\mathbb{E}\,X^{\hat\otimes k}=\mathbb{E}\,Y^{\hat\otimes k}$;

(b) $\mathbb{E}\,X^{\check\otimes k}=\mathbb{E}\,Y^{\check\otimes k}$;

(c) *For any $g_1,\dots,g_k\in L^q(\mathbb{P})$,*

$$\mathbb{E}\big(\langle g_1,X\rangle\cdots\langle g_k,X\rangle\big)=\mathbb{E}\big(\langle g_1,Y\rangle\cdots\langle g_k,Y\rangle\big);$$

(d) $\Phi_k=\Psi_k$ *a.e. on S^k.*

Proof. (i): X is Bochner measurable by Lemma 10.1, and thus $\mathbb{E}\, X^{\hat{\otimes}k}$ and $\mathbb{E}\, X^{\check{\otimes}k}$ exist in Bochner sense by Lemma 6.3 and Theorem 5.1 (or by Theorem 6.7). Furthermore, by (6.3) and Fubini's theorem, using Lemma 10.1,

$$\begin{aligned}\langle \mathbb{E}\, X^{\check{\otimes}k}, g_1 \otimes \cdots \otimes g_k\rangle &= \mathbb{E}\big(\langle g_1, X\rangle \cdots \langle g_k, X\rangle\big)\\ &= \mathbb{E}\int_{S^k} X(s_1)\cdots X(s_k)\, g_1(s_1)\cdots g_k(s_k)\,\mathrm{d}\mu(s_1)\cdots \mathrm{d}\mu(s_k)\\ &= \int_{S^k} \mathbb{E}\big(X(s_1)\cdots X(s_k)\big) g_1(s_1)\cdots g_k(s_k)\,\mathrm{d}\mu(s_1)\cdots \mathrm{d}\mu(s_k),\end{aligned}$$

showing (10.2).

(ii): The moments exist in Bochner sense by (i). Then (a) $\iff$ (b) by Theorem 8.3 and (6.1), (b) $\iff$ (c) by Corollary 6.14 and (b) $\iff$ (d) by (10.2). □

Example 10.3. In the special case $p = 2$, $L^2(\mu)$ is a Hilbert space, and we can also apply the results of Section 9. (Since X is Bochner measurable by Lemma 10.1, and thus a.s. separably valued, it suffices to consider a separable subspace of $L^2(\mu)$.) In particular, Theorem 9.8 and Theorem 9.9 give conditions for the existence of the injective second moment in the different senses, while Theorem 9.14 shows that for the second projective moment, the different senses coincide.

Example 10.4. Another interesting special case is $p = 1$. We have $L^1(S,\mu)^{\hat{\otimes}k} = L^1(S^k, \mu^k)$, see [57, Exercise 2.8]; thus $\mathbb{E}\, X^{\hat{\otimes}k} \in L^1(S^k,\mu^k)$ (when it exists in Bochner or Pettis sense); clearly $\mathbb{E}\, X^{\hat{\otimes}k}$ equals the function Φ_k in (10.1) when Theorem 10.2 applies.

Example 7.3 shows that $\mathbb{E}\, X^{\hat{\otimes}k}$ may exist in Pettis sense in $L^1(S,\mu)^{\hat{\otimes}k} = L^1(S^k,\mu^k)$ without existing in Bochner sense. In this case $\mathbb{E}\, X^{\hat{\otimes}k}$ is still given by a function Φ_k in $L^1(S^k,\mu^k)$ but the pointwise formula (10.1) may fail; in Example 7.3, we have $X(s,\omega) = \pm a_N$ for all s and ω, and thus $|X(s_1,\omega)X(s_2,\omega)| = a_N^2$, so the expectation in (10.1) exists (for any s_1, s_2) only when $\mathbb{E}\, a_N^2 = \sum_n p_n a_n^2 < \infty$, which in this case is the condition for $\mathbb{E}\, X^{\hat{\otimes}2}$ to exist in Bochner sense, see Example 7.3; we may choose $p_n a_n^2 = 1/n$ to obtain our counterexample. In this example, $\mathbb{E}\, X^{\hat{\otimes}2}$ equals $\sum_n p_n a_n^2 r_n \otimes r_n$, where the sum converges in L^2, and thus in L^1; the sum also converges pointwise a.e. (since it is a sum of independent random variables defined on $[0,1]^2$), but (when $p_n a_n^2 = 1/n$, say) does not converge absolutely at any point.

We do not know any necessary and sufficient conditions for the existence of $\mathbb{E}\, X^{\hat{\otimes}k}$ in Pettis sense in $L^1(S,\mu)^{\hat{\otimes}k} = L^1(S^k,\mu^k)$, nor for the existence of injective moments beyond Theorem 6.10, where (ii)(b) applies by Lemma 10.1. (Even for $k = 1$, the existence of the mean $\mathbb{E}\, X$ in Pettis sense in $L^1(\mu)$ seems difficult to characterize exactly, since it is essentially equivalent to $X \in L^1(\mu)\check{\otimes} L^1(\mathbb{P})$, which has no simple description, see Remark 5.20.)

Since $L^1(S^k,\mu^k)$ does not contain any subspace isomorphic to c_0 (e.g. as a consequence of [57, Theorem 6.31 and Corollary 6.21]), it follows from Theorem 5.13 that if X is Bochner measurable (e.g. by Lemma 10.1), then $\mathbb{E}\,X^{\hat\otimes k}$ exists in Pettis sense as soon as it exists in Dunford sense.

Example 10.5. Taking $S=\mathbb{N}$ with counting measure, we obtain ℓ^p, $1\leqslant p<\infty$. (See e.g. Examples 5.15, 5.16 and 7.2.) In this case, $X=(X_n)_{n=1}^\infty$ and Φ_k is by (10.1) the function $\mathbb{E}\bigl(X_{n_1}\cdots X_{n_k}\bigr)$ on $\mathbb{N}^k$. Consequently, by Theorem 10.2, if $\mathbb{E}\,\|X\|_{\ell^p}^k<\infty$ and $\mathbb{E}\,\|Y\|_{\ell^p}^k<\infty$, then $\mathbb{E}\,X^{\hat\otimes k}=\mathbb{E}\,Y^{\hat\otimes k}$ if and only if all mixed k:th moments of $(X_n)_{n=1}^\infty$ and $(Y_n)_{n=1}^\infty$ coincide.

Note that moments may exist in Pettis sense also under weaker assumptions, see Example 5.15.

Note that the projective tensor product $\ell^p\hat\otimes\ell^p$ with $2<p<\infty$ and the injective tensor product $\ell^p\check\otimes\ell^p$ with $1<p<2$ are reflexive, see [57, Corollary 4.24]. In these cases, at least, the second moment thus exists in Pettis sense as soon as it exists in Dunford sense.

11. $C(K)$

In this section we study the case $B=C(K)$ where K is a compact space. (By compact we mean compact Hausdorff.) The perhaps most important example is $C[0,1]$.

Note that several other Banach spaces are isomorphic to $C(K)$ for some compact K. Hence the results in this section apply to these spaces too.

Example 11.1. Let $C_{\mathsf b}(Z)$ be the space of bounded continuous functions on a completely regular topological space Z. Then $C_{\mathsf b}(Z)=C(\beta Z)$, where βZ is the Stone–Čech compactification of Z, see e.g. [27, Section 3.6], [30, Chapter 6] and [13, Section V.6 and Exercise VIII.2.7]. Note that Z is a dense subspace of βZ, and that every bounded continuous function on Z has a unique continuous extension to βZ.

One important example is

$$\ell^\infty=C_{\mathsf b}(\mathbb{N})=C(\beta\mathbb{N}). \tag{11.1}$$

Example 11.2. We shall see in Theorem 14.2 that $D[0,1]=C(\widehat I)$ for a compact space $\widehat I$.

Example 11.3. If B is any complex commutative C^*-algebra, then $B\cong C_{\mathbb{C}}(\Sigma)$, the space of complex-valued continuous functions on the maximal ideal space Σ, see [13, Section VIII.2], and thus the subset $B_{\mathbb{R}}:=\{f:f=f^*\}$ of hermitean (i.e., real) elements is isomorphic to $C(\Sigma)$. This includes Examples 11.1–11.2.

Example 11.4. Let L be a locally compact space and $L^*=L\cup\{\infty\}$ its one-point compactification. Then $C_0(L)=\{f\in C(L^*):f(\infty)=0\}$ is a subspace of codimension 1 in $C(L^*)$. Hence, if X is a random variable in $C_0(L)$, we can regard it as a random variable in $C(L^*)$. Note that $C_0(L)$ is a

complemented subspace. (Every subspace of finite codimension in a Banach space is complemented.)

In particular, $c_0 = C_0(\mathbb{N})$ is a (complemented) subspace of codimension 1 in $c = C(\mathbb{N}^*) = C(\mathbb{N} \cup \{\infty\})$. (In fact, c_0 is also isomorphic to c, by the mapping $(a_n)_1^\infty \mapsto (a_{n+1} + a_1)_1^\infty$, but it seems more convenient to use the inclusion.)

We begin by noting some well-known facts. See e.g. [57, Example 4.2], [57, Section 3.2] and [13, Theorem V.6.6], respectively, for proofs.

Theorem 11.5. *$C(K)$ has the approximation property, for any compact K.* □

Theorem 11.6. *If $K_1, \dots, K_k$ are compact spaces, then $C(K_1)\check{\otimes}\cdots\check{\otimes}C(K_k) = C(K_1 \times \cdots \times K_k)$ (isometrically) with the natural identification. In particular, $C(K)^{\check{\otimes}k} = C(K^k)$.* □

Theorem 11.7. *$C(K)$ is separable if and only if K is metrizable.* □

Corollary 11.8. *If K is a compact space, then $C(K)^{\hat{\otimes}k}$ can be regarded as a subspace of $C(K)^{\check{\otimes}k} = C(K^k)$. (As a vector space; typically, the norms differ.)*

Proof. By Theorem 11.5 and Theorem 8.3, $\iota : C(K)^{\hat{\otimes}k} \to C(K)^{\check{\otimes}k}$ is injective. Furthermore, $C(K)^{\check{\otimes}k} = C(K^k)$ by Theorem 11.6. □

Remark 11.9. Except in trivial cases ($k = 1$ or K finite), $C(K)^{\hat{\otimes}k}$ is not a closed subspace of $C(K)^{\check{\otimes}k}$, and thus the norms are different and not even equivalent on $C(K)^{\hat{\otimes}k}$. This is implicitly shown for c_0 by Littlewood [44], who showed (in our terminology) the existence of a bounded bilinear form $\alpha \in (c_0^{\hat{\otimes}2})^*$ which does not belong to $(c_0^{\check{\otimes}2})^* = c_0(\mathbb{N}^2)^* = \ell^1(\mathbb{N}^2)$; the result can be transfered to $C(K)$ for any infinite compact K. (See also the proof of Theorem 16.2 for an argument from [65] for $C[0,1]$.)

Theorem 11.10. *Let X be a $C(K)$-valued random variable, where K is a compact space. If $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Bochner or Pettis sense, then it is the function in $C(K)^{\check{\otimes}k} = C(K^k)$ given by*

$$\mathbb{E}\, X^{\check{\otimes}k}(t_1, \dots, t_k) = \mathbb{E}\big(X(t_1) \cdots X(t_k)\big). \tag{11.2}$$

Proof. $\mathbb{E}\, X^{\check{\otimes}k} \in C(K)^{\check{\otimes}k} = C(K^k)$ by Theorem 11.6. Furthermore, the point evaluations δ_t, $t \in K$, are continuous linear functionals on $C(K)$, and thus

$$\mathbb{E}\, X^{\check{\otimes}k}(t_1, \dots, t_k) = \langle \mathbb{E}\, X^{\check{\otimes}k}, \delta_{t_1} \otimes \dots \otimes \delta_{t_k} \rangle = \mathbb{E}\big(X(t_1) \cdots X(t_k)\big). \tag{11.3}$$

□

Note that the function (11.2) is the standard k:th moment function for a stochastic process; in particular, for $k = 2$ and $\mathbb{E}\, X = 0$, $\mathbb{E}\, X^{\check{\otimes}2}$ is the covariance function.

Example 11.11. Let W be standard Brownian motion in $C[0,1]$, see Example 5.12. Then all (projective and injective) moments exist in Bochner sense, e.g. by Theorem 11.25 below. All odd moments vanish by symmetry. $\mathbb{E}\, W^{\check{\otimes}2}$ is the covariance function $\mathbb{E}(W(s)W(t)) = \operatorname{Cov}(W(s),W(t)) = s\wedge t$ regarded as a continuous function in $C([0,1]^2)$, and $\mathbb{E}\, W^{\hat{\otimes}2}$ is the same function regarded as an element of the subspace $C[0,1]^{\hat{\otimes}2}$. Similarly, $\mathbb{E}\, W^{\hat{\otimes}4} = \mathbb{E}\, W^{\check{\otimes}4}$ is the function in $C([0,1]^4)$ given by, see e.g. [37],

$$\begin{aligned}\mathbb{E}\big(W(t_1)W(t_2)W(t_3)W(t_4)\big)&\\ = (t_1\wedge t_2)(t_3\wedge t_4) + (t_1\wedge t_3)(t_2\wedge t_4) + (t_1\wedge t_4)(t_2\wedge t_3).&\end{aligned}\tag{11.4}$$

Remark 11.12. If $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Dunford sense, then (11.2) still defines a bounded function on K^k, but the function is not necessarily continuous (Example 11.28); moreover, this function by itself does not in general determine $\mathbb{E}\, X^{\check{\otimes}k} \in C(K^k)^{**}$ uniquely, not even for $k=1$, and not even if it happens to be continuous. (Note that the point evaluations do not form a total set in $C(K^k)^*$.) Indeed, in Example 11.30 we shall see a K and a random variable $Z\in C(K)$ such that $\mathbb{E}\, Z$ exists in Dunford sense (but not Pettis sense) with $\mathbb{E}\, Z\neq 0$, but $\mathbb{E}(Z(t))=0$ for every $t\in K$. (This cannot happen when K is metric, see Corollary 11.17; see further Remarks 11.18 and 11.24.) We therefore prefer the Pettis or Bochner case for applications.

Apart from the previously defined σ-fields on $C(K)$, we let $\mathcal{C}$ be the σ-field generated by the point evaluations $f\mapsto f(t)$, $t\in K$. Thus $X:\Omega\to C(K)$ is $\mathcal{C}$-measurable if and only if $X(t)$ is measurable for every $t\in K$. Since $\delta_t\in C(K)^*$, it is immediate that every weakly measurable random variable X is $\mathcal{C}$-measurable; we shall see that the converse holds when K is metrizable but not in general (Example 11.31).

Further, let $\mathcal{B}(K)$ be the Borel σ-field on K.

We also let $M(K)$ be the space of signed Borel measures on K, and $\mathcal{M}$ be the σ-field on $M(K)$ generated by the maps $\mu\mapsto\langle\mu,f\rangle := \int_K f\,\mathrm{d}\mu$, $f\in C(K)$. Note that every $\mu\in M(K)$ defines a continuous linear functional on $C(K)$, so there is a bounded linear map $M(K)\to C(K)^*$. Moreover, by the Riesz representation theorem (see e.g. [22, Theorem IV.6.3], [12, Theorem 7.3.5], [13, Theorem III.5.7]), this map is an isometric bijection of the subspace $M_{\mathsf{r}}(K)$ of regular measures onto $C(K)^*$. (In many cases $M_{\mathsf{r}}(K)=M(K)$, for example when K is a metrizable compact space, see [12, Proposition 8.1.10].)

Theorem 11.7 implies that if K is a compact metric space, there are no measurability problems; some results are given in the following lemma. However, if K is not metrizable, and thus $C(K)$ is not separable, the situation is more complicated; we have to be more careful with measurability in statements, but even so, many results below do not hold for arbitrary compact K, see the (counter)examples at the end of the section. We thus state most of our results for the metrizable case only. (See Sections 12 and

15–16, together with Examples 11.4 and 11.2, for examples of non-separable spaces $C(K)$ where most results hold, although some new arguments are required.)

Remark 11.13. All spaces $C(K)$ where K is an uncountable compact metric space are isomorphic as Banach spaces [46], [50] (although not isometrically); hence they are all the same as $C[0,1]$ from an abstract point of view. We shall, however, not use this; we prefer to regard the spaces concretely.

Lemma 11.14. *Let K be a metrizable compact space.*

(i) *The mapping $(f,t) \to f(t)$ is jointly $(\mathcal{C}\times\mathcal{B}(K))$-measurable on $C(K)\times K$.*

(ii) *The mapping $(f,\mu) \to \int_K f \,\mathrm{d}\mu$ is jointly $(\mathcal{C}\times\mathcal{M})$-measurable on $C(K)\times M(K)$.*

In particular, $(f,x^*) \to \langle x^*, f\rangle$ is jointly $(\mathcal{C}\times\mathcal{M})$-measurable on $C(K)\times M_{\mathsf{r}}(K) = C(K)\times C(K)^*$.

Proof. Choose a metric d on K. For each $n \geqslant 1$ there is a finite covering of K by open sets U_{ni}, $1 \leqslant i \leqslant N_n$, of diameters $< 1/n$. We may find a partition of unity subordinate to $\{U_{ni}\}_i$, i.e., a set of functions $f_{ni} \in C(K)$ such that $f_{ni} \geqslant 0$, $\sum_i f_{ni} = 1$ and $\{t \in K : f_{ni}(t) > 0\} \subseteq U_{ni}$. (For example, let $g_{ni}(t) := d(t, U_{ni}^{\mathsf{c}})$ and $f_{ni} := g_{ni}/\sum_j g_{nj}$.)

Choose some $t_{ni} \in U_{ni}$. For any function $f \in C(K)$,

$$\begin{aligned}\sup_t \Bigl|f(t) - \sum_i f(t_{ni})f_{ni}(t)\Bigr| &= \sup_t \Bigl|\sum_i (f(t) - f(t_{ni}))f_{ni}(t)\Bigr| \\ &\leqslant \sup_t \sum_i |f(t) - f(t_{ni})|f_{ni}(t) \leqslant \sup\{|f(t)-f(u)| : |t-u| < 1/n\} \\ &\to 0 \end{aligned} \tag{11.5}$$

as $n \to \infty$. This shows first that $f(t) = \lim_{n\to\infty} \sum_i f(t_{ni})f_{ni}(t)$, which is $\mathcal{C}\times\mathcal{B}(K)$-measurable, showing (i).

Moreover, (11.5) implies that, for all $f \in C(K)$ and $\mu \in M(K)$,

$$\int_K f \,\mathrm{d}\mu = \lim_{n\to\infty} \sum_{i=1}^{N_i} f(t_{ni}) \int_K f_{ni} \,\mathrm{d}\mu,$$

where the right hand side evidently is $(\mathcal{C}\times\mathcal{M})$-measurable.

This proves the first claim in (ii), and the final claim follows by the Riesz representation theorem discussed before the lemma. □

Corollary 11.15. *If K is a metrizable compact, then $\mathcal{C}$ coincides with the σ-field $\mathcal{B}_w$ on $C(K)$ generated by the continuous linear functionals, and also with the Borel σ-field $\mathcal{B}$. Hence, if $X : \Omega \to C(K)$, then X is Bochner measurable $\iff$ X is weakly measurable $\iff$ X is $\mathcal{C}$-measurable.*

Proof. By Lemma 11.14, each $f \mapsto \langle x^*, f\rangle$ with $x^* \in C(K)^*$ is $\mathcal{C}$-measurable. Conversely, each point evaluation $f(t) = \langle \delta_t, f\rangle$ where $\delta_t \in C(K)^*$. This proves that $\mathcal{C}$ and $\mathcal{B}_w$ coincide.

Furthermore, $\mathcal{B}_w = \mathcal{B}$ in any separable Banach space, see Section 3, so Theorem 11.7 completes the proof. □

If K is not metrizable, i.e., when $C(K)$ is not separable, we cannot expect equivalence with Bochner measurability. Moreover, Examples 11.31 and 11.32 show that then in general $\mathcal{B}_w \neq \mathcal{C}$; hence we cannot expect equivalence between weak measurability and $\mathcal{C}$-measurability. Moreover, Example 11.31 constructs a random variable that is $\mathcal{C}$-measurable but not weakly measurable. (Recall that we only consider complete probability spaces.) Nevertheless, we shall see in Theorems 12.4 and 15.5 (using Examples 11.4 and 11.2) that there exist important cases of non-separable $C(K)$ such that $\mathcal{B}_w = \mathcal{C}$ and thus the last equivalence holds.

We can now complete the proof of Theorem 1.3 in the introduction.

Proof of Theorem 1.3. By Theorems 11.7 and 11.5, $C(K)$ is separable and has the approximation property. Hence, Theorem 1.1 applies and shows (together with its proof, or Corollary 6.14) that (1.2) $\iff$ (1.3) $\iff$ $\mathbb{E}\, X^{\check{\otimes} k} = \mathbb{E}\, Y^{\check{\otimes} k}$, where the moments exist in Bochner sense by Theorem 6.7. Finally, $\mathbb{E}\, X^{\check{\otimes} k} = \mathbb{E}\, Y^{\check{\otimes} k}$ $\iff$ (1.4) by (11.2). □

We continue with further results and next give a complete characterization of the existence of injective moments in the different senses in the metrizable case.

Theorem 11.16. *Let K be a metrizable compact space and suppose that X is a $\mathcal{C}$-measurable $C(K)$-valued random variable. Let $k \geqslant 1$.*

(i) $\mathbb{E}\, X^{\check{\otimes} k}$ *exists in Dunford sense* $\iff$ *the weak k:th moment exists* $\iff$ $\sup_{t\in K} \mathbb{E}\,|X(t)|^k < \infty$.

(ii) $\mathbb{E}\, X^{\check{\otimes} k}$ *exists in Pettis sense* $\iff$ *the family $\{|X(t)|^k : t \in K\}$ of random variables is uniformly integrable.*

(iii) $\mathbb{E}\, X^{\check{\otimes} k}$ *exists in Bochner sense* $\iff$ $\mathbb{E}\big(\sup_{t\in K} |X(t)|\big)^k < \infty$.

Proof. By Corollary 11.15, X is weakly measurable.

(i): Since $C(K)$ is separable by Theorem 11.7, Theorem 6.10(ii)(a) shows that $\mathbb{E}\, X^{\check{\otimes} k}$ exists in Dunford sense if and only if the weak k:th moment exists. Moreover, in this case $\mathbb{E}\,|x^*(X)|^k \leqslant C$ for some $C < \infty$ and all $x^* \in C(K)^*$ with $\|x^*\| \leqslant 1$. In particular, taking $x^* = \delta_t$, $\mathbb{E}\,|X(t)|^k \leqslant C$.

Conversely, if $\mathbb{E}\,|X(t)|^k \leqslant C$ for all $t \in K$, and $\mu \in C(K)^* \subseteq M(K)$ with $\|\mu\| = 1$, then, by Hölder's inequality and Fubini's theorem (using Lemma 11.14(i)),

$$\begin{aligned}\mathbb{E}\,|\langle \mu, X\rangle|^k = \mathbb{E}\left|\int_K X(t)\,\mathrm{d}\mu(t)\right|^k &\leqslant \mathbb{E}\int_K |X(t)|^k\,\mathrm{d}|\mu|(t) = \int_K \mathbb{E}\,|X(t)|^k\,\mathrm{d}|\mu|(t)\\ &\leqslant C.\end{aligned} \tag{11.6}$$

Hence Lemma 6.8 shows that the weak k:th moment exists.

(ii): Since $C(K)^{\check{\otimes}k} = C(K^k)$, its dual space is $C(K^k)^* = M_{\mathsf{r}}(K^k) \subseteq M(K^k)$. If $\mu \in M(K^k)$ with $\|\mu\| \leqslant 1$ and E is any event, then by the arithmetic-geometric inequality and Fubini (using Lemma 11.14(i))

$$\begin{aligned}
\mathbb{E}\big(\mathbf{1}_E|\langle \mu, X^{\otimes k}\rangle|\big) &= \mathbb{E}\left(\mathbf{1}_E\left|\int_{K^k} X(t_1)\cdots X(t_k)\,\mathrm{d}\mu(t_1,\ldots,t_k)\right|\right)\\
&\leqslant \mathbb{E}\left(\mathbf{1}_E\int_{K^k}\frac{1}{k}\big(|X(t_1)|^k+\cdots+|X(t_k)|^k\big)\,\mathrm{d}|\mu|(t_1,\ldots,t_k)\right)\\
&= \frac{1}{k}\int_{K^k}\sum_{i=1}^{k}\mathbb{E}\big(\mathbf{1}_E|X(t_i)|^k\big)\,\mathrm{d}|\mu|(t_1,\ldots,t_k)\\
&\leqslant \sup_{t\in K}\mathbb{E}\big(\mathbf{1}_E|X(t)|^k\big). \qquad (11.7)
\end{aligned}$$

It follows from (11.7), using (5.4), that if $\{|X(t)|^k : t \in K\}$ is uniformly integrable, then so is the family $\{\langle \mu, X^{\otimes k}\rangle : \|\mu\| \leqslant 1\}$. Since $C(K)^{\check{\otimes}k}$ is separable when $C(K)$ is, it follows by Theorem 5.8 that the moment $\mathbb{E}\,X^{\check{\otimes}k}$ exists in Pettis sense.

Conversely, if $\mathbb{E}\,X^{\check{\otimes}k}$ exists in Pettis sense, then $\{|X(t)|^k : t \in K\}$ is uniformly integrable by Theorem 6.19(i), taking $x^* = \delta_t$, $t \in K$.

(iii): Immediate by Theorem 6.7, since $C(K)$ is separable. □

Already the case $k = 1$ in Theorem 11.16 is non-trivial and gives the following characterisations of the existence of the expectation $\mathbb{E}\,X$ of a $C(K)$-valued random variable.

Corollary 11.17. *Let K be a metrizable compact space and suppose that X is a $\mathcal{C}$-measurable $C(K)$-valued random variable.*

(i) $\mathbb{E}\,X$ *exists in Dunford sense* $\iff \sup_{t\in K}\mathbb{E}\,|X(t)| < \infty$.
(ii) $\mathbb{E}\,X$ *exists in Pettis sense* $\iff$ *the family $\{X(t) : t \in K\}$ of random variables is uniformly integrable.*
(iii) $\mathbb{E}\,X$ *exists in Bochner sense* $\iff \mathbb{E}\big(\sup_{t\in K}|X(t)|\big) < \infty$.

In the Pettis and Bochner cases, $\mathbb{E}\,X$ is the continuous function $t \mapsto \mathbb{E}(X(t))$. Also when $\mathbb{E}\,X$ exists just in Dunford sense, it is given by this function (bounded by not necessarily continuous) in the sense that for any $\mu \in C(K)^* = M(K)$,

$$\langle \mu, \mathbb{E}\,X\rangle = \int_K \mathbb{E}(X(t))\,\mathrm{d}\mu(t). \qquad (11.8)$$

Proof. It remains only to verify (11.8). If X is Pettis or Bochner integrable, (11.8) follows by Theorem 11.10. In the more general Dunford case we have by Fubini, using Lemma 11.14,

$$\langle \mu, \mathbb{E}\,X\rangle = \mathbb{E}\langle \mu, X\rangle = \mathbb{E}\int_K X(t,\omega)\,\mathrm{d}\mu(t) = \int_K \mathbb{E}(X(t,\omega))\,\mathrm{d}\mu(t). \qquad (11.9)$$

□

Remark 11.18. For non-metrizable K, (11.8) is not true in general for Dunford integrable X, as seen in Example 11.30; cf. Remark 11.12. (The reason that the proof above fails in this case must be that we cannot use Fubini in (11.9), because $X(t,\omega)$ is not jointly measurable.)

We know that if the moment exists in Pettis sense, it is given by the function (11.2) which then has to be continuous. The next theorem shows that in the separable case (i.e., when K is metrizable compact), it is for the existence of an *even* injective moment in Pettis sense also sufficient that this function exists and is continuous. This is not true for odd moments, not even the first moment $\mathbb{E}\,X$, as is seen in Example 11.26 below; nor does this hold in general for non-metric K as is seen in Example 11.30.

Theorem 11.19. *Let K be a metrizable compact space and suppose that X is a $\mathcal{C}$-measurable $C(K)$-valued random variable such that $\sup_{t\in K}\mathbb{E}\,|X(t)|^2<\infty$. Suppose that $k\geqslant 2$ is even. Then the following are equivalent.*

(i) *$\mathbb{E}\,X^{\check{\otimes}k}$ exists in Pettis sense.*
(ii) *$(t_1,\dots,t_k)\mapsto\mathbb{E}\bigl(X(t_1)\cdots X(t_k)\bigr)$ is continuous on K^k.*
(iii) *$t\mapsto\mathbb{E}\,X(t)^k$ is continuous on K.*

In this case, $\mathbb{E}\,X^{\check{\otimes}k}$ is the function in $C(K^k)$ given in (ii).

Proof. (i) $\implies$ (ii): By Theorem 11.10, which also shows the final statement.

(ii) $\implies$ (iii): Trivial.

(iii) $\implies$ (i): If $t_n\to t$ in K, then $X(t_n)\to X(t)$ and thus $X(t_n)^k\to X(t)^k$. Furthermore, by (iii), and the fact that k is even, $\mathbb{E}\,|X(t_n)|^k\to\mathbb{E}\,|X(t)|^k$. Hence $\{|X(t_n)|^k\}$ is uniformly integrable and thus $|X(t_n)|^k\to|X(t)|^k$ in $L^1(\mathbb{P})$, see e.g. [34, Theorem 5.5.2]. Consequently, the map $t\mapsto|X(t)|^k$ is continuous $K\to L^1(\mathbb{P})$, and since K is compact, $\{|X(t)|^k:t\in K\}$ is a compact subset of $L^1(\mathbb{P})$, and in particular weakly compact and thus uniformly integrable [22, Theorem IV.8.11]. Thus (i) follows by Theorem 11.16(ii). □

We turn to projective moments. For the second moment, we can show that the conditions for the injective moment in Theorem 11.16 also imply the existence of the projective second moment. This uses the following result by Grothendieck [32] see e.g. [52, Theorem 5.5] or [8, Theorem V.2]. (These references also contain further related results. In particular, Grothendieck's theorem is essentially equivalent to *Grothendieck's inequality*, see [8, 52, 43].)

Let k_G denote *Grothendieck's constant*. (It is known that $\pi/2\leqslant k_G\leqslant\pi/(2\log(1+\sqrt{2}))$, but for us the value is not important.)

