THE SPACE D IN SEVERAL VARIABLES: RANDOM VARIABLES AND HIGHER MOMENTS

SVANTE JANSON

Abstract. We study the Banach space $D([0,1]^m)$ of functions of several variables that are (in a certain sense) right-continuous with left limits, and extend several results previously known for the standard case $m = 1$. We give, for example, a description of the dual space, and we show that a bounded multilinear form always is measurable with respect to the σ-field generated by the point evaluations. These results are used to study random functions in the space. (I.e., random elements of the space.) In particular, we give results on existence of moments (in different senses) of such random functions, and we give an application to the Zolotarev distance between two such random functions.

1. Introduction

Recall that $D([0,1])$ is the set of real-valued functions on $I := [0,1]$ that are right-continuous with left limits, see e.g. [1, Chapter 3]. Similarly, the m-dimensional analogue $D([0,1]^m)$ is defined as the set of real-valued functions f on $[0,1]^m$ such that at every $t = (t_1, \dots, t_m) \in [0,1]^m$, the limit of $f(s)$ exists (as a finite real number), as $s \to t$ in any of the octants of the form $J_1 \times \cdots \times J_m$ where each J_i is either $[t_i, 1]$ or $[0, t_i)$ (the latter only if $t_i > 0$). For example, take $m = 2$ for notational convenience; then $f \in D([0,1]^2)$ if and only if, for each $(x, y) \in [0,1]^2$, the limits

$$\begin{aligned}
f(x+, y+) &:= \lim_{\substack{x' \to x,\, x' \geqslant x \\ y' \to y,\, y' \geqslant y}} f(x', y'), \\
f(x+, y-) &:= \lim_{\substack{x' \to x,\, x' \geqslant x \\ y' \to y,\, y' < y}} f(x', y'), \\
f(x-, y+) &:= \lim_{\substack{x' \to x,\, x' < x \\ y' \to y,\, y' \geqslant y}} f(x', y'), \\
f(x-, y-) &:= \lim_{\substack{x' \to x,\, x' < x \\ y' \to y,\, y' < y}} f(x', y')
\end{aligned}$$

exist, except that we ignore all cases with an argument $0-$. Note the slight asymmetry; we use $\geqslant$ but $<$. Note also that necessarily $f(x+, y+) = f(x, y)$ when the limit exists.

The space $D([0,1]^m)$ was studied by Wichura [17, 18] and Neuhaus [13]; the latter extended the definition of the Skorohod topology from the case $m = 1$ and proved many basic results on it. (The definition of the space

Date: 1 April, 2020.
2010 *Mathematics Subject Classification.* 60B11; 46B28, 46E15, 46G10, 46J10.
Partly supported by the Knut and Alice Wallenberg Foundation.

in [13] differs slightly from the one above at the top and right parts of the boundary; this is not essential and his proofs are just as valid for the version considered here.) See also Straf [16] for an even more general setting.

In the present paper we study $D([0,1]^m)$ from a different point of view, as a normed (Banach) space. The space $D([0,1])$ was studied as a normed space in [11, Chapter 9] (together with $D([0,1]^m)$ to some minor extent) in order to show some results on second and higher moments of $D([0,1])$-valued random variables; these results were at least partly motivated by an application [12] where convergence in distribution of some $D([0,1])$-valued random variables was shown by the contraction method, which required some of these results as technical tools. The purpose of the present paper is to extend some of these results for $D([0,1])$ to $D([0,1]^m)$; one motivation is that this enables similar applications of the contraction method to $D([0,1]^m)$-valued random variables, see [3].

Functions in $D([0,1]^m)$ are bounded, and we define

$$\|f\| := \sup_{t\in[0,1]^m} |f(t)|. \tag{1.1}$$

$D([0,1]^m)$ is a Banach space with this norm. Note that the Banach space $D([0,1]^m)$ is not separable (already for $m = 1$), and that the space $C([0,1]^m)$ of continuous functions on $[0,1]^m$ is a closed, separable subspace.

Remark 1.1. We consider, for definiteness, real-valued functions. The definitions and results extend with no or trivial modifications to complex-valued functions and measures. It is also easy to extend the results to vector-valued functions with values in a fixed, finite-dimensional vector space.