Theorem 11.20 (Grothendieck)**.** *If K is a compact set and α is a bounded bilinear form on $C(K)$, then there exists a Borel probability measure μ on K such that*

$$|\alpha(f,g)|\leqslant 2k_G\|\alpha\|\|f\|_{L^2(K,\mu)}\|g\|_{L^2(K,\mu)}. \tag{11.10}$$

Hence, α extends to a bounded bilinear form on $L^2(K,\mu)$ of norm $\leqslant 2k_G\|\alpha\|$. □

Remark 11.21. The standard formulation is for a bilinear form α on $C(K_1) \times C(K_2)$ for two compact sets K_1 and K_2; then there exist probability measures μ_1 and μ_2 on K_1 and K_2 such that

$$|\alpha(f,g)| \leqslant k_G \|\alpha\| \|f\|_{L^2(K_1,\mu_1)} \|g\|_{L^2(K_2,\mu_2)}. \tag{11.11}$$

We are only interested in the special case $K_1 = K_2$, and we may then replace μ_1 and μ_2 by $\mu := \frac{1}{2}(\mu_1 + \mu_2)$ to obtain (11.10). (An inspection of the proof in e.g. [32] or [8] shows that we may take k_G in the version (11.10) too provided α is symmetric, but this is not enough in general; consider for example $\alpha(f,g) := f(0)g(1)$ on $C[0,1]$. Recall that we only have to consider symmetric α for our purposes, cf. Remark 6.16.)

This leads to the following improvement of Theorem 11.16 when $k = 2$. This does not extend to $k \geqslant 3$ by Example 11.27.

Theorem 11.22. *Let K be a metrizable compact space and suppose that X is a $\mathcal{C}$-measurable $C(K)$-valued random variable.*

(i) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Dunford sense* $\iff$ $\mathbb{E}\, X^{\check{\otimes}2}$ *exists in Dunford sense* $\iff$ *the weak second moment exists* $\iff$ $\sup_{t\in K} \mathbb{E}\,|X(t)|^2 < \infty$.

(ii) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Pettis sense* $\iff$ $\mathbb{E}\, X^{\check{\otimes}2}$ *exists in Pettis sense* $\iff$ *the family $\{|X(t)|^2 : t \in K\}$ of random variables is uniformly integrable.*

(iii) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Bochner sense* $\iff$ $\mathbb{E}\, X^{\check{\otimes}2}$ *exists in Bochner sense* $\iff$ $\mathbb{E}\big(\sup_{t\in K} |X(t)|\big)^2 < \infty$.

Proof. The forward implications are immediate, using Theorems 6.2, 6.10, 6.19, 6.7 and Lemma 6.8, see the proof of Theorem 11.16. It remains to show the converses. Note that $X(t,\omega)$ is jointly measurable by Lemma 11.14.

(i): Suppose that $\mathbb{E}\,|X(t)|^2 \leqslant C$ for every $t \in K$. Let α be a bounded bilinear form on $C(K)$. By Theorem 11.20, α extends to a bounded bilinear form on $L^2(K,\mu)$ for some probability measure μ on K. Since $C(K) \subseteq L^2(K,\mu)$, we can regard X as an $L^2(K,\mu)$-valued random variable. Since $X(t,\omega)$ is jointly measurable, it follows, see Lemma 10.1, that X is Bochner measurable in $L^2(K,\mu)$. Moreover, by Fubini,

$$\mathbb{E}\,\|X\|^2_{L^2(\mu)} = \mathbb{E}\int_K |X(t,\omega)|^2 \,\mathrm{d}\mu(t) = \int_K \mathbb{E}\,|X(t)|^2 \,\mathrm{d}\mu(t) \leqslant C < \infty. \tag{11.12}$$

Hence, Theorem 6.7 shows that $\mathbb{E}\, X^{\hat{\otimes}2} \in L^2(K,\mu)^{\hat{\otimes}2}$ exists (in Bochner sense). In particular, since α extends to $L^2(K,\mu)$, $\alpha(X,X)$ is measurable by Lemma 6.3 and

$$\mathbb{E}\,\alpha(X,X) = \mathbb{E}\langle \alpha, X^{\hat{\otimes}2}\rangle = \langle \alpha, \mathbb{E}\, X^{\hat{\otimes}2}\rangle$$

exists. Since α is an arbitrary bounded bilinear form on $C(K)$, Theorem 6.15 now shows that $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense in $C(K)^{\hat{\otimes}2}$ too.

(ii): Assume that the family $\{|X(t)|^2 : t \in K\}$ is uniformly integrable. Let α be a bounded bilinear form on $C(K)$ with $\|\alpha\| \leqslant 1$ and let as above μ be

as in Theorem 11.20. If $E \in \mathcal{F}$ is any measurable set, then by (11.10) and Fubini,

$$\begin{aligned}\mathbb{E}\,|\mathbf{1}_E \cdot \alpha(X,X)| &\leqslant \mathbb{E}\Big(\mathbf{1}_E 2k_G \int_K |X(t,\omega)|^2 \,\mathrm{d}\mu(t)\Big) \\ &= 2k_G \int_K \mathbb{E}\big(\mathbf{1}_E |X(t,\omega)|^2\big)\,\mathrm{d}\mu(t) \leqslant 2k_G \sup_{t\in K} \mathbb{E}\big(\mathbf{1}_E |X(t)|^2\big)\end{aligned}$$

It follows, using (5.4), that the family $\big\{\alpha(X,X) : \alpha \in L(C(K)^2;\mathbb{R}),\ \|\alpha\| \leqslant 1\big\}$ is uniformly integrable. Moreover, $X^{\hat{\otimes}2}$ is trivially weakly a.s. separably valued since $C(K)^{\hat{\otimes}2}$ is separable. Hence Theorem 5.8 shows, using Theorem 4.6, that $\mathbb{E}\,X^{\hat{\otimes}2}$ exists in Pettis sense.

(iii): This is (again) a special case of Theorem 6.7, included here for completeness. □

As an immediate corollary, we can weaken the integrability condition in Theorem 1.3 in the case $k = 2$ to $\{|X(t)|^2 : t \in K\}$ and $\{|Y(t)|^2 : t \in K\}$ being uniformly integrable, since it is enough for the proofs above of Theorems 1.1 and 1.3 that the moments $\mathbb{E}\,X^{\hat{\otimes}2}$ and $\mathbb{E}\,Y^{\hat{\otimes}2}$ exist in Pettis sense. By the methods in the proof of Theorem 11.19, we can weaken the condition further; the following theorem shows that in this case, it suffices that the moments exist in Dunford sense.

Theorem 11.23. *Let K be a metrizable compact space and suppose that X and Y are two $\mathcal{C}$-measurable $C(K)$-valued random variables such that $\sup_{t\in K}|X(t)|^2 < \infty$ and $\sup_{t\in K}|Y(t)|^2 < \infty$. Then, for $k = 2$, the following are equivalent.*

- (i) (1.2), *i.e.,* $\mathbb{E}\,\alpha(X,X) = \mathbb{E}\,\alpha(Y,Y)$ *for every bounded bilinear form α on $C(K)$.*
- (ii) (1.3), *i.e.,* $\mathbb{E}\big(x_1^*(X)x_2^*(X)\big) = \mathbb{E}\big(x_1^*(Y)x_2^*(Y)\big)$, *for any $x_1^*, x_2^* \in C(K)^*$.*
- (iii) (1.4), *i.e.,* $\mathbb{E}\big(X(t_1)X(t_2)\big) = \mathbb{E}\big(Y(t_1)Y(t_2)\big)$ *for any $t_1, t_2 \in K$.*
- (iv) $\mathbb{E}\,X^{\hat{\otimes}2} = \mathbb{E}\,Y^{\hat{\otimes}2}$ *in $C(K)^{\hat{\otimes}2}$, with the moments existing in Dunford sense.*
- (v) $\mathbb{E}\,X^{\check{\otimes}2} = \mathbb{E}\,Y^{\check{\otimes}2}$ *in $C(K)^{\check{\otimes}2}$, with the moments existing in Dunford sense.*

Proof. The implications (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) are trivial, and the equivalences (i) $\Longleftrightarrow$ (iv) and (ii) $\Longleftrightarrow$ (v) are Corollaries 6.17 and 6.12(a). It remains to show that (iii) $\Longrightarrow$ (i).

Thus, let $\alpha \in L(C(K)^2;\mathbb{R}) = \big(C(K)^{\hat{\otimes}2}\big)^*$ As in the proof of Theorem 11.22, there exists by Theorem 11.20 a probability measure μ on K such that α extends to $L^2(K,\mu)$ and $\mathbb{E}\,X^{\hat{\otimes}2}$ exists in $L^2(K,\mu)^{\hat{\otimes}2}$ in Bochner sense; similarly $\mathbb{E}\,Y^{\hat{\otimes}2}$ exists in $L^2(K,\mu)^{\hat{\otimes}2}$ in Bochner sense.

As a consequence, the injective moments $\mathbb{E}\, X^{\check{\otimes}2}$ and $\mathbb{E}\, Y^{\check{\otimes}2}$ exist in $L^2(K,\mu)^{\check{\otimes}2}$ in Bochner sense. Let $g, h \in L^2(K,\mu)^* = L^2(K,\mu)$. Then, by Fubini's theorem, with joint measurability by Lemma 11.14,

$$\begin{aligned}\mathbb{E}\big(\langle X, g\rangle\langle X, h\rangle\big) &= \mathbb{E}\iint X(t,\omega)g(t)X(u,\omega)h(u)\,\mathrm{d}\mu(t)\,\mathrm{d}\mu(u)\\ &= \iint \mathbb{E}\big(X(t)X(u)\big)g(t)h(u)\,\mathrm{d}\mu(t)\,\mathrm{d}\mu(u)\end{aligned}$$

The same applies to Y. Consequently, if (iii) holds, then $\mathbb{E}\big(\langle X, g\rangle\langle X, h\rangle\big) = \mathbb{E}\big(\langle Y, g\rangle\langle Y, h\rangle\big)$, which is (1.3) for $L^2(K,\mu)$ (with $k = 2$). Hence Corollary 6.14 yields $\mathbb{E}\, X^{\check{\otimes}2} = \mathbb{E}\, Y^{\check{\otimes}2}$ in $L^2(K,\mu)^{\check{\otimes}2}$. Furthermore, $L^2(K,\mu)$ has the approximation property by Theorem 9.1, and thus $\mathbb{E}\, X^{\hat{\otimes}2} = \mathbb{E}\, Y^{\hat{\otimes}2}$ in $L^2(K,\mu)^{\hat{\otimes}2}$ by Theorem 8.3. Consequently,

$$\mathbb{E}\,\alpha(X,X) = \big\langle \alpha, \mathbb{E}\, X^{\hat{\otimes}2}\big\rangle = \big\langle \alpha, \mathbb{E}\, Y^{\hat{\otimes}2}\big\rangle = \mathbb{E}\,\alpha(Y,Y).$$

Since α is an arbitrary bounded bilinear form on $C(K)$, this completes the proof. □

Remark 11.24. Theorem 11.23 shows that, when K is metrizable, the moment $\mathbb{E}\, X^{\hat{\otimes}2}$ (or $\mathbb{E}\, X^{\check{\otimes}2}$) is determined by the moment function $\mathbb{E}(X(t)X(u))$ as soon as the moment exists in Dunford sense. The corresponding result for $k = 1$ is in Corollary 11.17.

Example 11.30 yields an example of a non-metrizable K such that (iii) does not imply (ii).

We leave it as an open problem whether Theorem 11.23 extends to $k \geqslant 3$. In particular, for metrizable K and $k \geqslant 3$, if $\mathbb{E}\, X^{\check{\otimes}k}$ or $\mathbb{E}\, X^{\hat{\otimes}k}$ exists in Dunford sense, does the moment function (11.2) determine this moment uniquely?

For projective moments of order $k \geqslant 3$, we do not know any special results for Pettis or Dunford integrability, but we have as always a simple result for Bochner integrability. (Recall that this implies Pettis and Dunford integrability, so we have a sufficient condition for them too.)

Theorem 11.25. *Let K be a metrizable compact space and suppose that X is a $\mathcal{C}$-measurable $C(K)$-valued random variable. Then $\mathbb{E}\, X^{\hat{\otimes}k}$ exists in Bochner sense $\iff$ $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Bochner sense $\iff$ $\mathbb{E}\big(\sup_{t\in K}|X(t)|\big)^k < \infty$.*

Proof. Theorem 6.7. □

We end by a few counterexamples, partly taken or adapted from [24] and [62] where further related examples are given. See also Examples 5.14, 5.18 and 5.19, which by Examples 11.1–11.4 can be seen as examples in some $C(K)$.

Example 11.26. Let X be a $C(K)$-valued random variable such that $\mathbb{E}\,X$ exists in Dunford sense but not in Pettis sense, and let $Y := \eta X$, where $\eta = \pm 1$ with the sign uniformly random and independent of X. Then Y is Dunford integrable and $\mathbb{E}\,Y(t) = \mathbb{E}(\eta X(t)) = 0$ for every $t \in K$, so $\mathbb{E}\,Y(\cdot) \in C(K)$ but Y is not Pettis integrable (since otherwise $X = \eta Y$ would be too). We may for example take X as in Example 5.14 ($X \in c_0 \subset c = C(\mathbb{N}^*)$, so $K = \mathbb{N}^*$ is metric) or as in Example 5.19 (X is bounded). (We cannot take both K metric and X bounded by Theorem 11.16.)

Example 11.27 (cf. [8, Section V.5])**.** Let $\mathbb{T} := [0, 2\pi]$ (perhaps regarded as the unit circle) and $\mathbb{N}^* := \mathbb{N} \cup \{\infty\}$ (the usual one-point compactification as in Example 11.4), and take $K := \mathbb{T} \cup \mathbb{N}^*$, where we regard $\mathbb{T}$ and $\mathbb{N}^*$ as disjoint. If $f \in C(K)$, define $\hat f(n) := \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-\mathrm{i}nt}\,\mathrm{d}t$, i.e., the Fourier coefficients of $f|_{\mathbb{T}}$.

Define the trilinear form α on $C(K)$ by

$$\alpha(f,g,h) := \operatorname{Re}\sum_{n=1}^{\infty} \hat f(n)\hat g(n)h(n). \tag{11.13}$$

By Hölder's inequality and Parseval's identity,

$$\sum_{n=1}^{\infty}\bigl|\hat f(n)\hat g(n)h(n)\bigr| \leqslant \|h\|\biggl(\int_{\mathbb{T}}|f(t)|^2\frac{\mathrm{d}t}{2\pi}\biggr)^{1/2}\biggl(\int_{\mathbb{T}}|g(t)|^2\frac{\mathrm{d}t}{2\pi}\biggr)^{1/2} \leqslant \|f\|\|g\|\|h\|;$$

thus the sum in (11.13) converges and α is well-defined and bounded.

Let N be an $\mathbb{N}$-valued random variable with $\mathbb{P}(N = n) = p_n$, let a_n be some positive numbers and define the random variable $X \in C(K)$ by $X(t) := \cos(Nt)$, $t \in \mathbb{T}$, and $X(n) := a_n e_N(n)$, $n \in \mathbb{N}^*$. Then $\hat X(N) = \frac12$ and

$$\alpha(X,X,X) = \hat X(N)^2 a_N = \tfrac14 a_N.$$

Hence,

$$\mathbb{E}\,|\alpha(X,X,X)| = \tfrac14\,\mathbb{E}\,a_N = \tfrac14\sum_{n=1}^{\infty} p_n a_n. \tag{11.14}$$

Furthermore, $|X(t)| \leqslant 1$ for $t \in \mathbb{T}$ and $\mathbb{E}\,|X(n)|^3 = p_n a_n^3$ for $n \in \mathbb{N}$ (and 0 for $n = \infty$).

Choose, e.g., $p_n := n^{-4/3}$ (for n large) and $a_n := n^{1/3}$. Then $\mathbb{E}\,|X(n)|^3 \to 0$ as $n \to \infty$, and thus $\{|X(x)|^3 : x \in K\}$ is uniformly integrable. Hence $\mathbb{E}\,X^{\check\otimes 3}$ exists in Pettis sense by Theorem 11.16. On the other hand, $\mathbb{E}\,|\alpha(X,X,X)| = \infty$ by (11.14), and thus $\mathbb{E}\,X^{\hat\otimes 3}$ does not exist (even in Dunford sense). This shows that Theorem 11.22(i)(ii) do not hold for $k = 3$.

The example may be modified for any given $k \geqslant 3$ by taking

$$\alpha(f_1,\dots,f_k) := \operatorname{Re}\sum_{n=1}^{\infty} \hat f_1(n)\hat f_2(n) f_3(n)\cdots f_k(n), \tag{11.15}$$

and $p_n = n^{-(k+1)/3}$.

Example 11.28 (cf. [23, Example (2)] and [24, Example 5.5])**.** Let ω_1 be the first uncountable ordinal and let $K = [0, \omega_1]$, the set of all ordinals $\leqslant \omega_1$ with the order topology. It is easily seen that K is compact, Note that $K\backslash\{\omega_1\} = [0, \omega_1)$ is the (uncountable) set of all countable ordinals.

It is not difficult to see that every $f \in C[0, \omega_1]$ is constant on $[\alpha, \omega_1]$ for some $\alpha < \omega_1$. Hence, the Baire σ-field Ba on $[0, \omega_1]$ is the σ-field consisting of all subsets A such that either A or its complement A^{c} is countable.

Furthermore, every bounded increasing function $f : [0, \omega_1] \to \mathbb{R}$ is constant on $[\alpha, \omega_1)$ for some $\alpha < \omega_1$. It follows (by considering the positive and negative parts) that every regular signed Borel (or Baire) measure on $[0, \omega_1]$ is supported on some set $[0, \alpha] \cup \{\omega_1\}$; hence every regular signed Borel measure has countable support and is thus discrete. Consequently, $C[0, \omega_1]^* = \ell^1[0, \omega_1]$.

We define a probability measure $\mathbb{P}$ on $([0, \omega_1], \mathsf{Ba})$ by setting $\mathbb{P}(A) = 0$ if A is countable and $\mathbb{P}(A) = 1$ if A^{c} is countable. The mapping $X : \alpha \mapsto X_\alpha := \mathbf{1}_{[\alpha,\omega_1]}$ maps $(\Omega, \mathsf{Ba}, \mathbb{P}) \to C[0, \omega_1]$. If $\beta < \omega_1$, then $X_\alpha(\beta) = \mathbf{1}\{\alpha \leqslant \beta\} = 0$ for a.e. α, while $X_\alpha(\omega_1) = 1$ for all α. In other words, $X(\beta) = 0$ a.s. if $\beta < \omega_1$ but $X(\omega_1) = 1$.

We have seen that any $x^* \in C[0, \omega_1]^*$ is given by

$$x^*(f) = \sum_\alpha \xi(\alpha) f(\alpha) \tag{11.16}$$

for some $\xi \in \ell^1[0, \omega_1]$. (Note that the sum really is countable.) It follows that in this case,

$$x^*(X) = \xi(\omega_1) \qquad \text{a.s.} \tag{11.17}$$

Consequently, X is weakly measurable, and since X further is bounded, the Dunford integral $\mathbb{E}\, X$ exists. We have, by (11.17),

$$\langle \mathbb{E}\, X, x^* \rangle = \mathbb{E}\langle x^*, X \rangle = \xi(\omega_1), \tag{11.18}$$

when x^* is given by (11.16). Consequently, $\mathbb{E}\, X$ is given by the function $\mathbf{1}_{\{\omega_1\}}$ that is 1 at ω_1 and 0 on $[0, \omega_1)$. This function is not continuous, and thus $\mathbb{E}\, X \notin C[0, \omega_1]$, which shows that X is not Pettis integrable.

By Theorem 5.8, X is not weakly a.s. separably valued. (This is also easily seen directly. If M is any separable subspace of $C[0, \omega_1]$, then there exists an $\alpha < \omega_1$ such that every function in M is constant on $[\alpha, \omega_1]$. Hence, $X \notin M$ a.s.)

Note that Example 11.28 gives a uniformly bounded continuous random function $X(t)$ such that $\mathbb{E}\, X(t)$ is not continuous. This cannot happen on a metric space K, since dominated convergence shows that $\mathbb{E}\, X(t)$ always is sequentially continuous. (Indeed, if X is a uniformly bounded random variable in $C(K)$ with K compact metric, then Theorem 11.16 applies and shows that all injective moments exist in both Bochner and Pettis sense.)

Example 11.29 ([29] and [62])**.** Recall that $\ell^\infty = C_b(\mathbb{N}) \cong C(\beta\mathbb{N})$ where $\beta\mathbb{N}$ is the Stone–Čech compactification of $\mathbb{N}$, see Example 11.1. Furthermore,

$\beta\mathbb{N}$ can be regarded as the subset of the unit ball of $(\ell^\infty)^*$ consisting of all multiplicative linear functionals on ℓ^∞. (This holds also as topological spaces, with the weak* topology on $(\ell^\infty)^*$.) A point $n \in \mathbb{N}$ then is identified with the multiplicative linear functional $\delta_n : f \mapsto f(n)$.

Let $X^{(1)}$ and $X^{(2)}$ be two independent copies of the random variable $X \in \ell^\infty = C(\beta\mathbb{N})$ constructed in Example 5.19, and let $Y = X^{(1)} - X^{(2)}$. We follow Fremlin and Talagrand [29] and Talagrand [62, Theorem 4-2-5] to show that Y is Pettis integrable. (Recall from Example 5.19 that X is not.)

The coordinates Y_n are i.i.d. random variables, each with the distribution of $\eta_1 - \eta_2$ with independent $\eta_1, \eta_2 \sim \mathrm{Be}(1/2)$; hence Y_n has the centred binomial distribution $\mathrm{Bin}(2, 1/2) - 1$. It follows that $\sqrt{2}Y_n$ is an orthonormal sequence in $L^2(\mathbb{P})$. Consequently, by Bessel's inequality, if $\xi \in L^\infty(\mathbb{P}) \subset L^2(\mathbb{P})$, then the sequence $y_\xi := (\mathbb{E}(\xi Y_n))_1^\infty \in \ell^2 \subset c_0$.

If $x^* = (a_n)_1^\infty \in \ell^1$, then by Fubini,

$$\langle x^*, y_\xi \rangle = \sum_{n=1}^\infty a_n \, \mathbb{E}(\xi Y_n) = \mathbb{E} \sum_{n=1}^\infty a_n \xi Y_n = \mathbb{E}(\xi \langle x^*, Y \rangle). \tag{11.19}$$

If $x^* \in c_0^\perp$, then $x^*(X^{(1)})$ and $x^*(X^{(2)})$ are by Example 5.19 equal to some constant a.s., and thus $x^*(Y) = x^*(X^{(1)}) - x^*(X^{(2)}) = 0$ a.s. Furthermore, $\langle x^*, y_\xi \rangle = 0$ since $y_\xi \in c_0$, and thus (11.19) holds in this case too. Hence, (11.19) holds for all $x^* \in (\ell^\infty)^*$, and every $\xi \in L^\infty(\mathbb{P})$, which shows that Y is Pettis integrable and $\mathbb{E}(\xi Y) = y_\xi$.

Consider now the injective second moment $\mathbb{E}\, Y^{\check{\otimes} 2}$. Assume that this exists as a Pettis integral; it then belongs to $C(\beta\mathbb{N}) \check{\otimes} C(\beta\mathbb{N}) = C((\beta\mathbb{N})^2)$. Let us write $Q := \mathbb{E}\, Y^{\check{\otimes} 2} \in C((\beta\mathbb{N})^2)$. If $x^* \in \beta\mathbb{N}\backslash\mathbb{N}$, then $x^* \in c_0^\perp$, and thus $x^*(Y) = 0$ a.s.; consequently,

$$Q(x^*, x^*) = \langle x^* \otimes x^*, \mathbb{E}\, Y^{\check{\otimes} 2} \rangle = \mathbb{E} \langle x^* \otimes x^*, Y^{\otimes 2} \rangle = \mathbb{E} \langle x^*, Y \rangle^2 = 0. \tag{11.20}$$

On the other hand, if $n \in \mathbb{N}$, then, similarly,

$$Q(n, n) = \langle \delta_n \otimes \delta_n, \mathbb{E}\, Y^{\check{\otimes} 2} \rangle = \mathbb{E} \langle \delta_n \otimes \delta_n, Y^{\otimes 2} \rangle = \mathbb{E} \langle \delta_n, Y \rangle^2 = \tfrac{1}{2}, \tag{11.21}$$

since $\langle \delta_n, Y \rangle = Y_n \sim \mathrm{Bin}(2, \frac{1}{2}) - 1$. However, $\mathbb{N}$ is dense in $\beta\mathbb{N}$, so (11.21) implies by continuity that $Q(x^*, x^*) = \frac{1}{2}$ for all $x^* \in \beta\mathbb{N}$, which contradicts (11.20). Consequently, the second moment $\mathbb{E}\, Y^{\check{\otimes} 2}$ does *not* exist as a Pettis integral. (We do not know whether it exists as a Dunford integral.)

Example 11.30. Let $B := \ell^\infty / c_0$. We identify ℓ^∞ and $C(\beta\mathbb{N})$, see Example 11.1; then $c_0 = \{f \in C(\beta\mathbb{N}) : f(x) = 0 \text{ when } x \in \beta\mathbb{N}\backslash\mathbb{N}\}$, and it follows by the Tietze–Urysohn extension theorem [27, Theorem 2.1.8], since $\mathbb{N}$ is open in $\beta\mathbb{N}$, that we can identify $B = \ell^\infty / c_0 = C(\beta\mathbb{N}\backslash\mathbb{N})$. (This can also be seen from Example 11.3, since $\beta\mathbb{N}\backslash\mathbb{N}$ is the maximal ideal space of the complex version of ℓ^∞ / c_0.)

Let $\pi : \ell^\infty \to \ell^\infty / c_0$ be the quotient mapping. (Thus π is identified with the restriction mapping $C(\beta\mathbb{N}) \to C(\beta\mathbb{N}\backslash\mathbb{N})$.) Let X be the weakly

measurable ℓ^∞-valued random variable constructed in Example 5.19, and let $\hat{X} = \pi(X) \in \ell^\infty/c_0 = C(\beta\mathbb{N}\backslash\mathbb{N})$. Finally, let $Z := 1 - \hat{X}$. (Thus $0 \leqslant Z \leqslant 1$.)

Then $\hat{X}$ and Z are bounded and weakly measurable $C(K)$-valued random variable, with $K = \beta\mathbb{N}\backslash\mathbb{N}$. By Example 5.19, if $x^* \in K = \beta\mathbb{N}\backslash\mathbb{N}$, then

$$\hat{X}(x^*) = X(x^*) = \langle x^*, X\rangle = 1 \qquad \text{a.s.} \tag{11.22}$$

Hence,

$$Z(x^*) = 0 \qquad \text{a.s.} \tag{11.23}$$

for every $x^* \in K$ and

$$\mathbb{E}\big(Z(x_1^*)\cdots Z(x_k^*)\big) = 0 \tag{11.24}$$

for any $x_1^*, \dots, x_k^* \in K = \beta\mathbb{N}\backslash\mathbb{N}$, so the function defined in (11.2) exists and is continuous on K^k (in fact, constant 0).

However, let $x_n^* \in (\ell^\infty)^*$ be defined by $\langle x_n^*, (a_i)_i\rangle := \frac{1}{n}\sum_{i=1}^n a_i$, and let $T : \ell^\infty \to \ell^\infty$ be the linear map $x \to (x_n^*(x))_n$. Choose any $x_0^* \in \beta\mathbb{N}\backslash\mathbb{N}$ and define $\tilde{x}^* \in (\ell^\infty)^*$ as $T^*x_0^*$. i.e., by

$$\langle \tilde{x}^*, x\rangle := \langle x_0^*, Tx\rangle = \langle x_0^*, (x_n^*(x))_n\rangle, \qquad x \in \ell^\infty.$$

If $x \in c_0$, then $x_n^*(x) \to 0$ as $n \to \infty$, and thus $Tx = (x_n^*(x))_n \in c_0$, so $\langle \tilde{x}^*, x\rangle = \langle x_0^*, Tx\rangle = 0$; hence, $\tilde{x}^* \perp c_0$, and $\tilde{x}^* \in (\ell^\infty/c_0)^*$. Furthermore, by the law of large numbers, $\langle x_n^*, X\rangle = \frac{1}{n}\sum_{i=1}^n X_i \to \frac{1}{2}$ a.s., and thus

$$\langle \tilde{x}^*, \hat{X}\rangle = \langle \tilde{x}^*, X\rangle = \tfrac{1}{2} \qquad \text{a.s.} \tag{11.25}$$

and

$$\langle \tilde{x}^*, Z\rangle = \langle \tilde{x}^*, 1 - \hat{X}\rangle = 1 - \langle \tilde{x}^*, \hat{X}\rangle = \tfrac{1}{2} \qquad \text{a.s.} \tag{11.26}$$

For $k = 1$, $\mathbb{E}\,Z$ exists in Dunford sense, since Z is bounded and weakly measurable; (11.26) implies that $\langle \tilde{x}^*, \mathbb{E}\,Z\rangle = \mathbb{E}\langle \tilde{x}^*, Z\rangle = \frac{1}{2}$ and thus $\mathbb{E}\,Z \neq 0$, although (11.23) shows that $\langle \mathbb{E}\,Z, \delta_t\rangle = \mathbb{E}(Z(t)) = 0$ for every $t \in K$. In particular, it follows that $\mathbb{E}\,Z \in C(K)^{**}\backslash C(K)$, and thus Z is not Pettis integrable. (Cf. Example 5.19 which shows that $\mathbb{E}\,X$ does not exist in Pettis sense by essentially the same argument.) We see also that (11.8) fails for Z and $\mu = \tilde{x}^*$; hence, as said in Remark 11.18, $Z(t,\omega)$ is not jointly measurable.

For $k \geqslant 2$, we do not know whether $\mathbb{E}\,Z^{\check{\otimes}k}$ or $\mathbb{E}\,Z^{\hat{\otimes}k}$ exists in Dunford sense, but they do not exists in Pettis sense, by an extension of the argument for $\mathbb{E}\,Z$. Indeed, if $\mathbb{E}\,Z^{\check{\otimes}k}$ exists in Pettis (or just Dunford) sense, then by (11.25), with x^* as above,

$$\big\langle (\tilde{x}^*)^{\otimes k}, \mathbb{E}\,Z^{\check{\otimes}k}\big\rangle = \mathbb{E}\langle x^*, Z\rangle^k = \big(\tfrac{1}{2}\big)^k, \tag{11.27}$$

so $\mathbb{E}\,Z^{\check{\otimes}k} \neq 0$. On the other hand, by (11.24),

$$\big\langle \delta_{t_1}\otimes\cdots\otimes\delta_{t_k}, \mathbb{E}\,Z^{\check{\otimes}k}\big\rangle = \mathbb{E}\big(Z(t_1)\cdots Z(t_k)\big) = 0 \tag{11.28}$$

for all $t_1, \dots, t_k \in K$. Suppose that $\mathbb{E}\,Z^{\check{\otimes}k}$ exists in Pettis sense. Then, by Theorem 11.10, $\mathbb{E}\,Z^{\check{\otimes}k}$ is the function (11.2) on K^k, i.e., by (11.28),

$\mathbb{E}\, Z^{\check{\otimes} k} = 0$, a contradiction. Hence, $\mathbb{E}\, Z^{\check{\otimes} k}$ does not exist in Pettis sense. This also shows that Theorem 11.19 does not hold for $K = \beta\mathbb{N}\backslash\mathbb{N}$.

Example 11.31. Let $Z \in \ell^\infty/c_0$ and $\tilde{x}^* \in (\ell^\infty/c_0)^*$ be as in Example 11.30. Let $\xi \sim \mathrm{Be}(1/2)$ be independent of Z, and define $Z_1 := \xi Z \in \ell^\infty/c_0 = C(\beta\mathbb{N}\backslash\mathbb{N})$.

By (11.23), $Z_1(x^*) = 0$ a.s. for every $x^* \in K = \beta\mathbb{N}\backslash\mathbb{N}$. On the other hand, by (11.26),

$$\langle \tilde{x}^*, Z_1\rangle = \xi\langle \tilde{x}^*, Z\rangle = \tfrac{1}{2}\xi \qquad \text{a.s.} \tag{11.29}$$

Define the measure μ on the σ-field $\mathcal{C}$ as the distribution of Z_1, and regard the random variable Z_1 as defined by the identity map $C(K) \to C(K)$ on the probability space $(C(K), \mathcal{C}^*, \mu)$, where $\mathcal{C}^*$ is the completion of $\mathcal{C}$. (Recall that we want our probability space to be complete.) This version of Z_1 is $\mathcal{C}$-measurable. If $\varphi(Z_1)$ is any measurable functional, then φ is a $\mathcal{C}^*$-measurable function on $C(K)$, and thus φ is μ-a.e. equal to a $\mathcal{C}$-measurable function ψ on $C(K)$. By the definition of $\mathcal{C}$, this implies that $\psi(f) = \Psi\bigl(f(t_1), f(t_2), \dots\bigr)$ for some function F on $\mathbb{R}^\infty$ and some sequence of points $t_i \in K$. By (11.23), each $Z_1(t_i)$ is a.s. constant, and thus $\psi(Z_1)$ is a.s. constant; since $\varphi(Z_1) = \psi(Z_1)$ a.s., also $\varphi(Z_1)$ is a.s. constant. Since $\tilde{x}^*(Z_1)$ is not a.s. constant by (11.29), $\tilde{x}^*(Z_1)$ is not measurable; hence Z_1 is $\mathcal{C}$-measurable but not weakly measurable. It follows that $\mathcal{C} \neq \mathcal{B}_w$.

Example 11.32. For another (simpler) example with $\mathcal{C} \neq \mathcal{B}_w$, let $K := \{0,1\}^{\mathfrak{c}_1}$, where $\mathfrak{c}_1 = 2^{\mathfrak{c}}$ (or any cardinal number $> \mathfrak{c}$), and let as in Example 5.19 μ be the product measure $\mu := \bigl(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1\bigr)^{\mathfrak{c}_1}$.

Suppose that the linear functional $f \mapsto \chi(F) := \int_K f\,\mathrm{d}\mu$ is $\mathcal{C}$-measurable. Then there exist points $t_i \in K$, $i = 1, 2, \dots$, and a (measurable) function $\Phi : \mathbb{R}^\infty \to \mathbb{R}$ such that

$$\int_K f\,\mathrm{d}\mu = \Phi(f(t_1), f(t_2), \dots), \qquad f \in C(K). \tag{11.30}$$

Each $t \in K$ is a function $\mathfrak{c}_1 \to \{0,1\}$ which we denote by $\alpha \mapsto t(\alpha)$. Define an equivalence relation on $\mathfrak{c}_1$ by $\beta \equiv \gamma \iff t_i(\beta) = t_i(\gamma)$ for all i. The number of equivalence classes is $2^{\aleph_0} = \mathfrak{c} < \mathfrak{c}_1$, and thus there exists $\beta, \gamma \in \mathfrak{c}_1$ with $\beta \neq \gamma$ but $\beta \equiv \gamma$.

Consider the normalized coordinate functions $f_\alpha(t) := 2t(\alpha) - 1 : K \to \{-1, 1\}$, $\alpha \in \mathfrak{c}_1$. Since $\beta \equiv \gamma$, $f_\beta(t_i) = f_\gamma(t_i)$ for each i, and thus $f_\beta(t_i) f_\gamma(t_i) = 1$. Consequently, if $g := f_\alpha f_\gamma$ and $h := 1$, then $g, h \in C(K)$ and $g(t_i) = h(t_i)$ for each i, so by (11.30)

$$\int_K g\,\mathrm{d}\mu = \Phi(g(t_1), g(t_2), \dots) = \Phi(h(t_1), h(t_2), \dots) = \int_K h\,\mathrm{d}\mu,$$

which is a contradiction since $\int g\,\mathrm{d}\mu = 0$ and $\int h\,\mathrm{d}\mu = 1$.

This contradiction shows that the continuous linear functional χ is not $\mathcal{C}$-measurable, and thus $\mathcal{B}_w \supsetneq \mathcal{C}$.

12. $c_0(S)$

In this section we consider $B = c_0(S)$, for an arbitrary set S. As discussed in Example 11.4, we can regard $c_0(S)$ as a complemented subspace of codimension 1 in $C(S^*) = C(S \cup \{\infty\})$: $c_0(S) = \{f \in C(S^*) : f(\infty) = 0\}$. (The results below could easily be formulated for $C(S^*)$ instead, but we leave that to the reader.)

Note that $c_0(S)$ is separable (and S^* metrizable) if and only if S is countable. (The discrete space S is always metrizable, but that is not enough.) The case when S is countable is thus covered by the results (for $C(S^*)$) in Section 11. We shall see that these results extend to arbitrary S because of the special simple structure of $c_0(S)$. This illustrates that some non-separable Banach spaces can be handled without problems, and it is a background to Sections 15–16 where we (by technically more complicated arguments) obtain similar results for $D[0,1]$, which is more important for applications.

Theorem 12.1. *The injective tensor product $c_0(S)^{\check{\otimes} k} = c_0(S^k)$ (isometrically). Moreover, $c_0(S)$ has the approximation property, and thus $c_0(S)^{\hat{\otimes} k} \subseteq c_0(S)^{\check{\otimes} k} = c_0(S^k)$ (as a vector space).*

Proof. An easy consequence of Theorems 11.6 and 11.5, applied to $C(S^*)$. □

The dual space $c_0(S)^* = \ell^1(S)$. Note that every element in $\ell^1(S)$ has countable support. Thus every $x^* \in c_0(S)^*$ depends only on countably many coordinates. This extends to multilinear forms as follows. For a subset $A \subseteq S$, let P_A be the projection in $c_0(S)$ defined by

$$P_A f(x) := \mathbf{1}\{x \in A\} f(x). \tag{12.1}$$

Lemma 12.2. *If S is any set and α is a bounded k-linear form on $c_0(S)$, then there exist a countable subset $A \subseteq S$ such that*

$$\alpha(f_1, \dots, f_k) = \alpha(P_A f_1, \dots, P_A f_k). \tag{12.2}$$

Proof. Write, for convenience,

$$a(s_1, \dots, s_k) := \alpha(e_{s_1}, \dots, e_{s_k}), \qquad s_1, \dots, s_k \in S. \tag{12.3}$$

Fix a finite set $F \subseteq S$ and let $X \in c_0(S)$ be random with $X(s)$, $s \in F$, i.i.d. with $\mathbb{P}(X(s) = +1) = \mathbb{P}(X(s) = -1) = \frac{1}{2}$, while $X(s) = 0$ for $s \notin F$.

Let $X_1, \dots, X_k$ be independent copies of X. Then

$$\alpha(X_1, \dots, X_k) = \sum_{s_1, \dots, s_k \in F} a(s_1, \dots, s_k) X_1(s_1) \cdots X_k(s_k)$$

and thus, since different terms are orthogonal,

$$\mathbb{E}\,|\alpha(X_1, \dots, X_k)|^2 = \sum_{s_1, \dots, s_k \in F} |a(s_1, \dots, s_k)|^2.$$

Hence,

$$\sum_{s_1,\dots,s_k\in F} |a(s_1,\dots,s_k)|^2 \leqslant \|\alpha\|^2.$$

Since this holds for every finite F,

$$\sum_{s_1,\dots,s_k\in S} |a(s_1,\dots,s_k)|^2 \leqslant \|\alpha\|^2. \tag{12.4}$$

In particular, only a countable number of $a(s_1,\dots,s_k)$ are non-zero. Hence there exists a countable subset A of S such that $a(s_1,\dots,s_k) = 0$ unless $s_1,\dots,s_k \in A$. Then (12.2) holds for every $f_1,\dots,f_k$ with finite supports, and the general case follows by continuity. □

Remark 12.3. Bohnenblust and Hille [9] proved the stronger result

$$\sum_{s_1,\dots,s_k\in S} |a(s_1,\dots,s_k)|^{2k/(k+1)} < \infty,$$

where the case $k = 1$ is just $c_0(S)^* = \ell^1(S)$ and $k = 2$ had been proved earlier by Littlewood [44]; see also [8].

The integral forms are, by definition, the elements of the dual of $c_0(S)^{\check{\otimes} k}$; by Theorem 12.1 this equals $c_0(S^k)^* = \ell^1(S^k)$. Consequently, every integral k-linear form on $c_0(S)$ is nuclear.

We let, as in Section 11, $\mathcal{C}$ be the σ-field generated by the point evaluations. Thus, a $c_0(S)$-valued random variable X is $\mathcal{C}$-measurable if and only if $X(s)$ is measurable for every $s \in S$.

Theorem 12.4. *The σ-fields $\mathcal{C}$ and $\mathcal{B}_w$ on $c_0(S)$ coincide, for any S.*
Moreover, the following hold for any $c_0(S)$-valued random variable X:

(i) *X is weakly measurable if and only if X is $\mathcal{C}$-measurable.*
(ii) *If X is $\mathcal{C}$-measurable, then there exists a countable subset $S_0 \subseteq S$ such that for every $s \notin S_0$, $X(s) = 0$ a.s.*
(iii) *If X is $\mathcal{C}$-measurable, then X is weakly a.s. separably valued. Moreover, then $X^{\otimes k}$ is weakly a.s. separably valued in $c_0(S)^{\hat{\otimes} k}$ and $c_0(S)^{\check{\otimes} k}$ for every $k \geqslant 1$.*

Note that (i) is proved for any separable $C(K)$ in Corollary 11.15, but here $c_0(S)$ may be non-separable.

Proof. Since $c_0(S)^* = \ell^1(S)$, and every element of $\ell^1(S)$ has countable support, it follows that $\mathcal{B}_w$ is generated by the point evaluations, i.e., that $\mathcal{B}_w = \mathcal{C}$.

(i): This is immediate from $\mathcal{B}_w = \mathcal{C}$.

(ii): Suppose that X is $\mathcal{C}$-measurable, and let, for $\delta, \varepsilon > 0$,

$$S_{\delta\varepsilon} := \{s \in S : \mathbb{P}(|X(s)| > \delta) > \varepsilon)\}.$$

Suppose that $S_{\delta\varepsilon}$ is infinite for some $\delta, \varepsilon > 0$. Fix these δ and ε, and let s_i, $i = 1, 2, \dots$, be an infinite sequence of distinct elements of $S_{\delta\varepsilon}$. Let

$N := \sum_i \mathbf{1}\{|X(s_i)| > \delta\}$ be the number of points s_i where $|X| > \delta$. Since $X \in c_0(S)$, N is a finite random variable, and thus there exists $M < \infty$ such that $\mathbb{P}(N > M) < \varepsilon/2$. It follows that for every s_i,

$$\begin{aligned}\mathbb{E}\big(\mathbf{1}\{|X(s_i)| > \delta\}\mathbf{1}\{N \leqslant M\}\big) &= \mathbb{P}\big(|X(s_i)| > \delta) \text{ and } N \leqslant M\big) \\ &\geqslant \mathbb{P}\big(|X(s_i)| > \delta)\big) - \mathbb{P}(N > M) > \varepsilon/2.\end{aligned}$$

Summing over all i we obtain the contradiction

$$M \geqslant \mathbb{E}\big(N\mathbf{1}\{N \leqslant M\}\big) = \sum_{i=1}^{\infty} \mathbb{E}\big(\mathbf{1}\{|X(s_i)| > \delta\}\mathbf{1}\{N \leqslant M\}\big) \geqslant \sum_{i=1}^{\infty} \varepsilon/2 = \infty.$$

Consequently, each $S_{\delta\varepsilon}$ is finite. Let $S_0 = \bigcup_{n=1}^{\infty} S_{n^{-1},n^{-1}}$. Then S_0 is a countable subset of S and if $s \notin S_0$, then $X(s) = 0$ a.s.

(iii): Let S_0 be as in (ii) and let $B_1 := \{f \in c_0(S) : \operatorname{supp}(f) \subseteq S_0\}$. Then B_1 is separable. Moreover, if $x^* \in c_0(S)^* = \ell^1(S)$ with $x^* \perp B_1$, then $x^* = (a(s)) \in \ell^1(S)$ with $a(s) = 0$ for $s \in S_0$ and thus, since $\{s : a(s) \neq 0\}$ is countable,

$$x^*(X) = \sum_{s \notin S_0} a(s)X(s) = 0 \qquad \text{a.s.}$$

Thus X is weakly a.s. separably valued.

More generally, if $k \geqslant 1$, then $B_1^{\otimes k}$ is a separable subspace of $c_0(S)^{\hat{\otimes} k}$. Suppose that $\alpha \in (c_0(S)^{\hat{\otimes} k})^*$ with $\alpha \perp B_1^{\otimes k}$. By Theorem 4.6, α is a bounded multilinear form $c_0(S)^k \to \mathbb{R}$. Let A be the countable subset given by Lemma 12.2.

Since A is countable and $X(s) = 0$ a.s. for $s \in A \backslash S_0$, $P_A X = P_{A \cap S_0} X \in B_1$ a.s., and thus a.s.

$$\langle \alpha, X^{\otimes k} \rangle = \alpha(X, \dots, X) = \alpha(P_A X, \dots, P_A X) = \langle \alpha, (P_{A \cap S_0} X)^{\otimes k} \rangle = 0.$$

This hold for every $\alpha \perp B_1^{\otimes k}$, and thus $X^{\otimes k}$ is weakly a.s. separably valued in $c_0(S)^{\hat{\otimes} k}$. Since $\iota : c_0(S)^{\hat{\otimes} k} \to c_0(S)^{\check{\otimes} k}$ is continuous, $X^{\otimes k}$ is weakly a.s. separably valued in $c_0(S)^{\check{\otimes} k}$ too. □

However, X is not always a.s. separably valued; consider for example $X := e_U \in c_0[0,1]$ where $U \sim \mathrm{U}(0,1)$. (Cf. Example 7.4, where we consider a similar random variable in $\ell^2[0,1]$.)

Theorem 12.5. *If X is a random variable in $c_0(S)$, then X is a.s. separably valued if and only if there exists a countable subset $A \subseteq S$ such that* $\operatorname{supp}(X) \subseteq A$ *a.s.*

Proof. If A is countable, then $\{f \in c_0(S) : \operatorname{supp}(f) \subseteq A\}$ is a separable subspace of $c_0(S)$; conversely, every separable subspace of $c_0(S)$ is included in some such subspace with A countable. □

Theorem 12.6. *Suppose that X is a $\mathcal{C}$-measurable $c_0(S)$-valued random variable. Let $k \geqslant 1$.*

(i) $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Dunford sense* $\iff$ *the weak* k*:th moment exists* $\iff \sup_{s\in S}\mathbb{E}\,|X(s)|^k<\infty$.

(ii) $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Pettis sense* $\iff$ *the family* $\{|X(s)|^k : s\in S\}$ *of random variables is uniformly integrable* $\iff \mathbb{E}\,|X(s)|^k\in c_0(S)$.

(iii) $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Bochner sense* $\iff \mathbb{E}\bigl(\sup_{s\in S}|X(s)|\bigr)^k<\infty$ *and there exists a countable subset* $A\subseteq S$ *such that* $\operatorname{supp}(X)\subseteq A$ *a.s.*

If $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Bochner or Pettis sense, then it is the function in* $c_0(S)^{\check{\otimes}k}=c_0(S^k)$ *given by*

$$\mathbb{E}\,X^{\check{\otimes}k}(s_1,\dots,s_k)=\mathbb{E}\bigl(X(s_1)\cdots X(s_k)\bigr). \tag{12.5}$$

Proof. The proof of Theorem 11.16 holds with a few minor changes; we use Theorems 6.10(ii)(c), 12.4(iii) and 12.5, and note that (11.6) and (11.7) hold without measurability problem since μ now is a discrete measure with countable support.

Moreover, for (ii), if s_n is any sequence of distinct elements in S, then $|X(s_n)|^k\to 0$ as $n\to\infty$. Hence, if $\{|X(s)|^k : s\in S\}$ of random variables is uniformly integrable, then $\mathbb{E}\,|X(s_n)|^k\to 0$, and it follows that $\mathbb{E}\,|X(s)|^k\in c_0(S)$. The converse is obvious.

Finally, (12.5) follows as in Theorem 11.10. □

There is an obvious analogue of Corollary 11.17, which we leave to the reader. Note that (11.8) holds for $c_0(S)$, even when this space is non-separable, because each μ in (11.9) has a countable support.

For the second projective moment, we can again use Grothendieck's theorem, and obtain the following version of Theorem 11.22.

Theorem 12.7. *Suppose that* X *is a* $\mathcal{C}$*-measurable* $c_0(S)$*-valued random variable.*

(i) $\mathbb{E}\,X^{\hat{\otimes}2}$ *exists in Dunford sense* $\iff$ $\mathbb{E}\,X^{\check{\otimes}2}$ *exists in Dunford sense* $\iff$ *the weak second moment exists* $\iff \sup_{s\in S}\mathbb{E}\,|X(s)|^2<\infty$.

(ii) $\mathbb{E}\,X^{\hat{\otimes}2}$ *exists in Pettis sense* $\iff$ $\mathbb{E}\,X^{\check{\otimes}2}$ *exists in Pettis sense* $\iff \mathbb{E}\,|X(s)|^2\in c_0(S)$.

(iii) $\mathbb{E}\,X^{\hat{\otimes}2}$ *exists in Bochner sense* $\iff$ $\mathbb{E}\,X^{\check{\otimes}2}$ *exists in Bochner sense* $\iff \mathbb{E}\bigl(\sup_{s\in S}|X(s)|\bigr)^2<\infty$ *and there exists a countable subset* $A\subseteq S$ *such that* $\operatorname{supp}(X)\subseteq A$ *a.s.*

Proof. For any given bounded bilinear form α on $c_0(S)$, there exists by Lemma 12.2 (or by Theorem 11.20) a countable subset S_0 of S such that $\alpha(f,g)$ depends only on the restrictions of f and g to S_0. Thus α can be regarded as a bilinear form on $c_0(S_0)$, and can be extended to $C(S_0^*)$, and the proof of Theorem 11.22 applies, again using Theorems 12.4(iii) and 12.5. □

It is now easy to see that Theorems 1.3 and 11.23 hold for $\mathcal{C}$-measurable random variables in $c_0(S)$; we leave the details to the reader.

The results above show that the space $c_0(S)$ behaves very well even when S is uncountable. However, the following example shows that the moments

may be degenerate. We note also (Example 12.9) that the norm of a weakly measurable random variable in $c_0(S)$ may fail to be measurable.

Example 12.8. Let $B = c_0[0,1]$ and $X = e_U$, $U \sim \mathrm{U}(0,1)$. (Cf. Example 7.4.)

Let α be a bounded k-linear form on $c_0[0,1]$, and let $A \subset [0,1]$ be a countable set as in Lemma 12.2. Since $\mathbb{P}(P_A X \neq 0) = \mathbb{P}(U \in A) = 0$, it follows from (12.2) that

$$\langle \alpha, X^{\otimes k} \rangle = \alpha(X, \dots, X) = 0 \qquad \text{a.s.}$$

Consequently, every projective moment $\mathbb{E}\, X^{\hat{\otimes} k}$ exists in Pettis sense, with $\mathbb{E}\, X^{\hat{\otimes} k} = 0$, for every $k \geqslant 1$. Hence also the injective moments exist in Pettis sense with $\mathbb{E}\, X^{\check{\otimes} k} = 0$. (No moment exists in Bochner sense, since X is not a.s. separably valued, see Theorems 12.5 and 12.6(iii).)

Example 12.9. We modify Example 12.8 by still taking $B = c_0[0,1]$ but now, as in Example 7.4, $X = a(U)e_U$ where $a : [0,1] \to (0,1]$ is a non-measurable function. Then X is weakly measurable, as in Examples 7.4 and 12.8, but $\|X\| = a(U)$ is not measurable.

13. $D[0,1]$ as a Banach space

Recall that $D = D[0,1]$ denotes the linear space of functions $[0,1] \to \mathbb{R}$ that are right-continuous with left limits, see e.g. [6, Chapter 3]. In other words $f \in D[0,1]$ if $\lim_{s \searrow t} f(s) = f(t)$ for every $t \in [0,1)$, and furthermore the left limit of f at t, which we denote by

$$f(t-) := \lim_{s \nearrow t} f(s), \tag{13.1}$$

exists for every $t \in (0,1]$. We further define

$$\Delta f(t) := f(t) - f(t-), \tag{13.2}$$

the jump at t. We may for completeness define $f(0-) := f(0)$ and thus $\Delta f(0) := 0$.

It is easily seen that each function $f \in D$ is bounded. Hence $\|f\|_D := \sup_{t \in [0,1]} |f(t)|$ defines a norm on D; the resulting topology is the *uniform topology* on D, see [6, Section 15]. The norm is complete, so D is a Banach space. Note that D is not separable, since the uncountable set of functions $\mathbf{1}_{[t,1]}$, $t \in [0,1]$, all have distance 1 to each other; this leads to measurability problems when we consider D-values random variables, as discussed in Example 3.5 and [6, Section 15]. Note also that the (separable) space $C = C[0,1]$ of continuous functions on $[0,1]$ is a closed subspace of $D[0,1]$. The following relation between $D[0,1]$ and $C[0,1]$ was proved by Corson [14, Example 2].

Theorem 13.1 ([14])**.** *Δ is a bounded linear map of $D[0,1]$ onto $c_0(0,1]$, with kernel $C[0,1]$ and norm $\|\Delta\| = 2$. Furthermore, for any $f \in D[0,1]$,*

$$\|\Delta f\|_{c_0(0,1]} = 2 \inf\bigl\{ \|f + h\|_{D[0,1]} : h \in C[0,1] \bigr\}. \tag{13.3}$$

Hence, Δ can be regarded as an isomorphism $D/C \to c_0(0,1]$ and $\frac12\Delta$ is an isometric isomorphism $D/C \to c_0(0,1]$.

Proof. It is well-known that $f \in D$ implies $\Delta f \in c_0(0,1]$, but for completeness we repeat the proof: Given $\varepsilon > 0$, for every $x \in [0,1]$ we may find an open interval $U_x = (x-\delta_x, x+\delta_x)$ such that $|f(y) - f(x)| < \varepsilon/2$ for $y \in (x, x+\delta_x)$ and $|f(y) - f(x-)| < \varepsilon/2$ for $y \in (x-\delta, x)$. (We consider only $y \in [0,1]$, and ignore $y \in U_x \backslash [0,1]$, if such points exist.) Hence, $|\Delta f(y)| \leqslant \varepsilon$ for $y \in U_x \backslash \{x\}$. Since $[0,1]$ is compact, it can be covered by a finite set of such intervals $U_{x_1}, \dots, U_{x_n}$, and then $\{y : \Delta f(y) > \varepsilon\} \subseteq \{x_1, \dots, x_n\}$. Since $\varepsilon > 0$ is arbitrary, this shows that $\Delta f \in c_0(0,1]$.

Hence $\Delta : D[0,1] \to c_0(0,1]$. It is obvious that Δ is linear and that $f \in \ker(\Delta) \iff f \in C[0,1]$. Furthermore, $\|\Delta f\|_{c_0} \leqslant 2\|f\|_D$. The function $f_0(x) := \mathbf{1}_{[1/2,1)} - \mathbf{1}_{[0,1/2)}$ in D has $\|\Delta f_0\|_{c_0} = 2 = 2\|f_0\|_D$, showing that equality can hold and thus $\|\Delta\| = 2$.

If $g \in c_{00}(0,1]$, let $\operatorname{supp}(g) = \{x_1, \dots, x_n\}$ with $0 < x_1 < \dots < x_n \leqslant 1$. Let f be the function in D that is constant on $[0, x_1)$ and $[x_n, 1]$, linear on each $[x_i, x_{i+1})$ for $1 \leqslant i < n$, and satisfies $f(x_i-) = -g(x_i)/2$, $f(x_i) = g(x_i)/2$, $1 \leqslant i \leqslant n$. Then $\Delta f = g$ and $\|f\|_D = \frac12\|g\|_{c_0(0,1]}$. (The case $g = 0$ is trivial; take $f = 0$.)

For an arbitrary $g \in c_0(0,1]$ and $\varepsilon > 0$, we can find $g_n \in c_{00}(0,1]$ such that $g = \sum_{n=1}^\infty g_n$ and $\sum_{n=1}^\infty \|g_n\|_{c_0(0,1]} \leqslant \|g\|_{c_0(0,1]} + \varepsilon$. Taking $f_n \in D$ as just constructed with $\Delta f_n = g_n$ and $\|f_n\|_D = \frac12\|g_n\|_{c_0(0,1]}$, we see that $f := \sum_{n=1}^\infty f_n \in D$ satisfies $\Delta f = \sum_{n=1}^\infty g_n = g$. Hence, Δ is onto. Moreover,

$$\|f\|_D \leqslant \sum_{n=1}^\infty \|f_n\|_D = \frac12 \sum_{n=1}^\infty \|g_n\|_{c_0(0,1]} \leqslant \frac12\|g\|_{c_0(0,1]} + \frac{\varepsilon}{2}.$$

Consequently,

$$\inf\bigl\{\|f\|_D : \Delta f = g\bigr\} \leqslant \tfrac12\|g\|_{c_0(0,1]},$$

and since $\|\Delta\| = 2$, we have equality:

$$\inf\bigl\{\|f\|_D : \Delta f = g\bigr\} = \tfrac12\|g\|_{c_0(0,1]},$$

which easily is seen to be equivalent to (13.3). □

Corollary 13.2 (Pestman [51])**.** *Every continuous linear functional $\chi \in D[0,1]^*$ is given by*

$$\chi(f) = \int_0^1 f \,\mathrm{d}\mu + \sum_{t\in(0,1]} h(t)\Delta f(t) \tag{13.4}$$

for some unique $\mu \in M[0,1]$ and $h \in \ell^1(0,1]$; conversely, (13.4) *defines a continuous linear functional on $D[0,1]$ for every such μ and h. Furthermore,*

$$\tfrac12\|\chi\|_{D^*} \leqslant \|\mu\|_M + \|h\|_{\ell^1} \leqslant 2\|\chi\|_{D^*}. \tag{13.5}$$

Note that the formally uncountable sum in (13.4) really is the countable sum $\sum_{t\in\operatorname{supp}(h)} h(t)\Delta f(t)$.

Proof. It is clear that for any $\mu \in M[0,1]$ and $h \in \ell^1(0,1]$, (13.4) defines a linear functional χ with

$$|\chi(f)| \leqslant \|\mu\|_M \|f\|_D + \|h\|_{\ell^1(0,1]} \|\Delta f\|_{c_0(0,1]} \leqslant \big(\|\mu\|_M + 2\|h\|_{\ell^1(0,1]}\big)\|f\|_D.$$

Hence, $\chi \in D^*$ and $\|\chi\|_{D^*} \leqslant \|\mu\|_M + 2\|h\|_{\ell^1(0,1]}$, showing the first part of (13.5).

Conversely, if $\chi \in D^*$, then the restriction of χ to C is a continuous linear functional on C, which by the Riesz representation theorem is given by a real measure $\mu \in M[0,1]$ with $\|\mu\|_M = \|\chi\|_C \leqslant \|\chi\|_D$. As just said, $\chi_1(f) := \int_0^1 f \,\mathrm{d}\mu$ defines a continuous linear functional on D, with $\|\chi_1\|_{D^*} \leqslant \|\mu\|_M \leqslant \|\chi\|_D$ and $\chi_1(f) = \chi(f)$ if $f \in C$. Let $\chi_2 := \chi - \chi_1$. Then $\chi_2 \in D^*$ and $\chi_2(f) = 0$ if $f \in C$. Hence χ_2 can be regarded as an element of $(D/C)^*$. By Theorem 13.1,

$$(D/C)^* \cong c_0(0,1]^* \cong \ell^1(0,1],$$

using the isometric isomorphism $\frac{1}{2}\Delta$. Hence, there exists $g \in \ell^1(0,1]$ with

$$\|g\|_{\ell^1(0,1]} = \|\chi_2\|_{(D/C)^*} = \|\chi_2\|_{D^*} \leqslant \|\chi\|_{D^*} + \|\chi_1\|_{D^*} \leqslant 2\|\chi\|_{D^*} \tag{13.6}$$

and

$$\chi_2(f) = \langle g, \tfrac{1}{2}\Delta f\rangle = \frac{1}{2}\sum_{t\in(0,1]} g(t)\Delta f(t).$$

The decomposition (13.4) follows with $h = \frac{1}{2}g$; furthermore, by (13.6),

$$\|\mu\|_M + \|h\|_{\ell^1(0,1]} \leqslant \|\chi\|_{D^*} + \|\chi\|_{D^*} = 2\|\chi\|_{D^*}.$$

Finally, to see uniqueness, suppose that $\mu \in M[0,1]$ and $h \in \ell^1(0,1]$ are such that χ defined by (13.4) equals 0. Then $0 = \chi(f) = \int_0^1 f \,\mathrm{d}\mu$ for every $f \in C[0,1]$, and thus $\mu = 0$. This implies further $0 = \chi(f) = \sum_t h(t)\Delta f(t)$ for every $f \in D[0,1]$, and thus by Theorem 13.1 $\sum_t h(t)g(t) = 0$ for every $g \in c_0(0,1]$. Hence, $h = 0$. □

Remark 13.3. The decomposition (13.4) can also be written

$$\chi(f) = \int_0^1 f(t)\,\mathrm{d}\mu_1(t) + \int_0^1 f(t-)\,\mathrm{d}\mu_2(t), \tag{13.7}$$

where $\mu_2 := -\sum_{x\in(0,1]} h(x)\delta_x$ and $\mu_1 := \mu - \mu_2$. Conversely, every pair of measures $\mu_1, \mu_2 \in M[0,1]$ defines a continuous linear functional $\chi \in D[0,1]^*$ by (13.7). However, this representation is not unique unless we impose further conditions (for example that μ_2 is discrete with $\mu_2\{0\} = 0$, as in the construction above); note that $\int f(t)\,\mathrm{d}\mu(t) = \int f(t-)\,\mathrm{d}\mu(t)$ for every $f \in D$ and every continuous measure μ, since $f(t-) = f(t)$ except on the countable set $\operatorname{supp}(\Delta f)$.