Acknowledgement. I am indebted to Henning Sulzbach for initiating this work by asking me questions that led to it, and also for helpful comments; this almost led to a joint paper.

2. Preliminaries

2.1. **The split interval.** Define the *split interval* $\widehat{I}$ as the set consisting of two copies, t and $t-$, of every point in $(0,1]$, together with a single 0. There is a natural total order on $\widehat{I}$, with $s < t- < t$ when $s,t \in [0,1]$ with $s < t$. We define intervals in $\widehat{I}$ in the usual way, using this order, and equip $\widehat{I}$ with the order topology, which has a base consisting of all open intervals $[0,x)$, $(x,1]$, and (x,y) with $x,y \in \widehat{I}$; then $\widehat{I}$ is a compact Hausdorff space; see e.g. [6, Problems 1.7.4 and 3.12.3]. The compact space $\widehat{I}$ is separable (i.e., has a countable dense subset, for example the rational numbers) and first countable (every point has a countable neighbourhood basis), but not second countable and not metrizable, see e.g. [11, Section 9.2].

We regard $[0,1]$ as a subset of $\widehat{I}$, with the inclusion mapping $\iota : [0,1] \to \widehat{I}$ given by $\iota(t) = t$. This mapping is *not* continuous; the subspace topology on I induced by $\widehat{I}$ is stronger than the usual topology on I (which we continue to use for I). (The induced topology on $(0,1)$ yields a version of the *Sorgenfrey line*, see [6, Examples 1.2.2, 2.3.12, 3.8.14, 5.1.31].)

Every function $f \in D([0,1])$ has a (unique) extension to a continuous function on $\widehat{I}$, given by $f(t-) = \lim_{s\nearrow t} f(s)$. Conversely, the restriction to

I of any continuous function on $\widehat{I}$ is a function in $D([0,1])$. There is thus a bijection $D([0,1]) \cong C(\widehat{I})$, which is an isometric isomorphism as Banach spaces [5]. Another way to see this is to note that $D([0,1])$ is a Banach algebra with $\widehat{I}$ as its maximal ideal space, and that the Gelfand transform is this isomorphism $D([0,1]) \cong C(\widehat{I})$, see [11, Section 9.2].

These results extend immediately to $D([0,1]^m)$. The definition of $D([0,1]^m)$ shows that every function $f \in D([0,1]^m)$ has a (unique) extension to a continuous function on $\widehat{I}^m$, and, conversely, that the restriction to I^m of any continuous function on $\widehat{I}^m$ is a function in $D([0,1]^m)$; hence, there is a bijection $D([0,1]^m) \cong C(\widehat{I}^m)$, which is an isometric isomorphism as Banach spaces. Again, this can be regarded as the Gelfand transform for the Banach algebra $D([0,1]^m)$, with maximal ideal space $\widehat{I}^m$.

2.2. **Tensor products.** For definitions and basic properties of the injective and projective tensor products $X \check{\otimes} Y$ and $X \hat{\otimes} Y$ of two Banach spaces X and Y see e.g. [15], or the summary in [11].

In particular, recall that if K is a compact Hausdorff space, then $C(K)^{\check{\otimes} k} \cong C(K^k)$ (isometrically) by the natural identification of $f_1 \otimes \cdots \otimes f_k$ with the function

$$f_1 \otimes \cdots \otimes f_k(x_1, \ldots, x_k) := \prod_{i=1}^{k} f_i(x_i) \tag{2.1}$$

on K^k. In particular, linear combinations of such functions $\bigotimes_1^k f_i$ are dense in $C(K^k)$. Furthermore, $C(K)$ has the approximation property (see [11, Chapter 4]), and as a consequence, $C(K)^{\hat{\otimes} k}$ can be regarded as a linear subspace of $C(K)^{\check{\otimes} k} = C(K^k)$ (with a different norm).