Remark 13.4. $C[0,1]$ is not a complemented subspace of $D[0,1]$, i.e. there does not exist a bounded linear projection $P : D[0,1] \to C[0,1]$, see Corson [14, Example 2]. Equivalently, there does not exist a right inverse of Δ, i.e., a bounded linear map $T : c_0(0,1] \to D[0,1]$ such that $\Delta T g = g$ for every $g \in c_0(0,1]$. (The equivalence is standard, and follows because we can

factor $I-P$ through $D/C \cong c_0(0,1]$ and thus define T by $T\Delta = I-P$, and conversely.)

To see this, suppose that such a map T exists. Then $g \mapsto Tg(t)$ is a bounded linear functional on $c_0(0,1]$ for every $t \in [0,1]$, and is thus given by some $h_t \in \ell^1(0,1]$. Let $N := \bigcup_{t\in\mathbb{Q}\cap[0,1]} \operatorname{supp}(h_t)$. Then N is countable, so there exists $s \in (0,1]\backslash N$. Consider $e_s(t) := \mathbf{1}\{t = s\} \in c_0(0,1]$. Since $\operatorname{supp}(e_s) = \{s\}$ is disjoint from $\operatorname{supp}(h_t)$ for all rational t, it follows that $Te_s(t) = \langle e_s, h_t\rangle = 0$ for all rational t. Since $Te_s \in D[0,1]$, this implies that $Te_s = 0$, which contradicts $\Delta Te_s = e_s$. (See Corson [14, Example 2] for a different proof.)

Nevertheless, Corollary 13.2 shows that $C[0,1]^* = M[0,1]$ embeds as a complemented subspace of $D[0,1]^*$, and we have $D[0,1]^* \cong C[0,1]^* \oplus c_0(0,1]^*$. The crucial fact is that each bounded linear functional χ on $C[0,1]$ extends in a canonical (linear) way to a bounded linear functional on $D[0,1]$, because χ is given by a measure $\mu \in M[0,1]$ and we can define the extension by $\int f \,\mathrm{d}\mu$ for any $f \in D$. (This is similar to the decomposition $(\ell^\infty)^* = c_0^* \oplus c_0^\perp$ since $c_0^* = \ell^1$ embeds into $(\ell^\infty)^*$.) In general, a closed subspace B_0 of a Banach space B is said to be *weakly complemented* if its annihilator is complemented, i.e. if there is a projection $P : B^* \to B_0^\perp$; it is easy to see that this is equivalent to the existence of a bounded linear map $i : B_0^* \to B^*$ such that ix^* is an extension of x^* for every $x^* \in B_0^*$, and then $B^* = i(B_0^*) \oplus B_0^\perp \cong B_0^* \oplus B_0^\perp$. Thus $C[0,1]$ is a weakly complemented subspace of $D[0,1]$. (And c_0 is a weakly complemented subspace of ℓ^∞; in fact, c_0 is a weakly complemented subspace of any Banach space $B \supset c_0$.)

14. $D[0,1]$ as a Banach algebra

The product of two functions in $D[0,1]$ is also a function in $D[0,1]$, and thus $D[0,1]$ is a commutative Banach algebra. In order to use the general theory of complex Banach algebras (which is much more satisfactory than the theory for real Banach algebras), we consider in this section $D[0,1]$ as a complex space, consisting of all complex-valued right continuous functions on $[0,1]$ with left limits. The results below will be proved for this case, but it follows immediately that they hold for the real case too, by considering the subset of real-valued functions.

The (complex) Banach algebra $D[0,1]$ has an involution given by $f \mapsto \bar{f}$, and obviously $\|f\bar{f}\| = \|f\|^2$, so $D[0,1]$ is a commutative C^*-algebra. As said in Example 11.3, every such algebra A is isometric to the space $C(K)$ of continuous functions on its maximal ideal space K, see e.g. [13, Chapter VIII.2]. The maximal ideal space can be described as the set of all (non-zero) multiplicative linear functionals (i.e., homomorphisms) $h : A \to \mathbb{C}$ with the weak*-topology (in this context known as the *Gelfand topology*), and the isometry $A \to C(K)$ maps $f \in A$ to the function $\hat{f} : h \mapsto h(f)$ on K; moreover, K is a compact Hausdorff space.

In the case of $D[0,1]$, the maximal ideal space has a simple description.

Theorem 14.1. *The linear homomorphisms on $D[0,1]$ are given by $f \mapsto f(t)$, $t \in [0,1]$, and $f \mapsto f(t-)$, $t \in (0,1]$. The maximal ideal space $\widehat{I}$ thus consists of two copies, t and $t-$, of every point in $(0,1]$, together with a single 0.*

Proof. This is a simple adaption of the standard argument for the Banach algebra of continuous functions on a compact set. The mappings $f \mapsto f(t)$ and $f \mapsto f(t-)$ are non-zero homomorphisms, and they are obviously all distinct.

Suppose that there exists another homomorphism $h : D[0,1] \to \mathbb{C}$, and let $M := \ker(h)$; thus M is a maximal ideal in $D[0,1]$. Since h differs from all $f \mapsto f(t)$ and $f \mapsto f(t-)$, there exists for each $t \in (0,1]$ two functions $f_{t+}, f_{t-} \in M$ such that $f_{t+}(t) \neq 0$ and $f_{t-}(t-) \neq 0$. By taking a suitable linear combination of f_{t+} and f_{t-} we see that there exists $f_t \in M$ such that $f_t(t) \neq 0$ and $f_t(t-) \neq 0$; thus there exists an open set U_t such that $t \in U_t$ and $|f_t|$ is bounded below in U_t. For $t = 0$ we directly find f_0 with $f_0(0) \neq 0$ and thus an open set U_0 with $0 \in U_0$ and f_0 bounded below in U_0.

The sets U_t form an open cover of the compact set $[0,1]$, and thus there exists a finite set $\{t_1, \dots, t_n\}$ such that $\bigcup_{i=1}^n U_{t_i} \supseteq [0,1]$. The function $F := \sum_{i=1}^n |f_{t_i}|^2 = \sum_{i=1}^n f_{t_i}\bar{f}_{t_i} \in M$, since M is an ideal, and $\inf_{t\in[0,1]} F(t) > 0$, by the construction. Hence $1/F \in D[0,1]$, and $1 = F \cdot (1/F) \in M$, which is a contradiction.

(Alternatively, one can use the description in Corollary 13.2 of the continuous linear functionals and show that (13.4) is multiplicative only in the cases given in the theorem.) □

We give $\widehat{I}$ the Gelfand topology, i.e., the topology generated by the functions $\hat{f} : \widehat{I} \to \mathbb{C}$; as said above $\widehat{I}$ is compact. By Theorem 14.1, the points in $\widehat{I}$ are of two types, t and $t-$; we call the points t *ordinary*. (We may for symmetry define $t+ = t$; then $\widehat{I} = \{t\pm : 0 < t \leqslant 1\} \cup \{0\}$. We shall use the notation $t+$ for the ordinary points in $\widehat{I}$ when we want to distinguish between $t+$ as an element of $\widehat{I}$ and t as an element of [0,1].) We may thus regard $[0,1]$ as a subset of $\widehat{I}$, consisting of the ordinary points. (But note that $[0,1]$ does not have the subspace topology.) We then have $D[0,1] = C(\widehat{I})$ as noted by Edgar [24, Example 5.7] (and possibly known earlier); we state this in detail:

Theorem 14.2. $D[0,1] = C(\widehat{I})$. *More precisely, each function $f \in D[0,1]$ extends to a unique continuous function on $\widehat{I}$, with $f(t-)$ given by* (13.1)*, and, conversely, if f is a continuous function on $\widehat{I}$, then the restiction to the ordinary points $t \in [0,1]$ is a function in $D[0,1]$.*

Proof. This is just a reformulation of the fact that the Gelfand transform $f \mapsto \hat{f}$ is an isomorphism $D[0,1] \to C(\widehat{I})$, using the description of $\widehat{I}$ in Theorem 14.1. □

Corollary 14.3. $D[0,1]$ *has the approximation property.*

Proof. In fact, $C(K)$ has the approximation property for every compact K, see Theorem 11.5 or [57, Example 4.2]. □

The topological space $\widehat{I}$ is called the *split interval* or *two arrow space*. (Actually, this name is often used for a modification of $\widehat{I}$ obtained by either adding an extra point $0-$ or deleting 1; both modifications are symmetric with a natural involutive homeomorphism $t\pm \mapsto (1-t)\mp$.)

Note that there is a natural total order on $\widehat{I}$, with $x < y- < y$ when $x, y \in [0,1]$ with $x < y$. (This is the lexicographic order, if we regard $\widehat{I}$ as a subset of $[0,1] \times \{-,+\}$.) We define intervals in $\widehat{I}$ in the usual way, using this order. Recall that any totally ordered set can be given a topology, the *order topology*, with a base consisting of all open intervals $(-\infty, a)$, (a, ∞), and (a, b), see e.g. [27, Problems 1.7.4 and 3.12.3].

Recall that a compact Hausdorff space is *totally disconnected* or *zero-dimensional* if it has a base consisting of open and closed sets, and *extremally disconnected* if the closure of every open set is open. (These notions are used also for non-compact spaces, but then "totally disconnected" is used in several, non-equivalent, meanings, coinciding for compact spaces, see e.g. [27].)

Theorem 14.4. *The compact Hausdorff topology on $\widehat{I}$ equals the order topology. The space is totally disconnected but not extremally disconnected.*

Furthermore, 1 is isolated (i.e., $\{1\}$ is open), each $t \in [0,1)$ has a neighbourhood base consisting of the intervals $[t,u)$, $u > t$, and each $t-$ has a neighbourhood base consisting of the intervals $[u,t-]$, $u < t$. These neighbourhood bases consist of open and closed sets.

For $t-$, there is an alternative neighbourhood base consisting of the intervals $(u,t-]$; these sets are open but not closed. (Symmetrically, the intervals $[t,u-)$ form another open neighbourhood base at t.)

Proof. We first consider the order topology and note that the given collections of sets are neighbourhood bases; this is easily seen since $[t,u) = (t-,u)$, $[u,t-] = (u-,t)$ and $\{1\} = (1-,\infty)$, with the interpretation $(0-,u) = (-\infty,u)$.

In particular, these intervals form together a base for the order topology. If J is any of these intervals, then $\mathbf{1}_J$ is a function on $\widehat{I}$ whose restriction f to $[0,1]$ belongs to $D[0,1]$, and it is easily verified that $\hat{f} = \mathbf{1}_J$. Hence, by Theorem 14.2, $\mathbf{1}_J \in C(\widehat{I})$, and thus J is open and closed in $\widehat{I}$ with the Gelfand topology. This shows that the Gelfand topology is stronger than the order topology. Since the Gelfand topology is compact and the order topology Hausdorff, the topologies coincide.

We have seen that the given base consists of open and closed sets; hence the space is totally disconnected. On the other hand, $U = \bigcup_{n=1}^{\infty} [1/(2n), 1/(2n-1))$ is an open set whose closure $\overline{U} = U \cup \{0\}$ is not open; hence $\widehat{I}$ is not extremally disconnected. □

Corollary 14.5. *The compact space $\widehat{I}$ is separable and first countable (each point has a countable neighbourhood base), but not second countable ($\widehat{I}$ does not have a countable base).*

Proof. $\widehat{I}$ is separable, since the rational numbers are dense.

We obtain countable neighbourhood bases by taking rational u only in the intervals in Theorem 14.4.

On the other hand, if $\{U_\alpha\}$ is a base for the topology, then each set $[t,1]$, which is open, contains some $U_{\alpha(t)}$ with $t \in U_{\alpha(t)}$. Then $t = \min U_{\alpha(t)}$, and thus the sets $U_{\alpha(t)}$ are distinct. It follows that every base contains (at least) $\mathfrak{c}$ elements. □

Corollary 14.6. *The compact space $\widehat{I}$ is not metrizable.*

Proof. A compact metrizable space is second countable [27, Theorem 4.2.8].

Alternatively, this follows by Theorem 11.7 since $C(\widehat{I}) = D[0,1]$ is not separable. □

Remark 14.7. The topology on $\widehat{I}$ induces on the subspace $[0,1]$ the topology where $\{1\}$ is open and each $t \in [0,1)$ has a neighbourhood base consisting of the intervals $[t,u)$, $u > t$. This (or rather the corresponding topology on $\mathbb{R}$) is known as the *Sorgenfrey line*, and is a standard source of counterexamples in topology, see e.g. [58] and [27, Examples 1.2.2, 2.3.12, 3.8.14, 5.1.31].

For example, if $\tilde{I}$ denotes $[0,1]$ with this topology, so $\tilde{I} \subset \widehat{I}$, then $\tilde{I}$ and $\tilde{I} \times \tilde{I}$ are separable, but $\tilde{I} \times \tilde{I}$ contains the closed subspace $\{(t, 1-t)\}$ which is discrete and uncountable, and thus not separable. (In particular, $\tilde{I} \times \tilde{I}$ cannot be metrizable, yielding a third proof of Corollary 14.6.) Moreover, $\tilde{I}$ is paracompact and normal, but $\tilde{I} \times \tilde{I}$ is neither [58].

15. Measurability and random variables in $D[0,1]$

We equip $C[0,1]$, $D[0,1]$ and $c_0(0,1]$ with the σ-fields generated by point evaluations; we denote these by $\mathcal{C}$ (as in Section 11), $\mathcal{D}$ and $\mathcal{C}_0$ (to distinguish it from $\mathcal{C}$). We further, as in Section 11, equip $M[0,1]$ with the σ-field $\mathcal{M}$ generated by the mappings $\mu \to \int f \,\mathrm{d}\mu$, $f \in C[0,1]$. In the present section, we shall always use these σ-fields, even if we do not always say so explicitly.

Note that $C[0,1]$ is a separable Banach space, and thus $\mathcal{C}$ equals the Borel σ-field $\mathcal{B}(C)$ on $C[0,1]$; moreover, it equals the σ-field $\mathcal{B}_w(C)$ generated by the continuous linear functionals, see Corollary 11.15.

On the other hand, $D[0,1]$ is *not* separable, and $\mathcal{D}$ is *not* equal to the Borel σ-field $\mathcal{B}(D)$, see Example 3.5. (The non-separability of D causes several problems, and is the main source of complications in the proofs below.) However, we shall see that $\mathcal{D}$ equals the σ-field $\mathcal{B}_w(D)$ generated by the continuous linear functionals. It is also well-known that $\mathcal{D}$ equals the Borel σ-field for the Skorohod topology on $D[0,1]$, see [6, Section 12]. (This is a weaker topology which is separable metric but not a vector space

topology; it is the topology commonly used on D, but it is not used in the present paper where we consider Banach spaces.) Furthermore, $\mathcal{D}$ also equals the σ-field generated by balls in D [6, Section 15].

Also $c_0(0,1]$ is not separable, but we have seen in Section 12 that it nevertheless has several nice properties.

We begin by proving some lemmas.

Lemma 15.1. *The mappings $(f,t) \mapsto f(t)$, $(f,t) \mapsto f(t-)$ and $(f,t) \mapsto \Delta f(t)$ are measurable $D\times[0,1] \to \mathbb{R}$. Furthermore, the map Δ is measurable $D[0,1] \to c_0(0,1]$.*

Proof. Since each $f \in D$ is right-continuous, $f(t) = \lim_{n\to\infty} f(\lceil nt\rceil/n)$, and $(f,t) \mapsto f(\lceil nt\rceil/n)$ is measurable for each n. The measurability of $f(t-)$ is shown similarly, using $f((\lceil nt\rceil-1)/n)$ (for $t>0$), and $\Delta f(t) = f(t)-f(t-)$.

In particular, $f \mapsto \Delta f(t)$ is measurable for each fixed t, which shows that $\Delta : D[0,1] \to c_0(0,1]$ is measurable. □

Lemma 15.2. *There exists a sequence of measurable maps $x_k : D[0,1] \to [0,1]$, $k = 1,2,\dots$, such that if $f \in D[0,1]$, then the non-zero values of $x_k(f)$, $k = 1,2,\dots$ are the jump points of f, i.e., the points $x \in [0,1]$ with $|\Delta f(x)| > 0$, without repetition.*

Proof. If $f \in D[0,1]$ and $I \subset [0,1]$ is an interval, define

$$\operatorname{var}(f;I) := \sup\bigl\{|f(x)-f(y)| : x,y \in I\bigr\}. \tag{15.1}$$

Note that it suffices to consider x,y in (15.1) that are rational, or equal to the right endpoint of I; this implies that $f \mapsto \operatorname{var}(f;I)$ is measurable for each I.

Fix $\varepsilon > 0$ and $f \in D$ and consider for each n the sequence $x_1^{(n)},\dots,x_{m(n)}^{(n)}$ of all dyadic rationals $k/2^n$ such that $1 \leqslant k \leqslant 2^n$ and $\operatorname{var}\bigl(f;[(k-1)/2^n,k/2^n]\bigr) \geqslant \varepsilon$; we assume that this sequence is in increasing order and we extend it to an infinite sequence by defining $x_k^{(n)} = 0$ for $k > m(n)$. It is easily seen that, as $n\to\infty$, $x_k^{(n)} \to x_k$ for each k, where $x_1,x_2,\dots$ are the points x where $|\Delta f(x)| \geqslant \varepsilon$, taken in increasing order and extended by 0's to an infinite sequence. By construction, for each k, the maps $f \mapsto x_k^{(n)}$ are measurable, and thus each x_k is a measurable function of f. Repeating this construction for $\varepsilon = 2^{-m}$, $m = 0,1,2,\dots$, we may rearrange the resulting sequences in a single sequence $(x_i)_1^\infty$, eliminating any repetitions of non-zero values. □

Remark 15.3. In Lemma 15.2, we may further assume that $x_k(f)$ are arranged with $|\Delta f(x_1)| \geqslant |\Delta f(x_2)| \geqslant \dots$. We will not use this, and leave the proof (using Lemma 15.1) to the reader.

Lemma 15.4. *The map $(f,\mu) \mapsto \int_0^1 f\,\mathrm{d}\mu$ is measurable $D[0,1]\times M[0,1] \to \mathbb{R}$.*

Proof. First note that the Riemann–Stieltjes sums

$$S_n(f,\mu) := \sum_{i=0}^{n-1} f(i/n)\mu[i/n,(i+1)/n) + f(1)\mu\{1\} \tag{15.2}$$

are measurable, and that, using the notation in (15.1),

$$\left|S_n(f,\mu) - \int_0^1 f\,\mathrm{d}\mu\right| \leqslant \|\mu\| \max_i \operatorname{var}(f;[i/n,(i+1)/n]). \tag{15.3}$$

Define $S^*(f;\mu) := \limsup_{n\to\infty} S_n(f;\mu)$. Then (15.3) implies

$$\begin{aligned}\left|S^*(f,\mu) - \int_0^1 f\,\mathrm{d}\mu\right| &\leqslant \|\mu\| \limsup_{n\to\infty} \max_i \operatorname{var}(f;[i/n,(i+1)/n]) \\ &= \|\mu\| \max_{x\in[0,1]} |\Delta f(x)|. \end{aligned}\tag{15.4}$$

Let $x_k(f)$ be as in Lemma 15.2 and let V_n be the (non-linear) measurable map $D \to D$ given by

$$V_n f(x) := \sum_{k=1}^{n} \Delta f(x_k(f))\mathbf{1}\{x \geqslant x_k(f)\}. \tag{15.5}$$

Thus $\max_x |\Delta(f - V_n f)(x)| = \max_{k>n} |\Delta f(x_k(f))| \to 0$ as $n \to \infty$, since $\Delta f \in c_0(0,1]$ by Theorem 13.1, and it follows from (15.4) that, for any $f \in D[0,1]$ and $\mu \in M[0,1]$,

$$S^*(f - V_n f;\mu) - \int_0^1 (f - V_n f)\,\mathrm{d}\mu \to 0, \qquad \text{as } n \to \infty. \tag{15.6}$$

Consequently, $\int_0^1 f\,\mathrm{d}\mu$ is the limit as $n \to \infty$ of the measurable functions

$$S^*(f - V_n f;\mu) + \int_0^1 V_n f\,\mathrm{d}\mu = S^*(f - V_n f;\mu) + \sum_{k=1}^{n} \Delta f(x_k(f))\mu[x_k(f),1],$$

where the last factor is measurable by Lemma 15.1 since $\mu \mapsto G_\mu(t) := \mu(t,1]$ is measurable $M[0,1] \to D[0,1]$ and $\mu[x,1] = G_\mu(x-)$ for $x > 0$. □

Theorem 15.5. *Every continuous linear form $\chi \in D[0,1]^*$ is $\mathcal{D}$-measurable. Hence $\mathcal{D} = \mathcal{B}_w(D)$, and a D-valued random variable X is $\mathcal{D}$-measurable if and only if it is weakly measurable.*

Proof. By Lemma 15.4, $f \mapsto \int_0^1 f\,\mathrm{d}\mu$ is $\mathcal{D}$-measurable for every $\mu \in M[0,1]$. It follows from Corollary 13.2 that $f \mapsto \langle \chi, f\rangle$ is $\mathcal{D}$-measurable for every $\chi \in D[0,1]^*$.

This implies $\mathcal{B}_w(D) \subseteq \mathcal{D}$. Conversely, each point evaluation $f(t) = \langle \delta_t, f\rangle$ where $\delta_t \in D[0,1]^*$. Hence $\mathcal{D} = \mathcal{B}_w(D)$. □

Remark 15.6. We can extend the map in Lemma 15.4 to $(f,\chi) \mapsto \langle f,\chi\rangle$ for $f \in D$, $\chi \in D^*$; however, this map is *not* jointly measurable for $\mathcal{D}$ and the σ-field $\mathcal{D}^*$ on D^* generated by $\chi \mapsto \langle \chi, f\rangle$, $f \in D$. (The map is separably measurable by Theorem 15.5.) In fact, let $a(t) : [0,1] \to \mathbb{R}$ be a

non-measurable function. Then the maps $t \mapsto f_t := \mathbf{1}_{[t,1]}$ and $t \mapsto \chi_t := g \mapsto a(t)\Delta g(t)$ are measurable for $\mathcal{D}$ and $\mathcal{D}^*$, respectively; note that for any fixed g, $t \mapsto \langle \chi_t, g\rangle = a(t)\Delta g(t)$ is measurable since it has countable support. However, $t \mapsto \langle f_t, \chi_t\rangle = a(t)$ is non-measurable.

Remark 15.7. We have $D[0,1] = C(\widehat{I})$, and it follows from Lemma 15.1 that $\mathcal{D}$ also equals the σ-field generated by point evaluations on $\widehat{I}$. We can also consider the map $(f,t) \mapsto f(t)$ for $f \in D[0,1]$ and $t \in \widehat{I}$. However, in contrast to Lemma 15.1, this map is *not* jointly $\mathcal{D} \times \mathcal{B}(\widehat{I})$-measurable, where $\mathcal{B}(\widehat{I})$ is the Borel σ-field on $\widehat{I}$. To see this, let $A \subset [0,1]$ be a non-measurable set (with $0 \in A$) and let $\varphi : I \to \widehat{I}$ be the function defined by $\varphi(u) := u+$ when $u \in A$ and $\varphi(u) := u-$ when $u \notin A$; then φ is increasing and thus Borel measurable. Furthermore, the function $u \mapsto f_u := \mathbf{1}_{[u,1]}$ is a measurable map $[0,1] \to D[0,1] = C(\widehat{I})$. Hence, $u \mapsto (f_u, \varphi(u))$ is measurable $[0,1] \to C(\widehat{I}) \times \widehat{I}$. However, the composition $u \mapsto f_u(\varphi(u)) = \mathbf{1}_A(u)$ is non-measurable, showing that $(f,t) \mapsto f(t)$ is not measurable on $C(\widehat{I}) \times \widehat{I}$.

One of the main purposes of this section is to prove the following theorem, extending Theorem 15.5 to multilinear forms. (Note that the corresponding result for $C[0,1]$ is immediate, since $\mathcal{C}$ equals the Borel σ-field on $C[0,1]$ and $C[0,1]$ is separable, which imply that the product σ-field $\mathcal{C}^m$ on $(C[0,1])^m$ equals the Borel σ-field.)

Theorem 15.8. *Every bounded multilinear form $T : (D[0,1])^m \to \mathbb{R}$, for any $m \geqslant 1$, is $\mathcal{D}$-measurable.*

Proof. We shall prove the more general result that for any $m, \ell \geqslant 0$, any bounded multilinear form $\alpha : (D[0,1])^m \times (c_0(0,1])^\ell \to \mathbb{R}$ is measurable. We do this by induction over m.

First, assume that $m = 0$, so $\alpha : (c_0(0,1])^\ell \to \mathbb{R}$. Recall the projections P_A in $c_0(S)$ defined in (12.1) and let A be a countable subset as in Lemma 12.2 (with $S = (0,1]$ and $k = \ell$). If A is finite, then $\alpha(P_A f_1, \dots, P_A f_\ell)$ is a finite linear combination of products of point evaluations, and thus measurable.

If A is infinite, write $A = \{x_1, x_2, \dots\}$ and define $P_n := P_{\{x_1,\dots,x_n\}}$. For any $f \in c_0(0,1]$, $P_n f \to P_A f$ in $c_0(0,1]$ as $n \to \infty$, and thus $\alpha(P_n f_1, \dots, P_n f_\ell) \to \alpha(P_A f_1, \dots, P_A f_\ell)$. Each map $\alpha(P_n f_1, \dots, P_n f_\ell)$ is measurable, and the result in the case $m = 0$ follows by (12.2).

Now suppose that $m \geqslant 1$. Consider first the restriction α_0 of α to $C \times D^{m-1} \times c_0(0,1]^\ell$. For fixed $f_2, \dots, f_m \in D$ and $g_1, \dots, g_\ell \in c_0(0,1]$, α_0 is a bounded linear functional on $C[0,1]$; thus α_0 can be regarded as a multilinear map $T_0 : D^{m-1} \times c_0(0,1]^\ell \to C[0,1]^* = M[0,1]$. For a fixed $f_1 \in C[0,1]$, the map $(f_2, \dots, f_m, g_1, \dots, g_\ell) \mapsto \alpha_0(f_1, f_2, \dots, f_m, g_1, \dots, g_\ell)$ is measurable by induction, which by the definition of the σ-field $\mathcal{M}$ in $M[0,1]$ shows that T_0 is measurable.

Define

$$\alpha_1(f_1, f_2, \dots, f_m, g_1, \dots, g_\ell) := \int_0^1 f_1 \, \mathrm{d}T_0(f_2, \dots, f_m, g_1, \dots, g_\ell). \tag{15.7}$$

Then α_1 is a bounded multilinear form $(D[0,1])^m \times (c_0(0,1])^\ell \to \mathbb{R}$, and α_1 is measurable by Lemma 15.4 and the measurability of T_0. Moreover, if $f_1 \in C[0,1]$, then $\alpha_1(f_1, f_2, \dots, f_m, g_1, \dots, g_\ell) = \alpha(f_1, f_2, \dots, f_m, g_1, \dots, g_\ell)$.

Define $\alpha_2 := \alpha - \alpha_1$. Then $\alpha_2 = 0$ on $C \times D^{m-1} \times c_0(0,1]^\ell$, so α_2 can be regarded as a multilinear form on $(D/C) \times D^{m-1} \times c_0(0,1]^\ell$. By Theorem 13.1, $\Delta : D/C \to c_0(0,1]$ is an isomorphism, so there exists a bounded multilinear form $\alpha_3 : c_0(0,1] \times D^{m-1} \times c_0(0,1]^\ell \to \mathbb{R}$ such that

$$\alpha_2(f_1, f_2, \dots, f_m, g_1, \dots, g_\ell) = \alpha_3(\Delta f_1, f_2, \dots, f_m, g_1, \dots, g_\ell). \tag{15.8}$$

Then α_3 is measurable by induction, and since $\Delta : D \to c_0(0,1]$ is measurable by Lemma 15.1, (15.8) shows that α_2 is measurable.

Thus, $\alpha = \alpha_1 + \alpha_2$ is measurable. □

Corollary 15.9. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable. Then $X^{\otimes k}$ is weakly measurable in $D^{\hat{\otimes} k}$ and in $D^{\check{\otimes} k}$ for every $k \geqslant 1$.*

Proof. Immediate by Theorems 4.6 and 15.8 and the continuous map $\iota : D^{\hat{\otimes} k} \to D^{\check{\otimes} k}$. □

We have seen in Example 3.5 that a $\mathcal{D}$-measurable D-valued random variable X is not always a.s. separably valued. In fact, the following theorem describes the situation precisely: X is a.s. separably valued if and only if the jumps only occur in a fixed countable set. (Cf. Theorem 12.5 for $c_0(S)$.)

Lemma 15.10. *A subset $D_0 \subseteq D$ is separable if and only if there exists a countable subset N of $[0,1]$ such that $\operatorname{supp}(\Delta f) := \{t : |\Delta f(t)| \neq 0\} \subseteq N$ for each $f \in D_0$.*

Proof. If $N \subset [0,1]$ is countable, then

$$D_1 := \{f \in D[0,1] : \operatorname{supp}(\Delta f) \subseteq N\} \tag{15.9}$$

is a separable subspace of D, for example because Δ induces an isomorphism $D_1/C \cong c_0(N)$, and both C and $c_0(N)$ are separable.

Conversely, if $D_0 \subseteq D$ is separable and $\{f_n\}_{n=1}^\infty$ is a countable dense subset, then $N := \bigcup_n \operatorname{supp}(\Delta f_n)$ is countable and $\operatorname{supp}(\Delta f) \subseteq N$ for every $f \in D_0$. □

Theorem 15.11. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable. Then X is a.s. separably valued if and only if there exists a countable subset N of $[0,1]$ such that a.s. $\operatorname{supp}(\Delta X) := \{t : |\Delta X(t)| \neq 0\} \subseteq N$.*

Proof. Immediate from Lemma 15.10. □

However, a weak version always holds.

Lemma 15.12. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable. Then there is a countable set $N \subset [0,1]$ such that for each $t \in [0,1]\setminus N$, $\Delta X(t) = 0$ a.s.*

The exceptional null set may depend on t, in contrast to the condition in Theorem 15.11.

Proof. By Theorem 12.4(ii) applied to $\Delta X \in c_0(0,1]$, which is $\mathcal{C}_0$-measurable by Lemma 15.1. □

Theorem 15.13. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable. Then X is weakly a.s. separably valued.*

Proof. Let N be as in Lemma 15.12 and let $D_1 \subset D$ be the separable subspace defined in (15.9).

Suppose that $\chi \in D^*$ with $\chi \perp D_1$, and represent χ as in (13.4). Since $\chi \perp D_1$ and $C \subseteq D_1$, we see that if $f \in C$, then $0 = \chi(f) = \int_0^1 f \,\mathrm{d}\mu$; thus $\mu = 0$ and $\chi(f) = \sum_t h(t)\Delta f(t)$ for all $f \in D$. Moreover, if $t \in N$, then $\mathbf{1}_{[t,1]} \in D_1$, and thus $0 = \chi(\mathbf{1}_{[t,1]}) = h(t)$. Hence, any $\chi \perp D_1$ has the form

$$\chi(f) = \sum_{t\in M} h(t)\Delta f(t) \tag{15.10}$$

for some countable set $M = \{t : h(t) \neq 0\}$ with $M \cap N = \varnothing$.

For each $t \in M$, we have $t \notin N$ and thus $X(t) = 0$ a.s., and thus a.s. $\chi(X) = 0$ by (15.10). □

We extend this to $X^{\otimes k}$.

Theorem 15.14. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable. Then, for every $k \geqslant 1$, $X^{\otimes k}$ is weakly a.s. separably valued in the projective and injective tensor products $(D[0,1])^{\hat\otimes k}$ and $(D[0,1])^{\check\otimes k}$.*

Proof. It suffices to consider the projective tensor product since $\iota : (D[0,1])^{\hat\otimes k} \to (D[0,1])^{\check\otimes k}$ is continuous.

Let, again, N be as in Lemma 15.12 and let $D_1 \subset D$ be the separable subspace (15.9). Then $D_1^{\otimes k}$ is a separable subspace of $D^{\hat\otimes k}$. We claim that if $\chi \in (D^{\hat\otimes k})^*$ and $\chi \perp D_1^{\otimes k}$, then $\chi(X^{\otimes k}) = 0$ a.s., which proves the lemma. By Theorem 4.6, χ is a bounded k-linear form on D. We will prove the claim by induction, using a more elaborate claim.

For $\ell, m \geqslant 0$, let $L_{\ell,m}$ be the set of all $(\ell+m)$-linear forms $\alpha \in L(D^{\ell+m}, \mathbb{R})$ such that

$$\alpha(f_1, \dots, f_{\ell+m}) = 0 \quad \text{if } f_1, \dots, f_{\ell+m} \in D_1, \tag{15.11}$$

and

$$\alpha(f_1, \dots, f_{\ell+m}) = 0 \quad \text{if } f_i \in C \text{ for some } i \leqslant \ell. \tag{15.12}$$

Claim. If $\ell, m \geqslant 0$, then $\alpha(X, \dots, X) = 0$ a.s. for every $\alpha \in L_{\ell,m}$.

The case $\ell = 0$, $m = k$ yields the claim $\chi(X^{\otimes k}) = 0$ a.s. above, proving the lemma. We prove the claim by induction on m.

Step 1: $m = 0$. If $m = 0$, then (15.12) shows that α can be seen as an ℓ-linear form on $D/C \cong c_0(0,1]$. Hence there exists an ℓ-linear form $\beta : c_0(0,1]^\ell \to \mathbb{R}$ such that

$$\alpha(f_1, \dots, f_\ell) = \beta(\Delta f_1, \dots, \Delta f_\ell).$$

By (15.11), $\beta(g_1, \dots, g_\ell) = 0$ if $\operatorname{supp} g_i \subseteq N$ for every i. Let A be a countable subset of $(0,1]$ given by Lemma 12.2 (applied to β). Then, as in the proof of Theorem 12.4(iii), $P_A(\Delta X) = P_{A\cap N}(\Delta X)$ a.s., and thus a.s.

$$\alpha(X, \dots, X) = \beta(\Delta X, \dots, \Delta X) = \beta(P_A(\Delta X), \dots, P_A(\Delta X)) = 0.$$

Step 2: $m \geqslant 1$. Let $k := \ell + m \geqslant 1$. If $k = 1$, then the result is Theorem 15.13 (and its proof). Thus assume $k \geqslant 2$.

For fixed $f_1, \dots, f_{k-1}$, the map $f_k \mapsto \alpha(f_1, \dots, f_k)$ is a bounded linear form on D. The restriction to C gives an element of $C^* = M[0,1]$, i.e. a signed measure $\mu_{f_1,\dots,f_{k-1}}$ on $[0,1]$. Define, similarly to (15.7),

$$\alpha_1(f_1, \dots, f_k) := \int_0^1 f_k \, \mathrm{d}\mu_{f_1,\dots,f_{k-1}}.$$

Then α_1 is a bounded k-linear form on D, and thus so is $\alpha_2 := \alpha - \alpha_1$.

If $f_1, \dots, f_{k-1} \in D_1$ and $f_k \in C \subseteq D_1$, then $\alpha(f_1, \dots, f_k) = 0$ by (15.11); hence $\mu_{f_1,\dots,f_{k-1}} = 0$ and thus $\alpha_1(f_1, \dots, f_k) = 0$ for any $f_k \in D$. Similarly, by (15.12), if $f_i \in C$ for some $i \leqslant \ell$, then $\mu_{f_1,\dots,f_{k-1}} = 0$ and $\alpha_1(f_1, \dots, f_k) = 0$. Hence, (15.11) and (15.12) hold for α_1, i.e., $\alpha_1 \in L_{\ell,m}$. Consequently, $\alpha_2 = \alpha - \alpha_1 \in L_{\ell,m}$ too.

Moreover, if $f_k \in C$, then the definition of α_1 yields $\alpha_1(f_1, \dots, f_k) = \alpha(f_1, \dots, f_k)$ and thus

$$\alpha_2(f_1, \dots, f_k) = \alpha(f_1, \dots, f_k) - \alpha_1(f_1, \dots, f_k) = 0.$$

Hence (15.12) holds for $i = k = \ell + m$ too, so (after relabelling) $\alpha_2 \in L_{\ell+1,m-1}$, and by induction $\alpha_2(X, \dots, X) = 0$ a.s.

Return to α_1. For each fixed $f_k \in C$, $(f_1, \dots, f_{k-1}) \mapsto \alpha(f_1, \dots, f_{k-1}, f_k)$ defines a $(k-1)$-linear form on D, which by (15.11)–(15.12) belongs to $L_{\ell,m-1}$. By induction, thus $\alpha(X, \dots, X, f_k) = 0$ a.s. for each fixed $f_k \in C$. By taking f_k in a countable dense subset of C, it follows that a.s. $\alpha(X, \dots, X, f) = 0$ for every $f \in C$. Thus, a.s., $\mu_{X,\dots,X} = 0$ and $\alpha_1(X, \dots, X, f) = 0$ for every $f \in D$; in particular, $\alpha_1(X, \dots, X, X) = 0$.

We have shown that $\alpha(X, \dots, X) = \alpha_1(X, \dots, X) + \alpha_2(X, \dots, X) = 0$ a.s. when $\alpha \in L_{\ell,m}$, which proves the claim. □

16. Moments of D-valued random variables

After the preliminaries in the last sections, we can prove analogues of Theorem 11.16–11.23 for $D[0,1]$. (Note that although $D[0,1] = C(\widehat{I})$, we cannot

use these theorems in Section 11 directly, since $D[0,1]$ is non-separable and thus $\widehat{I}$ is not metrizable.)

We begin with describing the space $D[0,1]^{\check{\otimes}k}$ where the moments live. By Theorem 11.6, $D[0,1]^{\check{\otimes}k} = C(\widehat{I})^{\check{\otimes}k} = C(\widehat{I}^k)$. We return from $\widehat{I}$ to [0,1] by taking the restrictions of the functions in $C(\widehat{I})^k$ to the (dense) subset $[0,1]^k \subset \widehat{I}^k$ (i.e., the normal points in $\widehat{I}^k$).

Definition 16.1. $D([0,1]^k)$ is the Banach space of all functions $[0,1]^k \to \mathbb{R}$ that have a continuous extension to $\widehat{I}^k$; $D([0,1]^k)$ is equipped with the supremum norm. Thus $D([0,1]^k)$ is naturally isometric to $C(\widehat{I}^k)$.

This means that $f \in D([0,1]^k)$ if at each $(t_1,\dots,t_k) \in [0,1]^k$, f has limits in the 2^k octants (with obvious modifications at the boundary). More precisely, taking $k=2$ for notational convenience, $f \in D([0,1]^2)$ if and only if, for each $(s,t) \in [0,1]^2$, the limits

$$\begin{aligned}
f(s+,t+) &:= \lim_{\substack{s'\to s,\ s'\geqslant s\\ t'\to t,\ t'\geqslant t}} f(s',t'),\\
f(s+,t-) &:= \lim_{\substack{s'\to s,\ s'\geqslant s\\ t'\to t,\ t'< t}} f(s',t'),\\
f(s-,t+) &:= \lim_{\substack{s'\to s,\ s'< s\\ t'\to t,\ t'\geqslant t}} f(s',t'),\\
f(s-,t-) &:= \lim_{\substack{s'\to s,\ s'< s\\ t'\to t,\ t'< t}} f(s',t')
\end{aligned}$$

exist (as finite real numbers), except that we ignore all cases with an argument $0-$. Note the slight asymmetry; we use $\geqslant$ but $<$. Note also that necessarily $f(s+,t+) = f(s,t)$ when it exists.

Theorem 16.2. *If $k \geqslant 2$, then $D[0,1]^{\hat{\otimes}k} \subsetneq D[0,1]^{\check{\otimes}k} = D([0,1]^k)$. (The subspace $D[0,1]^{\hat{\otimes}k}$ of $D([0,1]^k)$ is not closed, and the norms are different and not equivalent on $D[0,1]^{\hat{\otimes}k}$.)*

Proof. The equality $D[0,1]^{\check{\otimes}k} = D([0,1]^k)$ follows by Theorem 11.6 applied to $C(\widehat{I})$, together with Definition 16.1, and the inclusion $D[0,1]^{\hat{\otimes}k} \subseteq D[0,1]^{\check{\otimes}k}$ by Corollary 11.8, or by Corollary 14.3 and Theorem 8.3.

The claim that $D[0,1]^{\hat{\otimes}k}$ is not a closed subspace of $D([0,1]^k)$, is a special case of the general fact in Remark 11.9, applied to $C(\widehat{I})$, but we give also a direct proof, using an argument from Varopoulos [65] (where $C(\mathbb{T})\hat{\otimes}C(\mathbb{T})$ is studied, and more generally $C(K)\hat{\otimes}C(K)$ for an abelian compact group K.)

By the closed graph theorem the claim is equivalent to the claim that that the norm on $D[0,1]^{\hat{\otimes}k}$ is not equivalent to the norm inherited from $D([0,1]^k)$, i.e., the sup norm.

It suffices to consider the case $k = 2$. If $F \in L^1([0,1]^2)$, let $\hat{F}(m,n) := \iint F(s,t)e^{2\pi \mathrm{i}(ms+nt)}\,\mathrm{d}s\,\mathrm{d}t$ be its Fourier coefficients. Since $D[0,1] \subset L^2[0,1]$, Parseval's identity and Hölder's inequality imply, together wih (4.4), that if $F \in D[0,1]^{\hat{\otimes}2}$, then $\sum_{n=-\infty}^{\infty} |\hat{F}(n,n)| \leqslant \|F\|_{D[0,1]^{\hat{\otimes}2}}$. However, there exist trigonometric polynomials $f(t) = \sum_n a_n e^{2\pi \mathrm{i}nt}$ such that $\sup_t |f(t)| = 1$ with $\sum_n |a_n|$ arbitrarily large. (For example Cesáro means of the Fourier series $\sum_2^\infty \sin 2\pi int/(n\log n)$, which represents a continuous function [70, Section V.1].) Taking $F(s,t) := f(s+t) \in C([0,1]^2)$ we have $\|F\|_{D([0,1]^2)} = \sup|F(s,t)| = 1$ and $\|F\|_{D[0,1]^{\hat{\otimes}k}} \geqslant \sum_n |a_n|$ which is arbitrarily large. □

Example 16.3. Let W be standard Brownian motion, regarded as a random variable in $D[0,1]$. Then all (projective and injective) moments exist (in Bochner sense), and are the same as for W regarded as an random variable in $C[0,1]$, see Example 11.11 and Remark 6.22. In particular, $\mathbb{E}\, W^{\check{\otimes}2}$ is the covariance function $s \wedge t$ regarded as an element of $D[0,1]^{\check{\otimes}2} = D([0,1]^2)$, and $\mathbb{E}\, W^{\hat{\otimes}2}$ is the same function regarded as an element of the subspace $D[0,1]^{\hat{\otimes}2}$. Similarly, the fourth moment is given by the function (11.4).

Example 16.4. The $D[0,1]$-valued random variable $X = \mathbf{1}_{[U,1]}$ in Examples 3.5 and 5.18 is not a.s. separably valued, see Example 5.18, and thus it has no moments in Bochner sense; however, all moments exist in Pettis sense by Theorem 16.10 below. The k:th moment is given by the function

$$\mathbb{E}\big(X(t_1)\cdots X(t_k)\big) = \mathbb{E}\,\mathbf{1}\{U \leqslant t_1, \dots, U \leqslant t_k\} = \min\{t_1, \dots, t_k\}.$$

This function is continuous, and thus belongs to $C([0,1]^k) \subset D([0,1]^k)$.

Note that the second moment is the same as for Brownian motion (Example 16.3), but not the fourth (or any other moment).

Example 16.5. For a simple example with a discontinuous function in $D([0,1]^2)$ as second moment, let $S_n = \sum_1^n \xi_i$ be a simple random walk, with ξ_i i.i.d. and $\mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{2}$. Let $X_n(t) := S_{\lfloor nt \rfloor}/\sqrt{n}$. Then the second moment of $X_n \in D[0,1]$ is the function in $D([0,1]^2)$ given by $\mathbb{E}(X_n(s)X_n(t)) = \lfloor n(t \wedge u)\rfloor/n$.

We turn to conditions for the existence of injective moments.

Theorem 16.6. *Suppose that X is a $\mathcal{D}$-measurable $D[0,1]$-valued random variable. Let $k \geqslant 1$.*

(i) $\mathbb{E}\, X^{\check{\otimes}k}$ *exists in Dunford sense* $\iff$ *the weak k:th moment exists* $\iff \sup_{t\in[0,1]} \mathbb{E}\,|X(t)|^k < \infty$.

(ii) $\mathbb{E}\, X^{\check{\otimes}k}$ *exists in Pettis sense* $\iff$ *the family $\{|X(t)|^k : t \in [0,1]\}$ of random variables is uniformly integrable.*

(iii) $\mathbb{E}\, X^{\check{\otimes}k}$ *exists in Bochner sense* $\iff \mathbb{E}\big(\sup_{t\in[0,1]} |X(t)|\big)^k < \infty$ *and there is a countable set $N \subset [0,1]$ such that* $\operatorname{supp}(\Delta X) \subseteq N$ *a.s.*

If $\mathbb{E}\, X^{\check{\otimes}k}$ exists in Bochner or Pettis sense, then it is the function in $D[0,1]^{\check{\otimes}k} = D([0,1]^k)$ given by

$$\mathbb{E}\, X^{\check{\otimes}k}(t_1,\dots,t_k) = \mathbb{E}\big(X(t_1)\cdots X(t_k)\big), \qquad t_1,\dots,t_k \in [0,1]. \tag{16.1}$$

Proof. We use $D[0,1] = C(\widehat{I})$ and try to argue as in the proof of Theorem 11.16; however, several modifications are needed.

(i): The forward implications are immediate as in Theorem 11.16, but for the remaining implication we argue somewhat differently.

Consider a linear form $\mu \in (D[0,1]^{\check{\otimes}k})^* = C(\widehat{I}^k)^* = M_{\mathsf{r}}(\widehat{I}^k)$ with $\|\mu\| \leqslant 1$. Then μ can be regarded as an integral multilinear form on $D[0,1]^k$; by Corollary 15.9, $\langle \mu, X^{\otimes k}\rangle$ is measurable.

A serious technical problem is that $X(t,\omega)$ is in general *not* jointly measurable on $\widehat{I}\times\Omega$, see Remark 15.7; hence we cannot use Fubini's theorem as in (11.6) and (11.7). We circumvent this as follows. Similarly to (11.7),

$$\begin{aligned}\big|\langle \mu, X^{\otimes k}\rangle\big| &= \left|\int_{\widehat{I}^k} X(t_1)\cdots X(t_k)\,\mathrm{d}\mu(t_1,\dots,t_k)\right| \\ &\leqslant \frac{1}{k}\int_{\widehat{I}^k}\sum_{i=1}^{k} |X(t_i)|^k\,\mathrm{d}|\mu|(t_1,\dots,t_k) \\ &= \int_{\widehat{I}} |X(t)|^k\,\mathrm{d}\nu(t)\end{aligned} \tag{16.2}$$

for some positive measure ν on $\widehat{I}$ with $\|\nu\|_{M(\widehat{I})} = \|\mu\| \leqslant 1$. We now regard ν as a continuous linear functional χ on $C(\widehat{I}) = D[0,1]$, and represent it by Corollary 13.2 as $f \mapsto \int_0^1 f\,\mathrm{d}\nu_1 + \sum_{t\in(0,1]} h(t)\Delta f(t)$ with

$$\|\nu_1\|_M + \|h\|_{\ell^1} \leqslant 2\|\nu\|_{D^*} \leqslant 2. \tag{16.3}$$

Thus, (16.2) yields

$$|\langle \mu, X^{\otimes k}\rangle| \leqslant \int_0^1 |X(t)|^k\,\mathrm{d}\nu_1(t) + \sum_{t\in(0,1]} h(t)\Delta|X|^k(t). \tag{16.4}$$

Since $X(t,\omega)$ is jointly measurable on $[0,1]\times\Omega$ by Lemma 15.1 and the sum in (16.4) is countable, we can take the expectation in (16.4) and use Fubini's theorem here, obtaining

$$\mathbb{E}\big|\langle \mu, X^{\otimes k}\rangle\big| \leqslant \int_0^1 \mathbb{E}\,|X(t)|^k\,\mathrm{d}|\nu_1|(t) + \sum_{t\in(0,1]} |h(t)|\big(\mathbb{E}\,|X(t)|^k + \mathbb{E}\,|X(t-)|^k\big). \tag{16.5}$$

Now suppose that $\sup_{t\in[0,1]} \mathbb{E}\,|X(t)|^k \leqslant C$. Then Fatou's lemma applied to a sequence $t_k \nearrow t$ yields also $\mathbb{E}\,|X(t-)|^k \leqslant C$ for every $t > 0$, and (16.5) implies

$$\mathbb{E}\big|\langle \mu, X^{\otimes k}\rangle\big| \leqslant C\|\nu_1\|_M + 2C\|h\|_{\ell^1} \leqslant 4C. \tag{16.6}$$

This shows that $\langle \mu, X^{\otimes k} \rangle$ is integrable for every $\mu \in (D[0,1]^{\check{\otimes} k})^*$, which shows that $X^{\check{\otimes} k}$ is Dunford integrable, completing the proof of (i).

(ii): As in Theorem 11.16, the forward implication follows by Theorem 6.19(i), taking $x^* = \delta_t$, $t \in [0,1]$.

For the converse, apply (16.6) to $\mathbf{1}_E X$, where $E \in \mathcal{F}$ is an event, and obtain

$$\mathbb{E}\big|\mathbf{1}_E \langle \mu, X^{\otimes k} \rangle\big| \leqslant 4 \sup_{t\in[0,1]} \mathbb{E}\big(\mathbf{1}_E |X(t)|^k\big), \tag{16.7}$$

which by (5.4) shows that if the family $\{|X(t)|^k\}$ is uniformly integrable, then the family $\{\langle \mu, X^{\check{\otimes} k} \rangle : \|\mu\| \leqslant 1\}$ is uniformly integrable. Since Theorem 15.14 shows that $X^{\check{\otimes} k}$ is weakly a.s. separably valued, it follows by Theorem 5.8 that $\mathbb{E}\, X^{\check{\otimes} k}$ exists in Pettis sense.

(iii): Immediate by Theorems 6.7 and 15.11.

The final claim follows by Theorem 16.2 and the argument in the proof of Theorem 11.10, which yields (11.3) in the present setting too. □

The case $k = 1$ in Theorem 16.6 gives the following characterisations of the existence of the expectation $\mathbb{E}\, X$ of a $D[0,1]$-valued random variable.

Corollary 16.7. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable.*

(i) *$\mathbb{E}\, X$ exists in Dunford sense* $\iff$ $\sup_{t\in[0,1]} \mathbb{E}\,|X(t)| < \infty$.

(ii) *$\mathbb{E}\, X$ exists in Pettis sense* $\iff$ *the family $\{X(t) : t \in [0,1]\}$ of random variables is uniformly integrable.*

(iii) *$\mathbb{E}\, X$ exists in Bochner sense* $\iff$ $\mathbb{E}\big(\sup_{t\in[0,1]} |X(t)|\big) < \infty$ *and there is a countable set $N \subset [0,1]$ such that $\operatorname{supp}(\Delta X) \subseteq N$ a.s.*

In the Pettis and Bochner cases, $\mathbb{E}\, X \in D[0,1]$ is the function $t \mapsto \mathbb{E}\, X(t)$. □

As in Theorem 11.19, an *even* injective moment exists in Pettis sense if and only if it exists in Dunford sense and it belongs to $D[0,1]^{\check{\otimes} k} = D([0,1]^k)$.

Theorem 16.8. *Suppose that X is a $\mathcal{D}$-measurable $D[0,1]$-valued random variable such that $\sup_{t\in[0,1]} \mathbb{E}\,|X(t)|^2 < \infty$. Suppose that $k \geqslant 2$ is even. Then the following are equivalent.*

(i) *$\mathbb{E}\, X^{\check{\otimes} k}$ exists in Pettis sense.*

(ii) *The function $g(t) := \mathbb{E}\, X(t)^k$ belongs to $D[0,1]$, and $g(t-) := \mathbb{E}\, X(t-)^k$, $0 < t \leqslant 1$.*

Proof. (i) $\Longrightarrow$ (ii): If t_n is a sequence in [0,1] and $t_n \searrow t$, then $X(t_n)^k \to X(t)^k$, while if $t_n \nearrow t$, then $X(t_n)^k \to X(t-)^k$. By Theorem 16.6, $\{|X(t)|^k : t \in [0,1]\}$ is uniformly integrable; thus it follows that $\mathbb{E}\, X(t_n)^k \to \mathbb{E}\, X(t)^k$ or $\mathbb{E}\, X(t_n)^k \to \mathbb{E}\, X(t-)^k$, respectively.

(ii) $\Longrightarrow$ (i): If t_n is any sequence in [0,1], there exist a subsequence (still denoted t_n for convenience) such that $t_n \to t$ for some $t \in [0,1]$; we may furthermore select the subsequence such that either $t_n \geqslant t$ or $t_n < t$ for all n. In the first case, $X(t_n) \to X(t)$ and thus $|X(t_n)|^k \to |X(t)|^k$;

furthermore, by (ii) and the fact that k is even, $\mathbb{E}\,|X(t_n)|^k = g(t_n) \to g(t) = \mathbb{E}\,|X(t)|^k$. In the second case, similarly $|X(t_n)|^k \to |X(t-)|^k$ and $\mathbb{E}\,|X(t_n)|^k = g(t_n) \to g(t-) = \mathbb{E}\,|X(t-)|^k$. In both cases it follows that $|X(t_n)|^k$ converges in $L^1(\mathbb{P})$ (to $|X(t)|^k$ or $|X(t-)|^k$), see [34, Theorem 5.5.2]. Consequently, $\{|X(t)|^k : t \in [0,1]\}$ is a relatively compact subset of $L^1(P)$, and in particular relatively weakly compact and thus uniformly integrable [22, Theorem IV.8.11]. Thus (i) follows by Theorem 16.6(ii). □

The extra condition $g(t-) := \mathbb{E}\,X(t-)^k$ in Theorem 16.8 cannot be omitted, as seen by the following example.

Example 16.9. Let $I_n := [1-2^{-n}, 1-2^{-n-1})$ and let X equal $2^{n/2}\mathbf{1}_{I_n}$ with probability 2^{-n}, $n \geqslant 1$. Then $\mathbb{E}\,X(t)^2 = \mathbf{1}_{[1/2,1)}(t) \leqslant 1$ but $\{X(t)^2\}$ is not uniformly integrable; hence it follows from Theorem 16.6 that $\mathbb{E}\,X^{\check{\otimes}2}$ exists in Dunford sense but not in Pettis sense. Note that $g(t) := \mathbb{E}\,X(t)^2 = \mathbf{1}_{[1/2,1)}(t) \in D[0,1]$ but $\mathbb{E}\,X(1-)^2 = 0 \neq g(1-)$.

For projective moments, we do not know any general necessary and sufficient conditions for existence in Pettis or Dunford sense, but we have a simple sufficient condition.

Theorem 16.10. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable, and suppose that $\mathbb{E}\,\|X\|^k < \infty$.*

(i) *Then $\mathbb{E}\,X^{\hat{\otimes}k}$ exists in Pettis sense.*
(ii) *$\mathbb{E}\,X^{\hat{\otimes}k}$ exists in Bochner sense $\iff$ there is a countable set $N \subset [0,1]$ such that $\operatorname{supp}(\Delta X) \subseteq N$ a.s.*

Proof. (i): Let $\alpha \in L(D^k, \mathbb{R})$ be a k-linear form. Then $\langle \alpha, X^{\otimes k}\rangle$ is measurable by Corollary 15.9, and

$$\bigl|\langle \alpha, X^{\otimes k}\rangle\bigr| \leqslant \|\alpha\|\|X\|^k.$$

It follows that the family $\bigl\{\langle \alpha, X^{\hat{\otimes}k}\rangle : \|\alpha\| \leqslant 1\bigr\}$ is uniformly integrable. Moreover, $X^{\hat{\otimes}k}$ is weakly a.s. separably valued by Theorem 15.14. Hence Theorem 5.8 shows, using Theorem 4.6, that $\mathbb{E}\,X^{\hat{\otimes}k}$ exists in Pettis sense.

(ii): This is another special case of Theorem 6.7. □

For the second moment, we can as for $C(K)$ use Grothendieck's theorem Theorem 11.20 to show that the conditions for the injective moment in Theorem 16.6 also imply the existence of the projective second moment, thus improving Theorem 16.10 when $k = 2$. Example 11.27 shows that this does not extend to $k \geqslant 3$.

Theorem 16.11. *Let X be a $\mathcal{D}$-measurable $D[0,1]$-valued random variable.*

(i) *$\mathbb{E}\,X^{\hat{\otimes}2}$ exists in Dunford sense $\iff$ $\mathbb{E}\,X^{\check{\otimes}2}$ exists in Dunford sense $\iff$ the weak second moment exists $\iff$ $\sup_{t\in[0,1]} \mathbb{E}\,|X(t)|^2 < \infty$.*

(ii) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Pettis sense* $\iff$ $\mathbb{E}\, X^{\check{\otimes}2}$ *exists in Pettis sense* $\iff$ *the family* $\{|X(t)|^2 : t \in [0,1]\}$ *of random variables is uniformly integrable.*

(iii) $\mathbb{E}\, X^{\hat{\otimes}2}$ *exists in Bochner sense* $\iff$ $\mathbb{E}\, X^{\check{\otimes}2}$ *exists in Bochner sense* $\iff$ $\mathbb{E}\big(\sup_{t\in[0,1]} |X(t)|\big)^2 < \infty$ *and there is a countable set* $N \subset [0,1]$ *such that* $\operatorname{supp}(\Delta X) \subseteq N$ *a.s.*

Proof. The forward implications follow directly, as in the proof of Theorem 11.22, using also Theorem 15.11.

(i): Let α be a bounded bilinear form on $D[0,1] = C(\widehat{I})$. By Theorem 11.20, α extends to a bounded bilinear form on $L^2(\widehat{I}, \nu)$ for some probability measure ν on $\widehat{I}$; more precisely, (11.10) yields

$$\big|\langle \alpha, X^{\otimes 2}\rangle\big| \leqslant 2k_G\|\alpha\| \int_{\widehat{I}} |X(t)|^2 \,\mathrm{d}\nu(t). \tag{16.8}$$

This is, apart from a constant, the same estimate as (16.2) (proved for integral forms), and the same argument as in the proof of Theorem 16.6 yields, cf. (16.6),

$$\mathbb{E}\big|\langle \alpha, X^{\otimes 2}\rangle\big| \leqslant 8k_G\|\alpha\| \sup_{t\in[0,1]} \mathbb{E}\, |X(t)|^2. \tag{16.9}$$

Furthermore, $\langle \alpha, X^{\otimes 2}\rangle$ is measurable by Corollary 15.9. It follows that if $\sup_{t\in[0,1]} \mathbb{E}\, |X(t)|^2 < \infty$, then $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Dunford sense.

(ii): Assume that the family $\{|X(t)|^2 : t \in K\}$ is uniformly integrable. By applying (16.9) to $\mathbf{1}_E X$ as in the proof of Theorem 16.6, we obtain from (5.4) that the family $\big\{\alpha(X,X) : \alpha \in L(D[0,1]^2;\mathbb{R}),\ \|\alpha\| \leqslant 1\big\}$ is uniformly integrable. Moreover, $X^{\hat{\otimes}2}$ is weakly a.s. separably valued by Theorem 15.14. Hence Theorem 5.8 shows, using Theorem 4.6, that $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Pettis sense.

(iii): This is again a special case of Theorem 6.7. □

Theorem 16.12. *Let* X *and* Y *be* $\mathcal{D}$*-measurable* $D[0,1]$*-valued random variables such that* $\sup_{t\in[0,1]} |X(t)|^2 < \infty$ *and* $\sup_{t\in[0,1]} |Y(t)|^2 < \infty$*. Then the following are equivalent.*

(i) $\mathbb{E}\,\alpha(X,X) = \mathbb{E}\,\alpha(Y,Y)$ *for every bounded bilinear form* α *on* $D[0,1]$*.*

(ii) $\mathbb{E}\big(x_1^*(X)x_2^*(X)\big) = \mathbb{E}\big(x_1^*(Y)x_2^*(Y)\big)$*, for any* $x_1^*, x_2^* \in D[0,1]^*$*.*

(iii) $\mathbb{E}\big(X(t_1)X(t_2)\big) = \mathbb{E}\big(Y(t_1)Y(t_2)\big)$ *for any* $t_1, t_2 \in [0,1]$*.*

(iv) $\mathbb{E}\, X^{\hat{\otimes}2} = \mathbb{E}\, Y^{\hat{\otimes}2}$ *in* $D[0,1]^{\hat{\otimes}2}$*, with the moments existing in Dunford sense.*

(v) $\mathbb{E}\, X^{\check{\otimes}2} = \mathbb{E}\, Y^{\check{\otimes}2}$ *in* $D[0,1]^{\check{\otimes}2}$*, with the moments existing in Dunford sense.*

Proof. We argue as in the proof of Theorem 11.23, with some changes. The implications (i) $\implies$ (ii) $\implies$ (iii) and (iv) $\implies$ (v) are trivial, the equivalence (i) $\iff$ (iv) is Corollary 6.17 and (v) $\implies$ (ii) follows by (6.3). It remains to show that (iii) $\implies$ (i).