Since $D([0,1]^m) \cong C(\widehat{I}^m)$, these results apply also to $D([0,1]^m)$. In particular, $D([0,1]^m)^{\check{\otimes} k} \cong C(\widehat{I}^{mk}) \cong D([0,1]^{mk})$, again by the natural identification (2.1) of $\bigotimes_1^k f_i$ and $\prod_{i=1}^k f_i(x_i)$. (From now on, we identify these spaces and write $=$ instead of $\cong$.) In particular, linear combinations of functions $\bigotimes_1^m f_i = \prod_{i=1}^m f_i(x_i)$ with $f_i \in D([0,1])$ are dense in $D([0,1]^m)$. Furthermore, $D([0,1]^m)$ has the approximation property and thus $D([0,1]^m)^{\hat{\otimes} k} \subset D([0,1]^m)^{\check{\otimes} k} = D([0,1]^{mk})$. Note that $D([0,1]^m)^{\hat{\otimes} k}$ is *not* a closed subspace of $D([0,1]^m)^{\check{\otimes} k} = D([0,1]^{mk})$ when $k \geqslant 2$, and thus the projective and injective norms are not equivalent on $D([0,1]^m)^{\hat{\otimes} k}$; see e.g. [11, Remark 7.9 and Theorem 9.27].

2.3. **Baire sets and measures.** If K is a topological space, then the *Borel σ-field* $\mathcal{B}(K)$ is the σ-field generated by the open sets in K, and if K is a compact Hausdorff space (the only case that we consider), the *Baire σ-field* $\mathsf{Ba}(K)$ is the σ-field generated by the continuous real-valued functions on K; see e.g. [2, §6.3], [9, §51] (with a somewhat different definition, equivalent in the compact case) and [4, Exercises 7.2.8–13]. Elements of $\mathcal{B}(K)$ are called *Borel sets* and elements of $\mathsf{Ba}(K)$ are called *Baire sets*. A *Baire (Borel) measure on* K is a measure on $\mathsf{Ba}(K)$ $(\mathcal{B}(K))$, and similarly for signed measures; we consider in this paper only finite measures.

We collect some basic properties.

Lemma 2.1. *Let K, K_1, K_2 be compact Hausdorff spaces.*

(i) $\mathsf{Ba}(K) \subseteq \mathcal{B}(K)$.
(ii) *If K is a compact metric space, then* $\mathsf{Ba}(K) = \mathcal{B}(K)$.
(iii) $\mathsf{Ba}(K_1 \times K_2) = \mathsf{Ba}(K_1) \times \mathsf{Ba}(K_2)$.
(iv) *If $(S, \mathcal{S})$ is any measurable space, then a function $f : (S, \mathcal{S}) \to (K, \mathsf{Ba}(K))$ is measurable if and only if $g \circ f : (S, \mathcal{S}) \to (\mathbb{R}, \mathcal{B})$ is measurable for every $g \in C(K)$.*
(v) *If $f : K_1 \to K_2$ is continuous, then f is Baire measurable, i.e., $f : (K_1, \mathsf{Ba}(K_2)) \to (K_2, \mathsf{Ba}(K_2))$ is measurable.*

Proof. (i) and (ii) are easy and well-known; for (iii) see [9, Theorem 51E] or [2, Lemma 6.4.2]. (iv) is a consequence of the definition of $\mathsf{Ba}(K)$, and (v) follows. □

By Lemma 2.1(ii), there is no reason to study Baire sets instead of the perhaps more well-known Borel sets for metrizable compact spaces, since they coincide. However, we shall mainly study non-metrizable compact spaces such as $\widehat{I}$, and then the Baire σ-field is often better behaved than the Borel σ-field. One example is seen in Lemma 2.1(iii) above; the corresponding result for Borel σ-fields is not true in general, and in particular not for $K = \widehat{I}$, see Proposition A.4. Another important example is the Riesz representation theorem, which takes the following simple form using Baire measures. Let $M_{\mathsf{Ba}}(K)$ be the Banach space of signed Baire measures on K, with $\|\mu\|$ the total variation of μ, i.e., $\|\mu\| := |\mu|(K)$, where the measure $|\mu|$ is the variation of μ.

Proposition 2.2 (The Riesz representation theorem)**.** *Let K be a compact Hausdorff space. There is an isometric bijection between the space $C(K)^*$ of bounded continuous linear functionals on $C(K)$ and the space $M_{\mathsf{Ba}}(K)$ of signed Baire measures on K, where a signed Baire measure μ corresponds to the linear functional $f \mapsto \int_K f \, \mathrm{d}\mu$.*

Remark 2.3. The Riesz representation theorem is perhaps more often stated in a version using Borel measures, but then one has to restrict to *regular* signed Borel measures, see e.g. [4, Theorem 7.3.5] or [2, Theorem 7.10.4]. The connection between the two versions is that every (signed) Baire measure on K has a unique extension to a regular (signed) Borel measure, see [9, Theorem 54.D] or [2, Corollary 7.3.4].