Thus, let $\alpha \in L(D[0,1]^2;\mathbb{R}) = L(C(\widehat{I})^2;\mathbb{R})$. By Theorem 11.20 there exists a probability measure ν on $\widehat{I}$ such that α extends to $L^2(\widehat{I},\nu)$. The main difference from Theorem 11.23 is that we cannot assert that $X(t,\omega)$ is jointly measurable on $\widehat{I}\times\Omega$, see Remark 15.7. It is, however, still possible to regard X as a map into $L^2(\widehat{I},\nu)$. We prefer to state this slightly differently, returning to [0,1] by the argument already used in the proof of Theorem 16.6. We thus regard ν as a continuous linear functional χ on $C(\widehat{I}) = D[0,1]$; we now use the representation (13.7) in Remark 13.3 and write it as

$$\int_{\widehat{I}} f\,\mathrm{d}\nu = \chi(f) = \int_0^1 f(t)\,\mathrm{d}\mu_1(t) + \int_0^1 f(t-)\,\mathrm{d}\mu_2(t), \tag{16.10}$$

where μ_2 is a discrete measure supported on a countable set $\{t_n\}_{n=1}^\infty$. It is easily seen that both μ_1 and μ_2 are positive measures. (We may alternatively omit this verification and replace them by $|\mu_1|$ and $|\mu_2|$, possibly increasing ν.)

Let $\Sigma := [0,1]\cup\{t_n-\}_{n=1}^\infty \subset \widehat{I}$. We define a σ-field $\mathcal{A}$ on Σ by $\mathcal{A} := \{A\subseteq \Sigma : A\cap[0,1]\in\mathcal{B}([0,1])\}$, and let μ be the measure $\mu_1+\mu_2'$ on $(\Sigma,\mathcal{A})$, where μ_2' is the measure on $\{t_n-\}_{n=1}^\infty$ given by $\mu_2'\{t_n-\} := \mu_2\{t_n\}$. Then (16.10) can be written

$$\int_{\widehat{I}} f\,\mathrm{d}\nu = \int_\Sigma f\,\mathrm{d}\mu. \tag{16.11}$$

Applying this to $|f|^2$, we see that $\|f\|_{L^2(\widehat{I},\nu)} = \|f\|_{L^2(\Sigma,\mu)}$ for all $f\in D[0,1]$, and α thus extends to a bounded bilinear form on $L^2(\Sigma,\mu)$. Moreover, Lemma 15.1 implies that $X(t,\omega)$ is jointly measurable on $\Sigma\times\Omega$. Consequently, X is a Bochner measurable random variable in $L^2(\Sigma,\mu)$ by Lemma 10.1. Furthermore, the assumption that $\sup_t \mathbb{E}\,|X(t)|^2<\infty$ implies by Fubini's theorem, as in (11.12), that $\mathbb{E}\,\|X\|^2_{L^2(\Sigma,\mu)}<\infty$, and thus $\mathbb{E}\,X^{\hat\otimes 2}$ exists in $L^2(\Sigma,\mu)^{\hat\otimes 2}$ in Bochner sense. The same holds for $\mathbb{E}\,Y^{\hat\otimes 2}$.

The proof is now completed as for Theorem 11.23, *mutatis mutandis*. □

Theorem 16.13. *Let X and Y be $\mathcal{D}$-measurable $D[0,1]$-valued random variables, and suppose that either*

(i) $\mathbb{E}\,\|X\|^k<\infty$ *and* $\mathbb{E}\,\|Y\|^k<\infty$, *or*
(ii) $k=2$ *and* $\sup\{\mathbb{E}\,|X(t)|^2 : t\in[0,1]\}<\infty$, $\sup\{\mathbb{E}\,|Y(t)|^2 : t\in[0,1]\}<\infty$.

Then (1.2) *is equivalent to* (1.3), *and further to*

$$\mathbb{E}\big(X(t_1)\cdots X(t_k)\big) = \mathbb{E}\big(Y(t_1)\cdots Y(t_k)\big), \qquad t_1,\dots,t_k\in[0,1]. \tag{16.12}$$

Proof. (i): By Theorem 16.10, $\mathbb{E}\,X^{\hat\otimes k}$ and $\mathbb{E}\,Y^{\hat\otimes k}$ exist in Pettis sense. Thus, $\mathbb{E}\,X^{\check\otimes k}$ and $\mathbb{E}\,Y^{\check\otimes k}$ too exist in Pettis sense, and the result follows by Theorem 8.3, Corollaries 6.17 and 6.14 and (16.1), similarly as in the proofs of Theorems 1.1 and 1.3.

(ii): A simplified version of Theorem 16.12. □

Remark 16.14. As for Theorem 11.23, we do not know whether the condition (i) can be weakened for $k \geqslant 3$.

17. Uniqueness

In the previous sections we have considered the k:th moment(s) for a fixed k. In this section and the next, we consider the sequence of all moments. In the present section we show that there are analogues of the classical results for real-valued random variables that the moments (under certain conditions) determine the distribution. In Section 18 we consider convergence, where the situation is more complicated and less satisfactory.

We suppose for simplicity that the Banach space B is separable. All random variables are tacitly assumed to be (Borel) measurable, cf. Theorem 3.3.

We begin with two simple results on the existence of all moments.

Theorem 17.1. *Let X be a B-valued random variable, where B is a separable Banach space. Then the following are equivalent.*

(i) $\mathbb{E}\,\|X\|^k < \infty$ *for every* $k \geqslant 1$,
(ii) *The projective moment* $\mathbb{E}\,X^{\hat{\otimes}k}$ *exists in Bochner sense for every* $k \geqslant 1$.
(iii) *The injective moment* $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Bochner sense for every* $k \geqslant 1$.

Proof. An immediate consequence of Theorem 6.7. □

We do not know any general weaker criterion for the existence of projective moments in Pettis or Dunford sense. (See Theorem 9.14 for one case where no weaker criterion exists, and Theorem 11.22 for a case when it does.) For injective moments we have the following.

Theorem 17.2. *Let X be a B-valued random variable, where B is a separable Banach space. Then the following are equivalent.*

(i) $\mathbb{E}\,|\langle x^*, X\rangle|^k < \infty$ *for every* $k \geqslant 1$,
(ii) *The injective moment* $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Dunford sense for every* $k \geqslant 1$.
(iii) *The injective moment* $\mathbb{E}\,X^{\check{\otimes}k}$ *exists in Pettis sense for every* $k \geqslant 1$.

Proof. (i) $\iff$ (ii): By Theorem 6.10(ii)(a).

(ii) $\iff$ (iii): If (ii) holds, then $\sup\{\mathbb{E}\,|\langle x^*, X\rangle|^k : \|x^*\| \leqslant 1\} < \infty$ for every $k \geqslant 1$ by Theorem 6.10(i). Thus $\sup\{\mathbb{E}\,|\langle x^*, X\rangle|^{k+1} : \|x^*\| \leqslant 1\} < \infty$, which implies that $\{|\langle x^*, X\rangle|^k : \|x^*\| \leqslant 1\}$ is uniformly integrable. Theorem 6.19(ii)(a) shows that (iii) holds. The converse is obvious. □

It is well-known that already on $\mathbb{R}$, there are random variables with the same moments but different distributions, see e.g. [34, Section 4.10]. A well-known sufficient condition for the distribution of X to be uniquely determined by the moments is the *Carleman condition* [10]

$$\sum_{n=1}^{\infty} \left(\mathbb{E}\,|X|^{2n}\right)^{-1/2n} = \infty. \tag{17.1}$$

Note that (17.1) is satisfied whenever $\mathbb{E}\,e^{t|X|} < \infty$ for some $t > 0$.

Theorem 17.3. *Let B be a separable Banach space and let X and Y be two B-valued random variables. Suppose that $\mathbb{E}\,\|X\|^k < \infty$ and $\mathbb{E}\,\|Y\|^k < \infty$ for every $k \geqslant 1$ and that $\|X\|$ satisfies the Carleman condition* (17.1)*. Then the following are equivalent.*

(i) $X \overset{\mathrm{d}}{=} Y$.
(ii) $\mathbb{E}\, X^{\hat{\otimes} k} = \mathbb{E}\, Y^{\hat{\otimes} k}$ *for every* $k \geqslant 1$.
(iii) $\mathbb{E}\, X^{\check{\otimes} k} = \mathbb{E}\, Y^{\check{\otimes} k}$ *for every* $k \geqslant 1$.
(iv) $\langle x^*, X\rangle \overset{\mathrm{d}}{=} \langle x^*, Y\rangle$ *for every* $x^* \in B^*$.

Proof. Note that the moments exist in Bochner sense by Theorem 17.1.

(i) $\implies$ (ii) is obvious and (ii) $\implies$ (iii) follows by Theorem 6.2.

(iii) $\implies$ (iv): If $x^* \in B^*$, then

$$\mathbb{E}\langle x^*, X\rangle^k = \langle \mathbb{E}\, X^{\check{\otimes} k}, (x^*)^{\otimes k}\rangle = \langle \mathbb{E}\, Y^{\check{\otimes} k}, (x^*)^{\otimes k}\rangle = \mathbb{E}\langle x^*, Y\rangle^k$$

for every $k \geqslant 1$ by (6.3) and (iii). Since $\mathbb{E}\,|\langle x^*, X\rangle|^k \leqslant \|x^*\|^k\, \mathbb{E}\,\|X\|^k$, the Carleman condition (17.1) holds for $\langle x^*, X\rangle$ too, and it follows that (iv) holds.

(iv) $\implies$ (i): This is well-known [42]. (Sketch of proof: Any finite linear combination of elements of B^* is another element of B^*. Hence (iv) implies, by the Cramér–Wold device, that (iv) holds jointly for any finite number of functionals x^*. A standard application of the monotone class theorem shows that $\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$ for every $A \in \mathcal{B}_w = \mathcal{B}$.) □

Remark 17.4. The proof of Theorem 17.3 shows that it suffices that the Carleman condition holds for each $\langle x^*, X\rangle$. Moreover, if we only consider injective moments, and the equivalences (i) $\iff$ (iii) $\iff$ (iv), then the moment assumptions may be weakened to $\mathbb{E}\,|\langle x^*, X\rangle|^k < \infty$ and $\mathbb{E}\,|\langle x^*, Y\rangle|^k < \infty$ for every $k \geqslant 1$, with the injective moments existing in Pettis sense by Theorem 17.2.

Remark 17.5. The assumption that B is separable is essential; the equivalence (i) $\iff$ (iv) is not true in general for non-separable B. Example 7.4 gives an X in $\ell^2[0,1]$ such that $\langle x^*, X\rangle = 0$ a.s. for every $x^* \in \ell^2[0,1]^* = \ell^2[0,1]$, and thus $\langle x^*, X\rangle \overset{\mathrm{d}}{=} \langle x^*, Y\rangle$ with $Y = 0$, although $X \neq 0$ a.s.

18. Convergence

As in the preceding section, we assume that the Banach space B is separable, and that all random variables are (Borel) measurable.

We consider a sequence X_n, $n \geqslant 1$ of B-valued random variables, and a potential limit X. For definition and general properties of convergence in distribution, denoted $X_n \xrightarrow{\mathrm{d}} X$, see [6]. In particular, recall that convergence in distribution can be described by a metric, at least when B is separable as here, see [6, Theorem 6.8]. (The non-separable case is more complicated and related to the existence of real-measurable cardinals, cf. Remark 9.13, see [5, Appendix III].)

As in the real-valued case, convergence in distribution implies convergence of the moments, provided some suitable integrability condition holds uniformly. We state one simple such result.

Theorem 18.1. *Suppose that X and X_n, $n \geqslant 1$, are B-valued random variables, where B is a separable Banach space. Suppose that $X_n \xrightarrow{\mathrm{d}} X$, and that $\sup_n \mathbb{E}\,\|X_n\|^k < \infty$ for every $k \geqslant 1$. Then $\mathbb{E}\,X_n^{\hat{\otimes}k} \to \mathbb{E}\,X^{\hat{\otimes}k}$ in $B^{\hat{\otimes}k}$ and $\mathbb{E}\,X_n^{\check{\otimes}k} \to \mathbb{E}\,X^{\check{\otimes}k}$ in $B^{\check{\otimes}k}$ as $n \to \infty$, for every $k \geqslant 1$.*

Proof. By Theorem 17.1, all moments $\mathbb{E}\,X_n^{\hat{\otimes}k}$ and $\mathbb{E}\,X_n^{\check{\otimes}k}$ exist. Furthermore, $\|X_n\| \xrightarrow{\mathrm{d}} \|X\|$, and thus by Fatou's lemma (for convergence in distribution [34, Theorem 5.5.8]), $\mathbb{E}\,\|X\|^k \leqslant \liminf_{n\to\infty} \mathbb{E}\,\|X_n\|^k < \infty$. Hence all moments $\mathbb{E}\,X^{\hat{\otimes}k}$ and $\mathbb{E}\,X^{\check{\otimes}k}$ also exist.

By the Skorohod representation theorem [6, Theorem 6.7], we may assume that $X_n \xrightarrow{\text{a.s.}} X$. Then $X_n^{\hat{\otimes}k} \xrightarrow{\text{a.s.}} X^{\hat{\otimes}k}$, since the (non-linear) mapping $x \mapsto x^{\otimes k}$ is continuous $B \to B^{\hat{\otimes}k}$. Thus, $\|X_n^{\hat{\otimes}k} - X^{\hat{\otimes}k}\| \xrightarrow{\text{a.s.}} 0$. Furthermore,

$$\|X_n^{\hat{\otimes}k} - X^{\hat{\otimes}k}\| \leqslant \|X_n^{\hat{\otimes}k}\| + \|X^{\hat{\otimes}k}\| = \|X_n\|^k + \|X\|^k, \tag{18.1}$$

and since $\sup_n \mathbb{E}\,\|X_n\|^{2k} < \infty$ and $\mathbb{E}\,\|X\|^k < \infty$, it follows that the family $\{\|X_n^{\hat{\otimes}k} - X^{\hat{\otimes}k}\| : n \geqslant 1\}$ is uniformly integrable (for any fixed k), see e.g. [34, Theorems 5.4.2–4.6]; hence $\mathbb{E}\,\|X_n^{\hat{\otimes}k} - X^{\hat{\otimes}k}\| \to 0$. Consequently,

$$\|\,\mathbb{E}\,X_n^{\hat{\otimes}k} - \mathbb{E}\,X^{\hat{\otimes}k}\| = \|\,\mathbb{E}(X_n^{\hat{\otimes}k} - X^{\hat{\otimes}k})\| \leqslant \mathbb{E}\,\|X_n^{\hat{\otimes}k} - X^{\hat{\otimes}k}\| \to 0,$$

as $n \to \infty$. This proves $\mathbb{E}\,X_n^{\hat{\otimes}k} \to \mathbb{E}\,X^{\hat{\otimes}k}$, and $\mathbb{E}\,X_n^{\check{\otimes}k} \to \mathbb{E}\,X^{\check{\otimes}k}$ follows similarly, or by Theorem 6.2. □

If B is finite-dimensional, then the converse to Theorem 18.1 holds, provided the moments determine the distribution of X (for example by the Theorem 17.3); this is the standard method of moments (in several variables).

In infinite dimensions, there is in general no converse. We begin with a simple example showing that convergence of the injective moments does not imply convergence in distribution. Moreover, the example shows that convergence of the injective moments does not imply convergence of the projective moments. (The converse implication is trivial by Theorem 6.2.)

Example 18.2. Regard $\mathbb{R}^n$ as a subspace of ℓ^2 by the isometric embedding $(a_1, \dots, a_n) \mapsto (a_1, \dots, a_n, 0, 0, \dots)$. Let $X_n := n^{-1/2}(\xi_1, \dots, \xi_n) \in \mathbb{R}^n \subset \ell^2$, where $\xi_1, \xi_2, \dots \sim N(0,1)$ are i.i.d. standard normal random variables.

Note first that $\|X_n\|^2 = \frac{1}{n}\sum_{i=1}^n \xi_i^2 \to 1$ a.s. as $n \to \infty$ by the law of large numbers. Thus X_n does not tend to 0 in distribution.

Next, consider the injective moment $\mathbb{E}\,X_n^{\check{\otimes}k}$. (It does not matter whether we regard $\mathbb{E}\,X_n^{\check{\otimes}k}$ as an element of $(\mathbb{R}^n)^{\check{\otimes}k}$ or $(\ell^2)^{\check{\otimes}k}$, since $(\mathbb{R}^n)^{\check{\otimes}k}$ is isometrically a subspace of $(\ell^2)^{\check{\otimes}k}$, see Remark 4.4.)

If $y = (a_1, \dots, a_k) \in \mathbb{R}^n$, then $\langle y, X\rangle = \sum_{i=1}^n a_i n^{-1/2}\xi_i$ is normal with mean 0 and variance $\sum_{i=1}^n a_i^2 n^{-1}\,\mathbb{E}\,\xi_i^2 = \|y\|^2/n$, i.e. $\langle y, X\rangle \sim N(0, \|y\|^2/n)$. Hence, for any $y_1, \dots, y_k \in \mathbb{R}^n$, by Hölder's inequality,

$$\begin{aligned}|\langle \mathbb{E}\, X_n^{\check\otimes k}, y_1 \otimes \cdots \otimes y_k\rangle| &= \bigl|\mathbb{E}\bigl(\langle X, y_1\rangle \cdots \langle X, y_k\rangle\bigr)\bigr| \\ &\leqslant \prod_{i=1}^k (\mathbb{E}\,|\langle X, y_i\rangle|^k)^{1/k} \leqslant C_k \prod_{i=1}^k \frac{\|y_i\|}{\sqrt{n}}\end{aligned}$$

for some constant $C_k := \mathbb{E}\,|\xi_1|^k$. Since $B^{\check\otimes k}$ can be seen (isometrically) as a subspace of $L(B^k; \mathbb{R})$, see Section 4, it follows that

$$\|\,\mathbb{E}\, X_n^{\check\otimes k}\| \leqslant C_k n^{-k/2}. \tag{18.2}$$

In particular, the injective moments $\mathbb{E}\, X_n^{\check\otimes k} \to 0$ as $n \to \infty$ for every $k \geqslant 1$.

Finally, the second moment $\mathbb{E}\, X_n^{\otimes 2}$ is given by the covariance matrix,

$$\mathbb{E}\, X_n^{\otimes 2} = \Bigl(\frac{1}{n}\,\mathbb{E}\,\xi_i\xi_j\Bigr)_{i,j=1}^n = \frac{1}{n} I, \tag{18.3}$$

where I is the identity matrix. Regarded as an operator, I is the identity operator in $\mathbb{R}^n$, which has trace norm $\|I\|_{\mathcal{N}(\mathbb{R}^n)} = n$. Since the projective tensor norm equals the trace norm (for Hilbert spaces), see Theorem 9.2, we obtain $\|\,\mathbb{E}\, X_n^{\hat\otimes 2}\| = \|\frac{1}{n} I\|_{\mathcal{N}(\mathbb{R}^n)} = 1$. (Since there exists a projection $\ell^2 \to \mathbb{R}^n$ of norm 1, it does not matter whether we regard $\mathbb{E}\, X_n^{\hat\otimes 2}$ as an element of $(\mathbb{R}^n)^{\hat\otimes 2}$ or $(\ell^2)^{\hat\otimes 2}$.) Hence the projective moments $\mathbb{E}\, X_n^{\hat\otimes 2}$ do not tend to 0.

In fact, this extends to every infinite-dimensional Banach space.

Theorem 18.3. *Let B be any infinite-dimensional Banach space. Then there exists a sequence of Bochner measurable random variables X_n in B such that the injective moments $\mathbb{E}\, X_n^{\check\otimes k} \to 0$ as $n \to \infty$ for every $k \geqslant 1$ but X_n does not tend to 0; in fact, $\|X_n\| \xrightarrow{\text{p}} 1$.*

Proof. Let $\varepsilon_n \to 0$ (for example, $\varepsilon_n := 1/n$). By Dvoretzky's theorem, see e.g. [42] or [53], for every n, there is a n-dimensional subspace B_n of B such that B_n is isomorphic to $\mathbb{R}^n$ by an isomorphism $T_n : \mathbb{R}^n \to B_n$ with $\|T_n\|, \|T_n^{-1}\| \leqslant 1 + \varepsilon_n$.

Let $X_n' \in \mathbb{R}^n$ be as X_n in Example 18.2, and let $X_n := T_n X_n' \in B$. □

Remark 18.4. It is easy to see that we may replace X_n in Example 18.2 or Theorem 18.3 by $X_n/\|X_n\|$, thus obtaining $\|X_n\| = 1$ a.s. and $\mathbb{E}\, X_n^{\check\otimes k} \to 0$ for every $k \geqslant 1$. (In Example 18.2, this means taking X_n uniformly distributed on the unit sphere of $\mathbb{R}^n$.)

For projective moments, the situation is more complicated. We next give another example (with $B = c_0$), showing that also convergence of the projective moments does not imply convergence in distribution in general. On the other hand, we then show that in a Hilbert space, it does (assuming a Carleman condition). Moreover, we shall show that in a Hilbert space,

even *weak* convergence (denoted $\xrightarrow{\mathrm{w}}$) of the projective moments suffices to imply convergence in distribution.

Example 18.5. Let $B = c_0$ and let X_n be the B-valued random variable given by $\mathbb{P}(X_n = e_i) = 1/n$, $i = 1, \dots, n$. Then $\|X_n\| = 1$.

If α is a k-linear form on c_0, then by (12.4), using the notation (12.3) and the Cauchy–Schwarz inequality,

$$|\langle \alpha, \mathbb{E}\, X^{\hat{\otimes} k}\rangle| = |\,\mathbb{E}\langle \alpha, X^{\hat{\otimes} k}\rangle| = \left|\frac{1}{n}\sum_{i=1}^{n} \alpha(e_i, \dots, e_i)\right| \leqslant \Big(\frac{1}{n}\sum_{i=1}^{n} |a(i,\dots,i)|^2\Big)^{1/2}$$
$$\leqslant \Big(\frac{1}{n}\sum_{s_1,\dots,s_k \in \mathbb{N}} |a(s_1,\dots,s_k)|^2\Big)^{1/2} \leqslant n^{-1/2}\|\alpha\|.$$

Hence, using Theorem 4.6,

$$\|X_n^{\hat{\otimes} k}\| \leqslant n^{-1/2} \to 0 \qquad \text{as } n \to \infty.$$

Consequently, $\mathbb{E}\, X_n^{\hat{\otimes} k} \to 0$ for every k, but $X_n \xrightarrow{\mathrm{p}} 0$.

Theorem 18.6. *Suppose that X and X_n, $n \geqslant 1$, are H-valued random variables, where H is a separable Hilbert space. Suppose further that $\|X\|$ satisfies the Carleman condition* (17.1)*. If $\mathbb{E}\, X_n^{\hat{\otimes} k} \xrightarrow{\mathrm{w}} \mathbb{E}\, X^{\hat{\otimes} k}$ in $B^{\hat{\otimes} k}$ for every $k \geqslant 1$, with the moments existing in Pettis sense, then $X_n \xrightarrow{\mathrm{d}} X$.*

Proof. For any fixed $y \in H$, and any $m \geqslant 1$,

$$\|x - y\|^{2m} = \langle x - y, x - y\rangle^m = \sum_{k=0}^{2m} \alpha_k(x, \dots, x),$$

where α_k is some bounded k-linear form (depending on y and m). Hence, if $\mathbb{E}\, X_n^{\hat{\otimes} k} \xrightarrow{\mathrm{w}} \mathbb{E}\, X^{\hat{\otimes} k}$ for every k, then

$$\mathbb{E}\,\|X_n - y\|^{2m} = \sum_{k=0}^{2m} \mathbb{E}\langle \alpha_k, X_n^{\otimes k}\rangle = \sum_{k=0}^{2m} \langle \alpha_k, \mathbb{E}\, X_n^{\otimes k}\rangle$$
$$\to \sum_{k=0}^{2m} \langle \alpha_k, \mathbb{E}\, X^{\otimes k}\rangle = \mathbb{E}\,\|X - y\|^{2m}. \tag{18.4}$$

Since $\|X\|$ satisfies the Carleman condition, it is straightforward to show that $\|X - y\| \leqslant \|X\| + \|y\|$ does too. It thus follows from (18.4), by the method of moments, that

$$\|X_n - y\|^2 \xrightarrow{\mathrm{d}} \|X - y\|^2. \tag{18.5}$$

(We use here the fact that for positive random variables, the Carleman condition can be relaxed to (17.1) for the square root, see [34, (4.10.2)]. Alternatively, we can introduce random signs and apply the method of moments to show that $\pm\|X_n - y\| \xrightarrow{\mathrm{d}} \pm\|X - y\|$, where all odd moments vanish and thus converge trivially.)

The argument extends to any linear combination of $\|X_n-y_1\|^2,\dots,\|X_n-y_\ell\|^2$ for any given $y_1,\dots,y_\ell\in H$; hence (18.5) holds with joint convergence for any finite set $y_1,\dots,y_\ell\in H$.

Consequently, if $A\subseteq B$ is a finite intersection of open balls $B(y_i,r_i):=\{x:\|x-y_i\|<r_i\}$ such that $\mathbb{P}(X\in\partial B(y_i,r_i))=0$, then $\mathbb{P}(X_n\in A)\to\mathbb{P}(X\in A)$; this implies $X_n\xrightarrow{\mathrm{d}}X$ by [6, Theorem 2.4]. □

Remark 18.7. The argument extends to the spaces $L^p(\mu)$ (assumed to be separable) provided p is an even integer (and $\|X\|^{p/2}$ satisfies the Carleman condition). We do not know whether there are further Banach spaces such that $\mathbb{E}\,X_n^{\hat\otimes k}\to\mathbb{E}\,X^{\hat\otimes k}$ for every k implies $X_n\xrightarrow{\mathrm{d}}X$ (provided X is bounded, say, for simplicity).

Specialising to $X=0$, we have the related problem: For which Banach spaces B does $\mathbb{E}\,X_n^{\hat\otimes k}\to0$ in $B^{\hat\otimes k}$ for every k imply $X_n\xrightarrow{\mathrm{p}}0$?

Next we show that *weak* convergence of the injective moments is equivalent to *weak convergence in distribution*, meaning convergence in distribution of $\langle x^*,X_n\rangle$ for every $x^*\in B^*$.

Theorem 18.8. *Suppose that X and X_n, $n\geqslant1$, are B-valued random variables, where B is a separable Banach space. Suppose further that*

$$\sup_{n\geqslant1}\mathbb{E}\,|\langle x^*,X_n\rangle|^k<\infty \tag{18.6}$$

for every $x^*\in B^*$ and every $k\geqslant1$, and that every $\langle x^*,X\rangle$, $x^*\in B^*$, satisfies the Carleman condition (17.1)*. Then the following are equivalent.*

(i) $\mathbb{E}\,X_n^{\check\otimes k}\xrightarrow{\mathrm{w}}\mathbb{E}\,X^{\check\otimes k}$ *in* $B^{\check\otimes k}$ *for every* $k\geqslant1$.

(ii) $\mathbb{E}\bigl(x_1^*(X_n)\cdots x_k^*(X_n)\bigr)\to\mathbb{E}\bigl(x_1^*(X)\cdots x_k^*(X)\bigr)$ *for every* $k\geqslant1$ *and* $x_1^*,\dots,x_k^*\in B^*$. *(In other words, the weak moments converge pointwise.)*

(iii) $\mathbb{E}\bigl(x^*(X_n)^k\bigr)\to\mathbb{E}\bigl(x^*(X)^k\bigr)$ *for every* $k\geqslant1$ *and* $x^*\in B^*$.

(iv) $x^*(X_n)\xrightarrow{\mathrm{d}}x^*(X)$ *for every* $x^*\in B^*$.

(v) $x^*(X_n)\xrightarrow{\mathrm{d}}x^*(X)$ *jointly for all* $x^*\in B^*$.

Note that all (injective) moments exist in Pettis sense by Theorem 17.2 together with (18.6) and the Carleman condition for $\langle x^*,X\rangle$ (which implies that $\mathbb{E}\,|\langle x^*,X\rangle|^k<\infty$ for every k).

Proof. By assumption, for every $k\geqslant1$, $T:x^*\mapsto(\langle x^*,X_1\rangle,\langle x^*,X_2\rangle,\dots)$ maps B^* into $\ell^\infty(L^k(\mathbb{P}))$. By the closed graph theorem, T is bounded, and thus

$$C_k:=\sup\bigl\{\mathbb{E}\,|\langle x^*,X_n\rangle|^k:\|x^*\|\leqslant1,\ n\geqslant1\bigr\}<\infty. \tag{18.7}$$

Note that, by definition, (i) holds if and only if

$$\langle\chi,\mathbb{E}\,X_n^{\check\otimes k}\rangle=\mathbb{E}\langle\chi,X_n^{\check\otimes k}\rangle\to\mathbb{E}\langle\chi,X^{\check\otimes k}\rangle=\langle\chi,\mathbb{E}\,X^{\check\otimes k}\rangle \tag{18.8}$$

for every $\chi\in(B^{\check\otimes k})^*$ and every $k\geqslant1$.

(i) $\implies$ (ii): Choosing $\chi = x_1^* \otimes \cdots \otimes x_k^*$ in (18.8), (ii) follows by (6.3).

(ii) $\implies$ (i): Let $\chi \in (B^{\check{\otimes} k})^*$. By Theorem 4.7, χ has a representation (4.10), and as shown in the proof of Theorem 6.10, (6.5) holds, together with the corresponding formula with X replaced by X_n. Thus, (18.8) can be written

$$\int_{K^k} \mathbb{E}\bigl(x_1^*(X_n)\cdots x_k^*(X_n)\bigr)\,\mathrm{d}\mu(x_1^*,\dots,x_k^*) \to \int_{K^k} \mathbb{E}\bigl(x_1^*(X)\cdots x_k^*(X)\bigr)\,\mathrm{d}\mu(x_1^*,\dots,x_k^*). \quad (18.9)$$

The integrand converges pointwise by (ii); furthermore, by Hölder's inequality and (18.7), $\bigl|\mathbb{E}\bigl(x_1^*(X_n)\cdots x_k^*(X_n)\bigr)\bigr| \leqslant C_k$. Consequently, (18.9) holds by dominated convergence.

(ii) $\implies$ (iii): A special case, obtained by taking $x_1^* = \cdots = x_k^* = x^*$.

(iii) $\implies$ (ii): A standard polarisation argument. Given $x_1^*,\dots,x_k^*$, use (iii) with $x^* := t_1x_1^* + \cdots + t_kx_k^*$, where $t_1,\dots,t_k$ are real numbers. Then both sides of (iii) are (homogeneous) polynomials in $t_1,\dots,t_k$, and since the left side converges to the right for every $t_1,\dots,t_k$, the coefficient of $t_1\cdots t_k$ converges too, which yields (ii) (after dividing by $k!$).

(iii) $\implies$ (iv): This is the usual method of moments for the real-valued random variables $\langle x^*, X_n\rangle$, using the Carleman condition.

(iv) $\implies$ (iii): For any fixed k, the random variables $\langle x^*, X_n\rangle^k$ are uniformly integrable, by (18.6) with $k+1$. Hence (iv) $\implies$ (iii).

(iv) $\iff$ (v): The joint convergence in (v) means, by definition, joint convergence for any finite set $x_1^*,\dots,x_k^* \in B^*$. Since any linear combination of $x_1^*,\dots,x_k^*$ is another element of B^*, this follows from (iv) by the Cramér–Wold device. (Cf. the proof of Theorem 17.3.) □

Remark 18.9. If all (injective) moments exist in Pettis sense and Theorem 18.8(i) holds, then (iii) holds by (6.3), see the proof above, and thus (18.6) holds for every even k, and thus for every k, so (18.6) is redundant in this case.

Example 18.10. Take $B = \ell^1$ and let X_n be as in Theorem 18.3 and Remark 18.4. Then the injective moments $\mathbb{E}\, X_n^{\check{\otimes} k} \to 0$, and Theorem 18.8 shows that $\langle x^*, X_n\rangle \xrightarrow{\mathrm{d}} 0$ as $n\to\infty$ for every $x^* \in (\ell^1)^*$; equivalently, $\langle x^*, X_n\rangle \xrightarrow{\mathrm{p}} 0$ as $n\to\infty$. However, $\|X_n\| = 1$ and thus $X_n \xrightarrow{\mathrm{p}} 0$.

By Schur's theorem [17, p. 85], a sequence in ℓ^1 converges weakly if and only if it converges strongly (i.e., in norm). We see that this does not extend to convergence in probability (or distribution) for sequences of random variables in ℓ^1. This also shows that Skorohod's representation theorem does not hold for weak convergence in distribution: there is no way to couple the random variables X_n such that $X_n \xrightarrow{\mathrm{w}} 0$ a.s., since this would imply $X_n \to 0$ a.s. by Schur's theorem, and thus $X_n \xrightarrow{\mathrm{p}} 0$.

If we know tightness by other means, weak convergence in distribution is equivalent to convergence in distribution.

Corollary 18.11. *Suppose that X and X_n, $n \geqslant 1$, are B-valued random variables, where B is a separable Banach space. Suppose further that all injective moments $\mathbb{E}\, X_n^{\check{\otimes} k}$ exist in Pettis sense, that $\mathbb{E}\, X_n^{\check{\otimes} k} \xrightarrow{\mathrm{w}} \mathbb{E}\, X^{\check{\otimes} k}$ in $B^{\check{\otimes} k}$ for every $k \geqslant 1$, that the sequence X_n is tight, and that every $\langle x^*, X\rangle$, $x^* \in B^*$, satisfies the Carleman condition* (17.1)*. Then $X_n \xrightarrow{\mathrm{d}} X$.*

Proof. By Theorem 18.8 and Remark 18.9, $x^*(X_n) \xrightarrow{\mathrm{d}} x^*(X)$ for every $x^* \in B^*$. Since X_n is tight, every subsequence has a subsubsequence that converges in distribution to some random variable Y in B [6]. Then, along the subsubsequence, $x^*(X_n) \xrightarrow{\mathrm{d}} x^*(Y)$ for every $x^* \in B^*$, and thus $x^*(Y) \overset{\mathrm{d}}{=} x^*(X)$, which implies $Y \overset{\mathrm{d}}{=} X$, see Theorem 17.3. Hence every subsequence has a subsubsequence converging (in distribution) to X, which implies that the full sequence converges. □

In Hilbert spaces, we can use the second moment to deduce tightness. We regard as usual the second moments as operators on H; recall that they always are positive operators by Theorem 9.4. Recall also that if the second injective moment $\mathbb{E}\, X^{\check{\otimes} 2}$ exists (in any sense, e.g. Dunford) and is a trace class operator, then the projective moment $\mathbb{E}\, X^{\hat{\otimes} 2}$ exists too, in Bochner (and thus Pettis) sense by Theorem 9.10. (And conversely, see Corollary 9.11.) Moreover, the projective and injective second moments are then given by the same operator, so it does not matter which of them we use.

We identify $H \hat{\otimes} H$ with the space $\mathcal{N}(H)$ of nuclear (trace class) operators on H, see Section 9. For operators T, U in a Hilbert space H, we let as usual $T \leqslant U$ mean $\langle Tx, x\rangle \leqslant \langle Ux, x\rangle$ for every $x \in H$; in particular, $T \geqslant 0$ (T is positive) if $\langle Tx, x\rangle \geqslant 0$.

Theorem 18.12. *Let H be a separable Hilbert space, and let $\{X_\alpha : \alpha \in A\}$ be a family of H-valued random variables. Suppose that there is a nuclear operator $T \in H \hat{\otimes} H = \mathcal{N}(H)$ such that $\mathbb{E}\, X_\alpha^{\check{\otimes} 2} \leqslant T$ for every $\alpha \in A$, with the moment existing in Dunford sense. Then $\{X_\alpha : \alpha \in A\}$ is tight.*

Proof. By the spectral theorem for compact self-adjoint operators, e.g. [13, Corollary II.5.4], there exists an ON basis (e_n) in H such that

$$T = \sum_{n=1}^{\infty} \lambda_n e_n \otimes e_n, \tag{18.10}$$

where λ_n is the eigenvalue corresponding to e_n, $\lambda_n \geqslant 0$ because $T \geqslant \mathbb{E}\, X_\alpha^{\check{\otimes} 2} \geqslant 0$, and $\sum_{n=1}^{\infty} \lambda_n = \|T\|_{\mathcal{N}(H)} < \infty$.

Choose a positive sequence $a_n \to \infty$ such that $\sum_{n=1}^{\infty} a_n \lambda_n < \infty$, and define

$$|||x|||^2 := \sum_{n=1}^{\infty} a_n |\langle X, e_n\rangle|^2 \leqslant \infty.$$

Define $K_r := \{x \in H : |||x||| \leqslant r\}$. It is well known (and easy to see, e.g. using [42, Lemma 2.2]) that each K_r is a compact subset of H. Moreover, for each α, using (6.3) and (18.10),

$$\begin{aligned}\mathbb{E}\,|||X_\alpha|||^2 &= \sum_{n=1}^{\infty} a_n\, \mathbb{E}\,|\langle X_\alpha, e_n\rangle|^2 = \sum_{n=1}^{\infty} a_n \langle \mathbb{E}\, X_\alpha^{\hat{\otimes}2} e_n, e_n\rangle \\ &\leqslant \sum_{n=1}^{\infty} a_n \langle T e_n, e_n\rangle = \sum_{n=1}^{\infty} a_n \lambda_n < \infty.\end{aligned}$$

Hence, by Markov's inequality, $\mathbb{P}(X_\alpha \notin K_r) \leqslant \sum_{n=1}^{\infty} a_n\lambda_n / r^2 \to 0$ as $r \to \infty$, uniformly in $\alpha \in A$, which shows that $\{X_\alpha\}$ is tight. □

Theorem 18.13. *Let H be a separable Hilbert space, and let $\{X_\alpha : \alpha \in A\}$ be a family of H-valued random variables. If the family $\{\mathbb{E}\, X_\alpha^{\hat{\otimes}2} : \alpha \in A\}$ of projective second moments is relatively compact in $H \hat{\otimes} H = \mathcal{N}(H)$, with the moments existing in Pettis sense, then $\{X_\alpha : \alpha \in A\}$ is tight.*

Proof. Consider any sequence $(X_{\alpha_n})_{n=1}^{\infty}$ with $\alpha_n \in A$, $n \geqslant 1$. By the compactness assumption, there is a subsequence, which we simply denote by (X_n), such that $\mathbb{E}\, X_n^{\hat{\otimes}2} \to V$ in $H \hat{\otimes} H = \mathcal{N}(H)$ for some $V \in \mathcal{N}(H)$. By taking a further subsequence, we may assume that

$$\|\,\mathbb{E}\, X_n^{\hat{\otimes}2} - V\|_{\mathcal{N}(H)} < 2^{-n}. \tag{18.11}$$

Let $T_n := \mathbb{E}\, X_n^{\hat{\otimes}2} - V$; this is a symmetric nuclear operator and the corresponding positive operator $|T_n| := (T_n^* T_n)^{1/2}$ satisfies

$$\|\,|T_n|\,\|_{\mathcal{N}(H)} = \|T_n\|_{\mathcal{N}(H)} < 2^{-n}. \tag{18.12}$$

Define $T := V + \sum_{n=1}^{\infty} |T_n| \in \mathcal{N}(H)$, where the sum converges in $\mathcal{N}(H)$ by (18.12). Note that for any $x \in H$,

$$\langle \mathbb{E}\, X_n^{\hat{\otimes}2} x, x\rangle = \langle T_n x, x\rangle + \langle V x, x\rangle \leqslant \langle |T_n| x, x\rangle + \langle V x, x\rangle \leqslant \langle T x, x\rangle.$$

Thus, $\mathbb{E}\, X_n^{\hat{\otimes}2} \leqslant T$. By Theorem 18.12, the sequence (X_n) is tight, and thus there is a subsequence that converges in distribution.

We have shown that every sequence (X_{α_n}) has a subsequence that converges in distribution; this shows that $\{X_\alpha\}$ is tight. □

This leads to the following convergence criterion, combining the second projective and arbitrary injective (or, equivalently, weak) moments. Compare Theorems 18.3 and 18.6.

Theorem 18.14. *Suppose that X and X_n, $n \geqslant 1$, are H-valued random variables, where H is a separable Hilbert space. If $\mathbb{E}\, X_n^{\hat\otimes 2} \to \mathbb{E}\, X^{\hat\otimes 2}$ in $H^{\hat\otimes 2} = \mathcal{N}(H)$ and $\mathbb{E}\, X_n^{\check\otimes k} \xrightarrow{\mathrm{w}} \mathbb{E}\, X^{\check\otimes k}$ in $H^{\check\otimes k}$ for every $k \geqslant 1$, with all moments existing in Pettis sense, and furthermore every $\langle y, X\rangle$, $y \in H$, satisfies the Carleman condition* (17.1)*, then $X_n \xrightarrow{\mathrm{d}} X$.*

Proof. Since the sequence $\mathbb{E}\, X_n^{\hat\otimes 2}$ converges, it is relatively compact, and thus Theorem 18.13 shows that the sequence (X_n) is tight. The result follows from Corollary 18.11. □

We do not know whether there are similar results for other Banach spaces. Example 18.5 shows that convergence of moments is not enough to provide tightness in general. Note that a commonly used sufficient condition for tightness of a family $\{X_\alpha\}$ in $C[0,1]$, assuming $\{X_\alpha(0)\}$ tight, is that $\mathbb{E}(X_\alpha(s) - X_\alpha(t))^4 \leqslant C|s-t|^\beta$ for some $C < \infty$, $\beta > 1$ (typically, $\beta = 2$) and all $s, t \in [0,1]$. (See [5, Theorem 12.3 and (12.51)] for a more general result, and [5, Theorem 15.6] for a similar result for $D[0,1]$.) By expanding the fourth power and using (11.2), this can be seen as a continuity condition on the fourth moments $\mathbb{E}\, X_\alpha^{\check\otimes 4} \in C([0,1]^4)$. This suggests that also for other spaces, it might be possible to find tightness criteria using suitable subspaces of $B^{\check\otimes k}$ or $B^{\hat\otimes k}$. We have, however, not explored this further.

Appendix A. The reproducing Hilbert space

In this appendix (partly based on [42, Chapter 8]) we study a construction closely related to the injective second moment, and explore the connection. We suppose that X is weakly measurable and furthermore that $x^*(X) \in L^2(\mathbb{P})$ for every $x^* \in B^*$; this is by Lemma 6.8 equivalent to the existence of the weak second moment $\mathbb{E}\big(x_1^*(X)x_2^*(X)\big)$, $x_1^*, x_2^* \in B^*$. Furthermore, this holds whenever the injective second moment $\mathbb{E}\, X^{\check\otimes 2}$ exists in Dunford sense, and the converse holds under weak conditions, for example when B is separable, see Theorem 6.10.

By Remark 5.11, then $T_X : B^* \to L^2(\mathbb{P})$ and $T_X^* : L^2(\mathbb{P}) \to B^{**}$, and thus the composition $T_X^* T_X : B^* \to B^{**}$. This operator is characterised by

$$\langle T_X^* T_X x^*, y^*\rangle = \langle T_X x^*, T_X y^*\rangle = \mathbb{E}\big(x^*(X)y^*(X)\big). \tag{A.1}$$

In other words, $T_X^* T_X : B^* \to B^{**}$ is the operator corresponding to the weak moment, seen as a bilinear form $B^* \times B^* \to \mathbb{R}$.

By (5.1), (5.3) and Remark 5.11, or by (5.2) and (A.1),

$$T_X^* T_X x^* = \mathbb{E}\big(x^*(X)X\big), \tag{A.2}$$

with the expectation existing in Dunford sense.

Lemma A.1. *Suppose that $x^*(X) \in L^2(\mathbb{P})$ for all $x^* \in B^*$. Then the following are equivalent.*

(i) *X is Pettis integrable, i.e., $\mathbb{E}\, X$ exists in Pettis sense.*

(ii) $T_X^* : L^2(\mathbb{P}) \to B$.
(iii) $T_X^* T_X : B^* \to B$.

Proof. (i) $\iff$ (ii): By Remarks 5.3 and 5.11.
(ii) $\implies$ (iii): Trivial.
(iii) $\implies$ (ii): If (iii) holds, then T_X^* maps $\operatorname{im}(T_X)$ into B, and thus $T_X^*(\overline{\operatorname{im}(T_X)}) \subseteq B$. Since $\operatorname{im}(T_X)^\perp = \ker(T_X^*)$, it follows that $T_X^* : L^2(\mathbb{P}) \to B$. □

By Remark 5.11, the assertions (i)–(iii) in Lemma A.1 hold whenever B is separable, and more generally when X is weakly a.s. separably valued.

Recall from Section 4 that $B \check{\otimes} B$ can be regarded as a subspace of $L(B^*; B)$. If the injective second moment $\mathbb{E}\, X^{\check{\otimes} 2}$ exists in Pettis sense, then by (4.9), (6.3) and (A.1), $\mathbb{E}\, X^{\check{\otimes} 2} \in B \check{\otimes} B$ corresponds to the operator $T_X^* T_X$. (In particular, $T_X^* T_X \in L(B^*, B)$ so X is Pettis integrable by Lemma A.1.) Hence $T_X^* T_X$ can be seen as a form of the injective second moment.

Remark A.2. Note that if, for example, B is separable, then $T_X^* T_X \in L(B^*, B)$ by Lemma A.1. However, it does not always correspond to an element of $B \check{\otimes} B$. For example, if $B = c_0$ and X is as in Example 7.5, then $\mathbb{E}\, X^{\check{\otimes} 2}$ is the infinite diagonal matrix $(p_n a_n^2 \delta_{mn})_{m,n=1}^\infty$, and the corresponding operator $T_X^* T_X : c_0^* = \ell^1 \to c_0$ is the multiplication operator $(b_n)_1^\infty \mapsto (p_n a_n^2 b_n)_1^\infty$. Choosing a_n such that $p_n a_n^2 = 1$, $\mathbb{E}\, X^{\hat{\otimes} 2}$ is thus the (infinite) identity matrix and $T_X^* T_X$ is the inclusion map $\ell^1 \to c_0$; in this case $\mathbb{E}\, X^{\check{\otimes} 2}$ exists in Dunford sense but not in Pettis sense by Example 7.5, and $T_X^* T_X \in L(c_0^*, c_0)$ but $T_X^* T_X \notin c_0 \check{\otimes} c_0$. (Recall from Theorem 12.1 that $c_0 \check{\otimes} c_0 = c_0(\mathbb{N}^2)$, so the identity matrix is not an element of $c_0 \check{\otimes} c_0$.)

Assume in the remainder of this appendix also that the assertions of Lemma A.1 hold. (Recall that this is the case for example if B is separable, or if $\mathbb{E}\, X^{\check{\otimes} 2}$ exists in Pettis sense.) Thus $T_X : B^* \to L^2(P)$, $T_X^* : L^2(\mathbb{P}) \to B$ and $T_X^* T_X : B^* \to B$. Furthermore, $\operatorname{im}(T_X)^\perp = \ker(T_X^*)$ and thus $\overline{\operatorname{im}(T_X)} = \ker(T_X^*)^\perp$.

T_X^* induces an bijection of $\ker(T_X^*)^\perp$ onto $\operatorname{im}(T_X^*) \subseteq B$. Let H_X be $\operatorname{im}(T_X^*)$ equipped with the inner product induced by this bijection, i.e.,

$$\langle x, y \rangle_{H_X} := \langle (T_X^*)^{-1} x, (T_X^*)^{-1} y \rangle_{L^2(\mathbb{P})} \tag{A.3}$$

where $(T_X^*)^{-1} : H_X \to \ker(T_X^*)^\perp \subseteq L^2(\mathbb{P})$. Thus H_X is a Hilbert space isomorphic to $\ker(T_X^*)^\perp$ and T_X^* is a Hilbert space isomorphism $\overline{\operatorname{im}(T_X)} = \ker(T_X^*)^\perp \to H_X$. H_X is called the *reproducing kernel Hilbert space* corresponding to X [42]. Note that $H_X \subseteq B$ with a continuous inclusion, since $T_X^* : \ker(T_X^*)^\perp \to B$ is continuous. Furthermore, $\operatorname{im}(T_X)$ is dense in $\overline{\operatorname{im}(T_X)} = \ker(T_X^*)^\perp$, and thus $\operatorname{im}(T_X^* T_X)$ is dense in H_X.

The unit ball K_X of H_X is the image under T_X^* of the unit ball of $L^2(\mathbb{P})$. The latter unit ball is weakly compact, and since $T_X^* : L^2(\mathbb{P}) \to B$ is continuous, and therefore weakly continuous, K_X is a weakly compact subset of B. In particular, K_X is closed in B.

If $x \in H_X$ and $x^* \in B^*$, then $T_X^* T_X x^* \in H_X$ and

$$\langle T_X^* T_X x^*, x\rangle_{H_X} = \langle T_X x^*, (T_X^*)^{-1} x\rangle = \langle x^*, T_X^* (T_X^*)^{-1} x\rangle = \langle x^*, x\rangle. \quad \text{(A.4)}$$

Hence, the operator $T_X^* T_X : B^* \to H_X \subseteq B$ is the adjoint of the inclusion $H_X \to B$. Furthermore, by (A.4) and (A.1), for $x^*, y^* \in B^*$,

$$\langle T_X^* T_X x^*, T_X^* T_X y^*\rangle_{H_X} = \langle x^*, T_X^* T_X y^*\rangle = \mathbb{E}\big(x^*(X) y^*(X)\big). \quad \text{(A.5)}$$

The operator $T_X^* T_X$ and the Hilbert space H_X determine each other; more precisely, we have the following.

Theorem A.3. *If X and Y are B-valued random variables such that $T_X, T_Y : B^* \to L^2(\mathbb{P})$ and $T_X^*, T_Y^* : L^2(\mathbb{P}) \to B$, then the following are equivalent:*

- (i) *$H_X = H_Y$ (as vector spaces with given inner products).*
- (ii) *The unit balls K_X and K_Y are the same (as subsets of B).*
- (iii) *$T_X^* T_X = T_Y^* T_Y$.*
- (iv) *$\mathbb{E}\big(x_1^*(X) x_2^*(X)\big) = \mathbb{E}\big(x_1^*(Y) x_2^*(Y)\big)$ for every $x_1^*, x_2^* \in B^*$.*
- (v) *$\mathbb{E}\, X^{\tilde{\otimes} 2} = \mathbb{E}\, Y^{\tilde{\otimes} 2}$, assuming that either both moments exist in Pettis sense, or that B is separable (in which case the moments exist at least in Dunford sense).*

Proof. (i) $\iff$ (ii): Each Hilbert space determines its unit ball. Conversely, the unit ball determines the space and its norm, and thus the inner product by the polarisation identity $\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$.

(i) $\implies$ (iii): Immediate from (A.4).

(iii) $\implies$ (i): $T_X^* T_X$ determines both the set $\operatorname{im}(T_X^* T_X) \subseteq H_X$ and, by (A.5), the inner product in H_X restricted to this subspace; since $\operatorname{im}(T_X^* T_X)$ is dense in H_X, and H_X continuously included in B, this determines H_X.

(iii) $\iff$ (iv): By (A.1).

(iv) $\iff$ (v): By Corollary 6.14 or 6.12. □

As said above, $T_X^* T_X$ can be seen as the second injective moment of X; by Theorem A.3, also the space H_X can be seen as a manifestation of the second injective moment.

The space H_X is important for the law of iterated logarithm in Banach spaces, since the unit ball K_X turns out to be the natural limit set, see Ledoux and Talagrand [42, Chapter 8] for a detailed discussion. In particular [42, Theorem 8.5], if B is separable and $S_n = \sum_{i=1}^n X_i$ where X_i are independent copies of X, with $\mathbb{E}\, X = 0$, and further the sequence $S_n/\sqrt{2n \log\log n}$ is a.s. relatively compact (which holds under rather general conditions, but not always), then its set of limit points is a.s. K_X. See also [1, 2] for exceptional cases where H_X still is important.

Example A.4. Let $W \in C[0,1]$ be standard Brownian motion, see Examples 5.12 and 11.11. By (5.5) and an integration by parts,

$$T_X(\mu) = \int_0^1 W(t)\,\mathrm{d}\mu(t) = \int_0^1 \mu[t,1]\,\mathrm{d}W(t). \tag{A.6}$$

Hence $\operatorname{Im}(T_X)$ is the space of stochastic integrals $\int_0^1 g(t)\,\mathrm{d}W(t)$ where g is a deterministic function of the type $g(t) = \mu[t,1]$ with $\mu \in M[0,1]$, i.e., a function g on $[0,1]$ of bounded variation. Since $\|\int g\,\mathrm{d}W\|_{L^2(\mathbb{P})} = \|g\|_{L^2[0,1]}$ and the functions of bounded variation are dense in $L^2[0,1]$, it follows that

$$\overline{\operatorname{Im}(T_X)} = \left\{ \int_0^1 g\,\mathrm{d}W : g \in L^2[0,1] \right\}. \tag{A.7}$$

Moreover, by (5.6),

$$T_X^*\Big(\int_0^1 g\,\mathrm{d}W\Big)(t) = \left\langle W(t), \int_0^1 g(s)\,\mathrm{d}W(s) \right\rangle = \int_0^t g(s)\,\mathrm{d}s. \tag{A.8}$$

Hence the reproducing Hilbert space H_X is given by

$$H_X = \left\{ \int_0^t g(s)\,\mathrm{d}s : g \in L^2[0,1] \right\}; \tag{A.9}$$

equivalently, this is the space of absolutely continuous functions f on $[0,1]$ with $f(0) = 0$ and $f' \in L^2[0,1]$; the norm is $\|f'\|_{L^2}$. This is the usual *Cameron–Martin space*, see e.g. [37, Example 8.19]. (See [37, Section VIII.4] for a generalization to more general Gaussian processes.) Note that the law of iterated logarithm for Brownian motion [38, Theorem 27.18] says that the cluster set of $Y_s(t) := W(st)/\sqrt{2s\log\log s} \in C[0,1]$ as $s \to \infty$ a.s. is the unit ball K_X of H_X; this is another example of the connection between the law of iterated logarithm and the reproducing Hilbert space.

Remark A.5. The name *reproducing Hilbert space* is in a more general context used for a Hilbert space H of functions on some set $\mathcal{T}$ such that each point evaluation $f \mapsto f(t)$ is a continuous linear functional on H, see Aronszajn [3] and e.g. [37, Appendix F]. The definition implies that for each $t \in \mathcal{T}$ there is an element $K_t \in H$ such that $f(t) = \langle f, K_t\rangle_H$ for all $f \in H$. In particular, $K_s(t) = \langle K_s, K_t\rangle_H$; the symmetric function $K(s,t) := K_s(t) = \langle K_s, K_t\rangle$ on $\mathcal{T} \times \mathcal{T}$ is known as the *reproducing kernel*, see [37, Theorem F.3] for some of its basic properties.

We can connect the space H_X constructed above to this general setting by taking $\mathcal{T} = B^*$ and regarding elements of B as functions on B^* in the usual way, regarding $x \in B$ as the function $x^* \mapsto \langle x, x^*\rangle$. The point evaluations are thus the elements $x^* \in B^*$, which are continuous on $H_X \subseteq B$, and (A.4) shows that

$$K_{x^*} = T_X^* T_X x^*; \tag{A.10}$$

hence the reproducing kernel is, using (A.5), given by

$$K(x^*, y^*) = \mathbb{E}\big(x^*(X)y^*(X)\big). \tag{A.11}$$

In other words, H_X is a reproducing Hilbert space of functions on B^*, and the reproducing kernel equals the weak second moment of X given by (1.1).

We mention also that the construction above of H_X is an instance of [37, Theorem F.5], taking (in the notation there) $\mathcal{T} = B^*$, $h_{x^*} = x^*(X)$, and $H = \overline{\mathrm{Im}(T_X)}$, the closed linear subspace of $L^2(\mathbb{P})$ spanned by $\{x^*(X) : x^* \in B^*\}$. (Then the operator R defined there equals our T_X^* by (5.3) and (5.2).)

We finally mention the following result, adapted from [42, Lemma 8.4]. For simplicity, we consider only separable B. (Note that then $T_X^* : L^2(\mathbb{P}) \to B$ by Remark 5.11.) We do not know whether the result extends to non-separable spaces.

Theorem A.6. *Suppose that B is separable and that $x^*(X) \in L^2(\mathbb{P})$ for every $x^* \in B^*$. Then the following are equivalent.*

(i) *K_X is a compact subset of B.*
(ii) *$T_X : B^* \to L^2(\mathbb{P})$ is a compact operator.*
(iii) *$T_X^* : L^2(\mathbb{P}) \to B$ is a compact operator.*
(iv) *$T_X^* T_X : B^* \to B$ is a compact operator.*
(v) *If $x_n^* \xrightarrow{\mathrm{w}^*} 0$ in B^*, then $\mathbb{E}\, x_n^*(X)^2 \to 0$. In other words, $T_X : B^* \to L^2(\mathbb{P})$ is sequentially weak*-norm continuous.*
(vi) *The family $\{x^*(X)^2 : \|x^*\| \leqslant 1\}$ of random variables is uniformly integrable.*
(vii) *The injective moment $\mathbb{E}\, X^{\check{\otimes}2}$ exists in Pettis sense.*

Proof. (i) $\iff$ (iii): By the definition of compact operators (which says that K_X is relatively compact, i.e., $\overline{K_X}$ is compact) and the fact shown above that K_X is a closed subset of B.

(ii) $\iff$ (iii): Standard operator theory [13, VI.3.4].

(iii) $\implies$ (iv): Immediate.

(iv) $\implies$ (v): Let $x_n^* \xrightarrow{\mathrm{w}^*} 0$ in B^*. If $\xi \in L^2(\mathbb{P})$, then $T_X^*\xi \in B$, and thus $\langle T_X x_n^*, \xi\rangle = \langle x_n^*, T_X^*\xi\rangle \to 0$; hence $T_X x_n^* \xrightarrow{\mathrm{w}} 0$ in $L^2(\mathbb{P})$ and thus $T_X^* T_X x_n^* \xrightarrow{\mathrm{w}} 0$ in B. Moreover, the sequence x_n^* is bounded by the uniform boundedness principle since it is weak* convergent. Since $T_X^* T_X$ is a compact operator, it follows that the sequence $T_X^* T_X x_n^*$ is relatively compact in B, and thus the weak convergence $T_X^* T_X x_n^* \xrightarrow{\mathrm{w}} 0$ implies norm convergence, i.e., $\|T_X^* T_X x_n^*\| \to 0$. Consequently, (A.1) implies

$$\mathbb{E}\, x_n^*(X)^2 = \langle T_X^* T_X x_n^*, x_n^*\rangle \leqslant \|T_X^* T_X x_n^*\|_B \|x_n^*\|_{B^*} \to 0.$$

(v) $\implies$ (ii): Let K^* be the closed unit ball of B^* with the weak* topology; then K^* is compact by Alaoglu's theorem [13, Theorem V.3.2]. Moreover, since B is separable, K^* is metrizable, and thus sequential continuity on K^* is equivalent to ordinary continuity. Hence (v) implies that T_X is a continuous map from K^* into $L^2(\mathbb{P})$. Consequently its image is compact, which means that T_X is a compact operator.

(ii) $\implies$ (vi): By (ii), the family $\{x^*(X) : \|x^*\| \leqslant 1\}$ is relatively compact in $L^2(\mathbb{P})$, which implies that $\{x^*(X)^2 : \|x^*\| \leqslant 1\}$ is uniformly integrable.

(For example because $f \mapsto f^2$ is a continuous map $L^2(\mathbb{P}) \to L^1(\mathbb{P})$, so the latter family is relatively compact in $L^1(\mathbb{P})$, and thus uniformly integrable by [22, Corollary IV.8.11].)

(vi) $\implies$ (v): If $x_n^* \xrightarrow{\text{w}*} 0$, then $x_n^*(X)^2 \to 0$ a.s., and thus uniform integrability implies $\mathbb{E}\, x_n^*(X)^2 \to 0$.

(vi) $\iff$ (vii): By Theorem 6.19(ii)(a). □

Appendix B. The Zolotarev distances

The Zolotarev distance $\zeta_s(X,Y)$, where the parameter $s > 0$ is a fixed real number, is a measure of the distance (in some sense) between the distributions of two random variables X and Y with values in the same Banach space B. We give the definition of ζ_s in Subsection B.2 below and explain the connection to tensor products and moments, but first we recall some preliminaries on Fréchet derivatives in Subsection B.1.

Note that $\zeta_s(X,Y)$ depends only on the distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$; we may write $\zeta_s(\mathcal{L}(X),\mathcal{L}(Y)) = \zeta_s(X,Y)$, and regard ζ_s as a distance between probability distributions on B, but it is often convenient to use the notation $\zeta_s(X,Y)$ with random variables X and Y.

It is important to note that $\zeta_s(X,Y)$ may be infinite. Hence ζ_s defines a metric only on suitable classes of probability distributions on B, where we know *a priori* that $\zeta_s(X,Y) < \infty$.

B.1. **Fréchet differentiablity.** We recall some well-known facts about derivatives of Banach space valued functions, see e.g. [11] for details.