For a proof of Proposition 2.2, see [9, §56], or the references in Remark 2.3 above.

Example 2.4. Although not needed for our results, it is interesting to note that $\mathsf{Ba}(\widehat{I}) = \mathcal{B}(\widehat{I})$, but $\mathsf{Ba}(\widehat{I}^m) \subsetneq \mathcal{B}(\widehat{I}^m)$, when $m \geqslant 2$. See Appendix A.

2.4. **Some further notation.** Let $[m] := \{1, \dots, m\}$.

If $x \in \mathbb{R}$, then $\lfloor x \rfloor$ and $\lceil x \rceil$ denote x rounded down or up to the nearest integer, respectively.

Recall that $t-$ is a point in $\widehat{I} \setminus I$ for $t \in (0, 1]$. For completeness we define $0- := 0$.

Recall also that $\iota : I \to \widehat{I}$ denotes the inclusion mapping. Conversely, define the projection $\rho : \widehat{I} \to I$ by $\rho(t) = t$ and $\rho(t-) = t$ for $t \in [0,1]$. Let

$$\phi := \iota \circ \rho : \widehat{I} \to \widehat{I} \tag{2.2}$$

be the composition of ι and ρ, i.e., the projection

$$\begin{cases} \phi(t) = t, \\ \phi(t-) = t. \end{cases} \tag{2.3}$$

Note that $\phi \circ \phi = \phi$, i.e., ϕ is a projection map.

If $A \subseteq [0,1]$, let $A- := \{t- : t \in A\} \subset \widehat{I}$, and $\widehat{A} := \rho^{-1}(A) = A \cup (A-) \subset \widehat{I}$. In particular, if $s \in [0,1]$, then $\widehat{\{s\}} = \{s, s-\}$.

We sometimes denote elements of $\widehat{I}^m$ by $\hat{\mathbf{t}} = (\hat{t}_1, \dots, \hat{t}_m)$. Let $\pi_i : \widehat{I}^m \to \widehat{I}$ denote the projection on the i-th coordinate: $\pi_i(\hat{\mathbf{t}}) = \hat{t}_i$.

If $f \in D([0,1])$ and $t \in (0,1]$, let

$$\Delta f(t) := f(t) - f(t-). \tag{2.4}$$

This defines a bounded linear map $\Delta : D([0,1]) \to c_0((0,1])$, with norm $\|\Delta\| = 2$; see [11, Theorem 9.1] for further properties.

We extend this to several dimensions by defining, for $f \in D([0,1]^m)$ and $i \in [m] := \{1, \dots, m\}$,

$$\Delta_i f(t_1, \dots, t_m) := f(t_1, \dots, t_m) - f(t_1, \dots, t_i-, \dots, t_m), \tag{2.5}$$

i.e., the jump along the i-th coordinate at $t = (t_1, \dots, t_m)$. (This is 0 when $t_i = 0$, by our definition $0- = 0$.)

We further define, for any subset $J = \{j_1, \dots, j_\ell\} \subseteq [m]$,

$$\Delta_J f := \Delta_{j_1} \cdots \Delta_{j_\ell} f. \tag{2.6}$$

Note that the operators Δ_i commute, so their order in (2.6) does not matter.

Remark 2.5. In particular, (2.5) shows that for $f_1, \dots, f_m \in D([0,1])$,

$$\Delta_i(f_1 \otimes \cdots \otimes f_m) = f_1 \otimes \ldots \otimes (\Delta f_i) \otimes \ldots \otimes f_m. \tag{2.7}$$

Consequently, identifying $D([0,1]^m)$ and $D([0,1])^{\check{\otimes} m}$ as in Section 2.2,

$$\Delta_i = \mathrm{I} \check{\otimes} \cdots \check{\otimes} \mathrm{I} \check{\otimes} \Delta \check{\otimes} \mathrm{I} \check{\otimes} \cdots \check{\otimes} \mathrm{I}, \tag{2.8}$$

where I is the identity operator and there is a single Δ in the i-th position. Thus Δ_i can be regarded as a bounded linear map into $D([0,1]) \check{\otimes} \cdots \check{\otimes} c_0((0,1]) \check{\otimes} \cdots \check{\otimes} D([0,1])$, and similarly for Δ_J. However, we will not use this point of view; we just regard $\Delta_i f$ and $\Delta_J f$ as the functions on I^m given by (2.5)–(2.6).