Let B and B_1 be Banach spaces, and let U be a non-empty open subset of B. A function $f : U \to B_1$ is *(Fréchet) differentiable at a point* $x \in U$ if there exists a bounded linear operator $T : B \to B_1$ such that

$$f(y) - f(x) - T(y-x) = o\bigl(\|y-x\|\bigr) \tag{B.1}$$

as $y \to x$. The linear operator $T \in L(B;B_1)$ then is uniquely determined; it is called the *derivative* of f at x and is denoted by $f'(x)$ or $Df(x)$.

The function $f : U \to B_1$ is said to be *differentiable* if it is differentiable at every $x \in U$. In this case, the derivative f' is a function $U \to L(B;B_1)$. If furthermore $f' : U \to L(B;B_1)$ is continuous, f is said to be *continuously differentiable*.

Since $L(B;B_1)$ is a Banach space, we may iterate: If the derivative $f' : U \to L(B;B_1)$ is differentiable, its derivative f'' is called the second derivative of f, and so on. Note that f'' then is a function on U with values in $L(B;L(B;B_1)) = L(B^2;B_1)$, so the second derivative $f''(x)$ at a point $x \in U$ is a bilinear map $B \times B \to B_1$. More generally, the k:th derivative $f^{(k)}(x)$, if it exists, is a k-linear map $B^k \to B_1$. It can be shown [11, Théorème 5.1.1] that this map is symmetric. Since $L(B^k;B_1) \cong L(B^{\hat{\otimes}k};B_1)$ by an extension of Theorem 4.6, we can also regard the k:th derivative $f^{(k)}(x)$ as a (symmetric) linear map $B^{\hat{\otimes}k} \to B_1$. We may take advantage of this by writing $f^{(k)}(x)(y,\dots,y)$ as $f^{(k)}(x)(y^{\otimes k})$.

$C^0(U;B_1)$ is defined to be the linear space of all continuous functions $f : U \to B_1$, and $C^1(U;B_1)$ is the linear space of all continuously differentiable functions $f : U \to B_1$. More generally, $C^k(U;B_1)$ is the linear space of all k times continuously differentiable functions $U \to B_1$; this may be defined formally by induction for $k \geqslant 1$ as the space of all differentiable functions $f : U \to B_1$ such that $f' \in C^{k-1}(U; L(B;B_1))$.

Note that (B.1) implies continuity at x; thus a differentiable function is continuous. Hence, $C^0(U;B_1) \supset C^1(U;B_1) \supset C^2(U;B_1) \supset \cdots$.

For $0 < \gamma \leqslant 1$, we define $\operatorname{Lip}_\gamma(U;B_1)$ to be the linear space of all functions $f : U \to B_1$ such that

$$\|f\|_{\operatorname{Lip}_\gamma} := \sup_{x\neq y} \frac{\|f(x)-f(y)\|}{\|x-y\|^\gamma} \tag{B.2}$$

is finite. More generally, for $s > 0$, we write $s = m + \gamma$ with $m \in \{0,1,2\dots\}$ and $0 < \gamma \leqslant 1$ (thus $m = \lceil s \rceil - 1$) and define

$$\operatorname{Lip}_s(U;B_1) := \bigl\{ f \in C^m(U;B_1) : f^{(m)} \in \operatorname{Lip}_\gamma\bigl(U; L(B^m;B_1)\bigr)\bigr\}. \tag{B.3}$$

with

$$\|f\|_{\operatorname{Lip}_s} := \|f^{(m)}\|_{\operatorname{Lip}_\gamma}. \tag{B.4}$$

It follows from a Taylor expansion [11, Théorème 5.6.1] at 0 that if $f \in \operatorname{Lip}_s(B;B_1)$, then

$$f(x) = \sum_{k=0}^{m} \frac{1}{k!} f^{(k)}(0)(x^{\otimes k}) + O\bigl(\|f\|_{\operatorname{Lip}_s}\|x\|^s\bigr), \tag{B.5}$$

where the implicit constant is universal (it can be taken as 1). (The term with $k = 0$ in (B.5) is just the constant $f(0)$.) In particular,

$$\|f(x)\| = O\bigl(1 + \|x\|^s\bigr), \qquad x \in B. \tag{B.6}$$

Note that $\|\ \|_{\operatorname{Lip}_s}$ is only a seminorm. In fact, we have the following.

Lemma B.1. *Let $f : U \to B_1$ where $U \subseteq B$ is connected, and let $s > 0$. Then the following are equivalent, with $m = \lceil s \rceil - 1$,*

(i) $\|f\|_{\operatorname{Lip}_s} = 0$.
(ii) $f^{(m)}$ *is constant.*
(iii) $f^{(m+1)}(x) = 0$ *for all* $x \in U$.
(iv) $f(x) = \sum_{k=0}^{m} \alpha_k(x,\dots,x)$, *where* $\alpha_k \in L(B^k;B_1)$ *is a k-linear map.*

Proof. (i) $\iff$ (ii): It follows from (B.2) and (B.4) that

$$\|f\|_{\operatorname{Lip}_s} = 0 \iff \|f^{(m)}\|_{\operatorname{Lip}_\gamma} = 0 \iff f^{(m)}(x) = f^{(m)}(y) \text{ for any } x,y \in U.$$

(ii) $\Longrightarrow$ (iii): Obvious by the definition of the derivative.

(iii) $\Longrightarrow$ (iv): Suppose for simplicity that U is convex. (The general case follows easily, but we omit this.) By a translation we may assume that $0 \in U$, and then (iv) follows from Taylor's formula [11, Théorème 5.6.1]. (We have $\alpha_k = f^{(k)}(0)/k!$.)

(iv) $\Longrightarrow$ (ii): It is easily seen by induction that if each α_k is symmetric, as we may assume by symmetrization, then for $1 \leqslant j \leqslant m$,

$$f^{(j)}(x)(y_1, \dots, y_j) = \sum_{k=j}^{m} \frac{k!}{(k-j)!} \alpha_k(x, \dots, x, y_1, \dots, y_j)$$

(with $k - j$ arguments of α_k equal to x); in particular, $f^{(m)}(x) = m!\, \alpha_m$, which does not depend on x. □

B.2. **Zolotarev distances.** We now define the Zolotarev distance ζ_s as follows, for any $s > 0$; see Zolotarev [66, 67, 68, 69] and e.g. [48, 49, 60] for further details.

Let B be a Banach space, and suppose that every bounded multilinear form $B^k \to \mathbb{R}$ is $\mathcal{B}_w$-measurable (i.e., measurable for the product σ-field $\mathcal{B}_w^k$), for any $k \geqslant 1$. (For $k = 1$, this holds by the definition of $\mathcal{B}_w$.) If B is separable, this assumption always holds, since every bounded multilinear form is continuous and thus Borel measurable on B^k; moreover, when B is separable the Borel σ-field on B^k equals the product σ-field $\mathcal{B}^k$ and $\mathcal{B} = \mathcal{B}_w$. The main example with B non-separable is $D[0,1]$, where $\mathcal{B}_w = \mathcal{D}$ by Theorem 15.5 and every multilinear form is $\mathcal{D}$-measurable by Theorem 15.8. (Another example is $c_0(S)$, where $\mathcal{B}_w = \mathcal{C}$ by Theorem 12.4 and every multilinear form is $\mathcal{C}$-measurable as a consequence of Lemma 12.2.)

We let $\mathrm{Lip}_s^w(B;\mathbb{R})$ be the set of all $\mathcal{B}_w$-measurable functions in $\mathrm{Lip}_s(B;\mathbb{R})$; this is evidently a subspace of $\mathrm{Lip}_s(B;\mathbb{R})$. If B is separable, then $\mathrm{Lip}_s^w(B;\mathbb{R}) = \mathrm{Lip}(B;\mathbb{R})$, since every function in $\mathrm{Lip}_s(B;\mathbb{R})$ is continuous and thus Borel measurable, and $\mathcal{B} = \mathcal{B}_w$.

Let X and Y be two weakly measurable B-valued random variables, and suppose that $\bar{\mathbb{E}}\|X\|^s, \bar{\mathbb{E}}\|Y\|^s < \infty$. (We use upper integrals here, since the norms $\|X\|, \|Y\|$ are not necessarily measurable.) Define

$$\zeta_s(X,Y) := \sup\bigl\{ |\,\mathbb{E}\, f(X) - \mathbb{E}\, f(Y)| : f \in \mathrm{Lip}_s^w(B;\mathbb{R}) \text{ with } \|f\|_{\mathrm{Lip}_s} \leqslant 1 \bigr\}. \tag{B.7}$$

By assumption, f is $\mathcal{B}_w$-measurable, and thus $f(X)$ is measurable. Moreover, by (B.6), $|f(X)| = O(1 + \|X\|^s)$ and by assumption $\bar{\mathbb{E}}\|X\|^s < \infty$; hence $\mathbb{E}\,|f(X)| < \infty$. Similarly, $\mathbb{E}\,|f(Y)| < \infty$, and thus $|\,\mathbb{E}\, f(X) - \mathbb{E}\, f(Y)|$ is well-defined and finite for every $f \in \mathrm{Lip}_s^w(B;\mathbb{R})$. Thus $\zeta_s(X,Y)$ is a well-defined number in $[0,\infty]$. Moreover, $\mathbb{E}\, f(X)$ and $\mathbb{E}\, f(Y)$ depend only on the distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$, and thus ζ_s is really a distance between the distributions; we may write $\zeta_s(\mathcal{L}(X), \mathcal{L}(Y)) = \zeta_s(X,Y)$.

By (B.7) and homogeneity, for any $f \in \mathrm{Lip}_s^w(B;\mathbb{R})$, if $\zeta_s(X,Y) < \infty$,

$$|\,\mathbb{E}\, f(X) - \mathbb{E}\, f(Y)| \leqslant \|f\|_{\mathrm{Lip}_s}\, \zeta_s(X,Y). \tag{B.8}$$

It is clear that $\zeta_s(X,Y) \geqslant 0$ with $\zeta_s(X,X) = 0$, and that ζ_s is symmetric and that the triangle inequality holds. Moreover, if $x^* \in B^*$, then $e^{\mathrm{i}x^*} \in \mathrm{Lip}_s(B;\mathbb{C})$ for any $s > 0$ as is easily seen [68]. Hence, by taking real and

imaginary parts, (B.8) implies that if $\zeta_s(X,Y)=0$, then

$$\mathbb{E}\, e^{\mathrm{i}x^*(X)} = \mathbb{E}\, e^{\mathrm{i}x^*(Y)} \tag{B.9}$$

and it is well-known that this implies that $\mathcal{L}(X)=\mathcal{L}(Y)$ on $\mathcal{B}_w$. (In fact, (B.9) implies, by replacing x^* by tx^* with $t\in\mathbb{R}$, that $x^*(X)$ and $x^*(Y)$ have the same characteristic function, and thus $x^*(X)\overset{\mathrm{d}}{=}x^*(Y)$ for every $x^*\in B^*$. This implies $\mathcal{L}(X)=\mathcal{L}(Y)$ on $\mathcal{B}_w$ as seen in Theorem 17.3 and its proof.) Consequently, ζ_s is a metric on any set of probability distributions on $(B,\mathcal{B}_w)$ such that ζ_s is finite.

However, for $s>1$, $\zeta_s(X,Y)$ may be infinite. The following lemma says exactly when ζ_s is finite.

Lemma B.2. *Suppose that X and Y are weakly measurable B-valued random variables such that $\bar{\mathbb{E}}\|X\|^s,\bar{\mathbb{E}}\|Y\|^s<\infty$. Then the moments $\mathbb{E}\,X^{\hat{\otimes}k}$ and $\mathbb{E}\,Y^{\hat{\otimes}k}$ exist in Dunford sense for $1\leqslant k\leqslant m=\lceil s\rceil-1$, and the following are equivalent.*

- (i) $\zeta_s(X,Y)<\infty$.
- (ii) $\mathbb{E}\,\alpha(X,\dots,X)=\mathbb{E}\,\alpha(Y,\dots,Y)$ *for every* $k=1,\dots,m$ *and every* $\alpha\in L(B^k;\mathbb{R})$.
- (iii) $\mathbb{E}\,X^{\hat{\otimes}k}=\mathbb{E}\,Y^{\hat{\otimes}k}$ *for* $1\leqslant k\leqslant m$.

In particular, if $0<s\leqslant 1$, then (ii) *and* (iii) *are vacuous, and thus $\zeta_s(X,Y)<\infty$ for all such X and Y.*

Proof. If $\alpha\in L(B^k;\mathbb{R})$ is any k-linear form, then $x\mapsto\alpha(x,\dots,x)$ is $\mathcal{B}_w$-measurable by assumption and thus $\alpha(X,\dots,X)$ is measurable. Furthermore, if $1\leqslant k\leqslant m<s$, then $|\alpha(X,\dots,X)|\leqslant\|\alpha\|\|X\|^k=O(1+\|X\|^s)$, and thus $\alpha(X,\dots,X)$ is integrable. Theorem 6.15 shows that $\mathbb{E}\,X^{\hat{\otimes}k}$ exists in Dunford sense for $1\leqslant k\leqslant m$, and the same holds for $\mathbb{E}\,Y^{\hat{\otimes}k}$.

(i) $\implies$ (ii): Suppose that $\alpha\in L(B^k;\mathbb{R})$ with $k\leqslant m$, and let $f(x):=\alpha(x,\dots,x)$. By assumption, f is $\mathcal{B}_w$-measurable, and by Lemma B.1 $f\in\mathrm{Lip}_s(B;\mathbb{R})$ with $\|f\|_{\mathrm{Lip}_s}=0$. Consequently, $f\in\mathrm{Lip}_s^w(B;\mathbb{R})$, and if $\zeta_s(X,Y)<\infty$, then (B.8) yields $\mathbb{E}\,f(X)=\mathbb{E}\,f(Y)$.

(ii) $\implies$ (i) Suppose that $f\in\mathrm{Lip}_s^w(B;\mathbb{R})$ with $\|f\|_{\mathrm{Lip}_s}\leqslant 1$. By (B.5), there exist k-linear forms $\alpha_k=f^{(k)}(0)/k!\in L(B^k;\mathbb{R})$ and a function $g:B\to\mathbb{R}$ with $|g(x)|\leqslant\|x\|^s$ such that

$$f(x)=\sum_{k=0}^{m}\alpha_k(x,\dots,x)+g(x). \tag{B.10}$$

(Here, $\alpha_0=f(0)$ is just a real constant.) By assumption, $f(X)$ and all $\alpha_k(X,\dots,X)$ are measurable, and thus $g(X)$ is measurable, and similarly $g(Y)$ is measurable. Hence we can use the decomposition (B.10) and obtain using (ii) (and the fact that α_0 is a constant), with all terms finite by the

assumptions $\bar{\mathbb{E}}\|X\|^s, \bar{\mathbb{E}}\|Y\|^s < \infty$,

$$\begin{aligned}&\bigl|\mathbb{E}\, f(X) - \mathbb{E}\, f(Y)\bigr| \\ &\quad= \left|\sum_{k=0}^{m} \mathbb{E}\, \alpha_k(X,\dots,X) + \mathbb{E}\, g(X) - \sum_{k=0}^{m} \mathbb{E}\, \alpha_k(Y,\dots,Y) - \mathbb{E}\, g(Y)\right| \\ &\quad= \bigl|\mathbb{E}\, g(X) - \mathbb{E}\, g(Y)\bigr| \leqslant \mathbb{E}\,|g(X)| + \mathbb{E}\,|g(Y)| \leqslant \bar{\mathbb{E}}\|X\|^s + \bar{\mathbb{E}}\|Y\|^s,\end{aligned}$$

Taking the supremum over all such f yields

$$\zeta_s(X,Y) \leqslant \bar{\mathbb{E}}\|X\|^s + \bar{\mathbb{E}}\|Y\|^s < \infty. \tag{B.11}$$

(ii) $\iff$ (iii): This is Corollary 6.17. □

Example B.3. If $0 < s \leqslant 1$, then Lemma B.2 shows that ζ_s is a metric on the set of all probability distributions on $(B, \mathcal{B}_w)$ with a finite s:th moment of the norm $\bar{\mathbb{E}}\|X\|^s$.

If $1 < s \leqslant 2$, and we still assume $\bar{\mathbb{E}}\|X\|^s, \bar{\mathbb{E}}\|Y\|^s < \infty$, Lemma B.2 shows that $\zeta_s(X,Y) < \infty$ if and only if $\mathbb{E}\, X = \mathbb{E}\, Y$. Hence ζ_s is a metric on the set of probability distributions with $\bar{\mathbb{E}}\|X\|^s$ finite and a given expectation. In this case it is natural to work in the set of probability distributions with expectation 0, which easily is achived by subtracting the means from the variables, so this is no serious restriction.

The next case $2 < s \leqslant 3$ is substantially more complicated. Lemma B.2 shows that we need not only $\mathbb{E}\, X = \mathbb{E}\, Y$ but also $\mathbb{E}\, X^{\hat{\otimes}2} = \mathbb{E}\, Y^{\hat{\otimes}2}$. For real-valued random variables, it is standard to obtain this by considering the standardized variable $(\operatorname{Var} X)^{-1/2}(X - \mathbb{E}\, X)$. This extends to finite-dimensional spaces, where $\operatorname{Var} X$ is the covariance matrix, see e.g. [48], but not to infinite-dimensional ones, and this is a serious problem when using ζ_s with $s > 2$ in Banach spaces. Nevertheless, it is at least sometimes possible to modify the variables to achieve the desired exact equality of the second moments, see e.g. [49], and results like Theorems 1.1 and 1.3 then are useful.

For $s > 3$ we need not only equal first and second moments, but also equal third moments $\mathbb{E}\, X^{\hat{\otimes}3} = \mathbb{E}\, Y^{\hat{\otimes}3}$. This can in general not be achieved by any norming, not even for real-valued random variables. If the variables are symmetric, however, all odd moments vanish. Thus, for $s \leqslant 4$, ζ_s is a metric of the set of all symmetric probability distributions with a given second projective moment, and $\bar{\mathbb{E}}\|X\|^s$ finite. It seems likely that there might be applications with symmetric random variables and $3 < s \leqslant 4$, but we do not know of any such cases, or of any other applications of ζ_s with $s > 3$.

The main use of the Zolotarev distances is to prove convergence in distribution; the idea is that if $(X_n)_{n=1}^{\infty}$ is a sequence of B-valued random variables, we can try to prove $X_n \xrightarrow{\mathrm{d}} X$ (for a suitable random variable X in B) by first proving

$$\zeta(X_n, X) \to 0. \tag{B.12}$$

It turns out that (B.12) is by itself sufficient for convergence in distributions in some Banach spaces, for example when B is a Hilbert space [31, 19], but

not always, for example not in $C[0,1]$ where extra conditions are needed [49, 60].

Example B.4. One example where this approach has been particularly successful is the contraction method used to prove convergence in distribution of a sequence X_n of random variables when there is a recursive structure of the type

$$X_n \overset{\mathrm{d}}{=} \sum_{r=1}^{K} A_{r,n} X_{I_{r,n}}^{(r)} + b_n, \tag{B.13}$$

where $(X_j^{(1)})_{j=0}^{\infty}, \dots, (X_j^{(K)})_{j=0}^{\infty}$ are i.i.d. copies of $(X_n)_{n=0}^{\infty}$, and $A_{r,n}$, $I_{r,n}$ and b_n are given random variables independent of all $X_j^{(r)}$. (The idea is, roughly, to first find a good candidate X for the limit by formally letting $n \to \infty$ in (B.13), assuming that $A_{r,n}$, $I_{r,n}$ and b_n converge in some suitable way. Then one uses (B.13) to obtain a recursive estimate of the distance $\zeta_s(X_n, X)$, and use this to show that $\zeta_s(X_n, X) \to 0$.) This method was introduced for real-valued random variables by Rösler [54, 55, 56], and has been extended to variables with values in $\mathbb{R}^d$ [47, 48], $L^p[0,1]$ [25], Hilbert spaces [19], $C[0,1]$ and $D[0,1]$ [49, 60]; see further [49, 60] and the further references given there. These papers use not only the Zolotarev distance ζ_s considered here but also some other probability metrics; nevertheless the Zolotarev distances are essential in several of the applications. Moreover, it is easily verified that for any real constant t,

$$\zeta_s(tX, tY) = |t|^s \zeta_s(X, Y) \tag{B.14}$$

and, for any bounded linear operator $T \in L(B; B)$,

$$\zeta_s(TX, TY) \leqslant \|T\|^s \zeta_s(X, Y). \tag{B.15}$$

In applications, these relations are typically applied with $|t|$ or $\|T\|$ small, and it is then advantageous to take s large. In fact, in some applications (see e.g. [49]) one is forced to take $s > 2$ in order to obtain the required estimates, and then, as seen in Lemma B.2 and Example B.3, it is essential to have (or arrange) equalities of the second moments $\mathbb{E}\, X_n^{\hat{\otimes} 2} = \mathbb{E}\, X^{\hat{\otimes} 2}$ in order for $\zeta_s(X_n, X) < \infty$ (which is necessary in order to even start the recursion sketched above).

References

[1] K. S. Alexander, Characterization of the cluster set of the LIL sequence in Banach space. *Ann. Probab.* **17** (1989), no. 2, 737–759.

[2] K. S. Alexander, Unusual cluster sets for the LIL sequence in Banach space. *Ann. Probab.* **17** (1989), no. 3, 1170–1185.

[3] N. Aronszajn, Theory of reproducing kernels. *Trans. Amer. Math. Soc.* **68** (1950), 337–404.

[4] S. Banach & C. Kuratowski, Sur une généralisation du problème de la mesure. *Fundamenta Mathematicae*, **14** (1929), 127–131.

[5] P. Billingsley, *Convergence of Probability Measures*. 1st ed., Wiley, New York, 1968.
[6] P. Billingsley, *Convergence of Probability Measures*. 2nd ed., Wiley, New York, 1999.
[7] G. Birkhoff, Integration of functions with values in a Banach space, *Trans. Amer. Math. Soc.* **38** (1935), no. 2, 357–378.
[8] R. Blei, *Analysis in Integer and Fractional Dimensions*. Cambridge University Press, Cambridge, 2001.
[9] H. F. Bohnenblust & E. Hille, On the absolute convergence of Dirichlet series. *Ann. Math.* **32** (1931), no. 3, 600–622.
[10] T. Carleman, *Les fonctions quasi analytiques*. Gauthier-Villars, Paris, 1926.
[11] H. Cartan, *Calcul différentiel*. Hermann, Paris, 1967.
[12] D. L. Cohn, *Measure Theory*, Birkhäuser, Boston, 1980.
[13] J. B. Conway, *A Course in Functional Analysis*. 2nd ed., Springer-Verlag, New York, 1990.
[14] H. H. Corson, The weak topology of a Banach space. *Trans. Amer. Math. Soc.* **101** (1961), 1–15.
[15] A. M. Davie, The approximation problem for Banach spaces. *Bull. London Math. Soc.* **5** (1973), 261–266.
[16] A. M. Davie, The Banach approximation problem. *J. Approximation Theory* **13** (1975), 392–394.
[17] J. Diestel, *Sequences and Series in Banach Spaces*. Springer-Verlag, New York, 1984.
[18] J. Diestel & J. J. Uhl, jr, *Vector Measures*. American Mathematical Society, Providence, R.I., 1977.
[19] M. Drmota, S. Janson & R. Neininger, A functional limit theorem for the profile of search trees. *Ann. Appl. Probab.* **18** (2008), 288–333.
[20] R. M. Dudley, Weak convergences of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.* **10** (1966), 109–126.
[21] R. M. Dudley, Measures on non-separable metric spaces. *Illinois J. Math.* **11** (1967), 449–453.
[22] N. Dunford & J. T. Schwartz, *Linear Operators. I. General Theory*. Interscience Publishers, New York, 1958.
[23] G. A. Edgar, Measurability in a Banach space. *Indiana Univ. Math. J.* **26** (1977), no. 4, 663–677.
[24] G. A. Edgar, Measurability in a Banach space. II. *Indiana Univ. Math. J.* **28** (1979), no. 4, 559–579.
[25] K. Eickmeyer & L. Rüschendorf, A limit theorem for recursively defined processes in L^p. *Statist. Decisions* **25** (2007), no. 3, 217–235.
[26] P. Enflo, A counterexample to the approximation problem in Banach spaces. *Acta Math.* **130** (1973), 309–317.
[27] R. Engelking, *General topology*. 2nd ed., Heldermann Verlag, Berlin, 1989.

[28] D. H. Fremlin & J. Mendoza, On the integration of vector-valued functions. *Illinois J. Math.* **38** (1994), no. 1, 127–147.
[29] D. H. Fremlin & M. Talagrand, A decomposition theorem for additive set-functions, with applications to Pettis integrals and ergodic means. *Math. Z.* **168** (1979), no. 2, 117–142.
[30] L. Gillman & M. Jerison, *Rings of Continuous Functions*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960.
[31] E. Giné & J. R. León, On the central limit theorem in Hilbert space. *Stochastica* **4** (1980), no. 1, 43–71.
[32] A. Grothendieck, Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo* **8** (1953), 1–79.
[33] A. Grothendieck, Produits tensoriels topologiques et espaces nucléaires. *Mem. Amer. Math. Soc.* **16** (1955).
[34] A. Gut, *Probability: A Graduate Course*, Springer, New York, 2005.
[35] T. H. Hildebrandt, Integration in abstract spaces. *Bull. Amer. Math. Soc.* **59**, (1953). 111–139.
[36] R. Huff, Remarks on Pettis integrability. *Proc. Amer. Math. Soc.* **96** (1986), no. 3, 402–404.
[37] S. Janson, *Gaussian Hilbert Spaces*. Cambridge Univ. Press, Cambridge, 1997.
[38] O. Kallenberg, *Foundations of Modern Probability*. 2nd ed., Springer, New York, 2002.
[39] A. Kanamori, *The Higher Infinite. Large Cardinals in Set Theory from their Beginnings*. Springer-Verlag, Berlin, 1994.
[40] K. Kuratowski & A. Mostowski, *Set Theory*. PWN, Warszawa, and North-Holland, Amsterdam, 1968.
[41] P. Lax, *Functional Analysis*. Wiley-Interscience, New York, 2002.
[42] M. Ledoux & M. Talagrand, *Probability in Banach Spaces*. Springer-Verlag, Berlin, 1991.
[43] J. Lindenstrauss & L. Tzafriri, *Classical Banach Spaces. I. Sequence Spaces*. Springer-Verlag, Berlin-New York, 1977.
[44] J. E. Littlewood, On bounded bilinear forms in an infinite number of variables, *Quart. J. Math. Oxford* **1** (1930), 164–174.
[45] E. Marczewski & R. Sikorski, Measures in non-separable metric spaces. *Colloquium Math.* **1** (1948), 133–139.
[46] A. A. Miljutin, Isomorphism of the spaces of continuous functions over compact sets of the cardinality of the continuum. (Russian) *Teor. Funkciĭ Funkcional. Anal. i Priložen. Vyp.* **2** (1966), 150–156.
[47] R. Neininger, On a multivariate contraction method for random recursive structures with applications to Quicksort. Random Structures Algorithms 19 (2001), no. 3–4, 498–524.
[48] R. Neininger & L. Rüschendorf, A general limit theorem for recursive algorithms and combinatorial structures. *Ann. Appl. Probab.* **14** (2004), no. 1, 378–418.

[49] R Neininger & H. Sulzbach, On a functional contraction method. Preprint, 2012. arXiv:1202.1370.

[50] A. Pełczyński, Linear extensions, linear averagings, and their applications to linear topological classification of spaces of continuous functions. *Dissertationes Math. Rozprawy Mat.* **58**, 1968.

[51] W. R. Pestman, Measurability of linear operators in the Skorokhod topology. *Bull. Belg. Math. Soc. Simon Stevin* **2** (1995), no. 4, 381–388.

[52] G. Pisier, *Factorization of Linear Operators and Geometry of Banach Spaces*. Amer. Math. Soc., Providence, RI, 1986.

[53] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge University Press, Cambridge, 1989.

[54] U. Rösler, A limit theorem for "Quicksort". *RAIRO Inform. Théor. Appl.* **25** (1991), no. 1, 85–100.

[55] U. Rösler, A fixed point theorem for distributions. *Stochastic Process. Appl.* **42** (1992), no. 2, 195–214.

[56] U. Rösler, On the analysis of stochastic divide and conquer algorithms. *Algorithmica* **29** (2001), 238–261.

[57] R. A. Ryan, *Introduction to Tensor Products of Banach Spaces*. Springer-Verlag, London, 2002.

[58] R. H. Sorgenfrey, On the topological product of paracompact spaces. *Bull. Amer. Math. Soc.* **53** (1947), 631–632.

[59] G. F. Stefánsson, Pettis integrability. *Trans. Amer. Math. Soc.* **330** (1992), no. 1, 401–418.

[60] H. Sulzbach, *On a functional contraction method*. Ph. D. thesis, Goethe-Universität, Frankfurt am Main, 2012.

[61] A. Szankowski, $B(H)$ does not have the approximation property. *Acta Math.* **147** (1981), no. 1–2, 89–108.

[62] M. Talagrand, Pettis integral and measure theory. *Mem. Amer. Math. Soc.* **51** (1984), no. 307.

[63] F. Treves, *Topological Vector Spaces, Distributions and Kernels*. Academic Press, New York, 1967.

[64] S. Ulam, Zur Masstheorie in der allgemeinen Mengenlehre. *Fundamenta Mathematicae* **16** (1930), 140–150.

[65] N. Th. Varopoulos, Tensor algebras and harmonic analysis. *Acta Math.* **119** (1967), 51–112.

[66] V. M. Zolotarev, Metric distances in spaces of random variables and their distributions. (Russian.) *Mat. Sb. (N.S.)* **101(143)** (1976), no. 3, 416–454, 456. English transl.: *Math. USSR-Sb.* **30** (1976), no. 3, 373–401.

[67] V. M. Zolotarev, Approximation of the distributions of sums of independent random variables with values in infinite-dimensional spaces. (Russian.) *Teor. Veroyatnost. i Primenen.* **21** (1976), no. 4, 741–758. Erratum *ibid* **22** (1977), no. 4, 901. English transl.: *Theory Probab. Appl.* **21** (1976), no. 4, 721–737 (1977); *ibid* **22** (1977), no. 4, 881

(1978).
[68] V. M. Zolotarev, Ideal metrics in the problem of approximating distributions of sums of independent random variables. (Russian.) *Teor. Verojatnost. i Primenen.* **22** (1977), no. 3, 449–465. English transl.: *Theor. Probability Appl.* **22** (1977), no. 3, 433–449.
[69] V. M. Zolotarev, Ideal metrics in the problems of probability theory and mathematical statistics. *Austral. J. Statist.* **21** (1979), no. 3, 193–208. Corrigenda: *ibid* **22** (1980), no. 1, 97.
[70] A. Zygmund, *Trigonometric Series*. 2nd ed., Cambridge Univ. Press, Cambridge, 1959.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: svante.janson@math.uu.se

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: sten.kaijser@math.uu.se