3. Some projections

Recall the mappings ι, ρ and ϕ from Section 2.4.

Lemma 3.1. (i) *$\iota : I \to \widehat{I}$ is Baire measurable (but not continuous).*

(ii) *$\rho : \widehat{I} \to I$ is continuous, and thus Baire measurable.*

(iii) *$\phi : \widehat{I} \to \widehat{I}$ is Baire measurable (but not continuous).*

Proof. (i): We have already remarked that ι is not continuous.

To see that ι is Baire measurable, i.e., that $\iota : (I, \mathsf{Ba}) \to (\widehat{I}, \mathsf{Ba})$ is measurable, let $g \in C(\widehat{I})$. Then $g \circ \iota \in D([0,1])$, see Section 2.1, and thus $g \circ \iota : (I, \mathsf{Ba}) = (I, \mathcal{B}) \to (\mathbb{R}, \mathcal{B})$ is measurable. Thus ι is Baire measurable by Lemma 2.1(iv).

(ii): The continuity follows from the definitions.

Alternatively, we may note that if $f \in C(I)$, then $f \in D(I)$, so it has by Section 2.1 a continuous extension (which also is its Gelfand transform) $\hat{f} \in C(\widehat{I})$, given by $\hat{f}(t-) = f(t-) = f(t)$; hence $\hat{f} = f \circ \rho$. In particular, taking f to be the identity i with $i(x) = x$, we have $\hat{i} = \rho$, and thus $\rho \in C(\widehat{I})$.

(ii): That $\phi = \iota \circ \rho$ is Baire measurable follows by (i) and (ii). To see that ϕ is not continuous, it suffices to note that $\phi(\widehat{I}) = I$ is a proper dense subset of $\widehat{I}$, and thus not a compact subset of $\widehat{I}$. $\square$

For a fixed m and $1 \leqslant i \leqslant m$, define $\phi_i = \phi_{i,m} : \widehat{I}^m \to \widehat{I}^m$ by

$$\phi_i(\hat{t}_1, \dots, \hat{t}_m) := (\hat{t}_1, \dots, \phi(\hat{t}_i), \dots, \hat{t}_m) \tag{3.1}$$

(with the identity in all coordinates except the i-th). Then $\phi_1, \dots, \phi_m$ are commuting projections $\widehat{I}^m \to \widehat{I}^m$.

Since ϕ is Baire measurable, and $\mathsf{Ba}(\widehat{I}^m) = \mathsf{Ba}(\widehat{I})^m$ by Lemma 2.1(iii), each ϕ_i is Baire measurable. Hence, ϕ_i induces a map $\Phi_i : M_{\mathsf{Ba}}(\widehat{I}^m) \to M_{\mathsf{Ba}}(\widehat{I}^m)$ such that

$$\int_{\widehat{I}^m} f \,\mathrm{d}\Phi_i(\mu) = \int_{\widehat{I}^m} f \circ \phi_i \,\mathrm{d}\mu, \tag{3.2}$$

for Baire measurable and, say, bounded $f : \widehat{I}^m \to \mathbb{R}$. Letting $\tau^\sharp$ denote the map $M_{\mathsf{Ba}} \to M_{\mathsf{Ba}}$ induced as in (3.2) by a function $\tau : \widehat{I}^m \to \widehat{I}^m$, we thus have $\Phi_i = \phi_i^\sharp$, and $\Phi_i \circ \Phi_i = \phi_i^\sharp \circ \phi_i^\sharp = (\phi_i \circ \phi_i)^\sharp = \phi_i^\sharp = \Phi_i$. Hence, Φ_i is a projection in $M_{\mathsf{Ba}}(\widehat{I}^m)$. Similarly, $\Phi_i \circ \Phi_j = \Phi_j \circ \Phi_i$, so the projections Φ_i commute (beacuse the projections ϕ_i do).

Let $\Psi_i := \mathrm{I} - \Phi_i$, where I is the identity operator; thus, by (3.2),

$$\int_{\widehat{I}^m} f \,\mathrm{d}\Psi_i(\mu) = \int_{\widehat{I}^m} \bigl(f(\hat{\mathbf{t}}) - f(\phi_i(\hat{\mathbf{t}}))\bigr) \,\mathrm{d}\mu(\hat{\mathbf{t}}), \tag{3.3}$$

for bounded Baire measurable f on $\widehat{I}^m$.

Note that Ψ_i also is a projection in $M_{\mathsf{Ba}}(\widehat{I}^m)$. It follows immediately that $\Phi_1, \dots, \Phi_m$ and $\Psi_1, \dots, \Psi_m$ are commuting projections in $M_{\mathsf{Ba}}(\widehat{I}^m)$. Furthermore, for any $\mu \in M_{\mathsf{Ba}}(\widehat{I}^m)$,

$$\|\Phi_i(\mu)\| \leqslant \|\mu\|, \tag{3.4}$$

$$\|\Psi_i(\mu)\| \leqslant \|\mu\| + \|\Phi_i(\mu)\| \leqslant 2\|\mu\|. \tag{3.5}$$

Lemma 3.2. *If $\mu \in M_{\mathsf{Ba}}(\widehat{I}^m)$, then for each $i \in \{1, \dots, m\}$ there is a countable subset $A_i \subset (0,1]$ such that $\Psi_i(\mu)$ is supported on the set $\pi_i^{-1}(\widehat{A}_i) = \{(t_1, \dots, t_m) \in \widehat{I}^m : t_i \in \widehat{A}_i\}$.*

Furthermore, if $s \in (0,1]$, let $\Psi_i(\mu)_{s-}$ and $\Psi_i(\mu)_s$ denote the restrictions of $\Psi_i(\mu)$ to $\pi_i^{-1}(s-)$ and $\pi_i^{-1}(s)$, respectively, regarded as measures on $\widehat{I}^{m-1}$;

then

$$\Psi_i(\mu)_{s-} = -\Psi_i(\mu)_s. \tag{3.6}$$

Proof. For notational convenience, assume $i = 1$, and let $\bar\mu := \Psi_1(\mu)$.

For each $s \in (0,1]$, let $E_s := \pi_1^{-1}(\widehat{\{s\}}) = \{s-, s\} \times \widehat{I}^{m-1}$, and let $\bar\mu_s$ denote the restriction of $\bar\mu$ to E_s. (E_s is a Baire set in $\widehat{I}^m$, since $\{s\}$ is measurable in I and π_1 and ρ are Baire measurable as shown above; hence $\bar\mu_s$ is well defined.)

For any finite set $F \subset (0,1]$,

$$\sum_{s\in F} \|\bar\mu_s\| = \sum_{s\in F} |\bar\mu|(E_s) = |\bar\mu|\Bigl(\bigcup_{s\in F} E_s\Bigr) \leqslant |\bar\mu|(\widehat{I}^m) = \|\bar\mu\|. \tag{3.7}$$

Hence, $\sum_{s\in(0,1]} \|\bar\mu_s\| \leqslant \|\bar\mu\|$ and $\bar\mu_s \neq 0$ only for countably many s.

Let $A := \{s : \bar\mu_s \neq 0\}$ and $\nu := \sum_{s\in(0,1]} \bar\mu_s = \sum_{s\in A} \bar\mu_s$, where the sum converges in $M_{\mathsf{Ba}}(\widehat{I}^m)$ by (3.7). We shall prove that $\bar\mu = \nu$. In order to show this, recall that $C(\widehat{I}^m) = C(\widehat{I}) \check{\otimes} C(\widehat{I}^{m-1})$, see Section 2.2, and thus linear combinations of functions f of the form

$$f(\hat t_1, \dots, \hat t_m) = f_1(\hat t_1) g(\hat t_2, \dots, \hat t_m), \qquad f_1 \in C(\widehat{I}),\ g \in C(\widehat{I}^{m-1}), \tag{3.8}$$

are dense in $C(\widehat{I}^m)$. Hence, it suffices to show that

$$\int_{\widehat{I}^m} f \, \mathrm{d}\bar\mu = \int_{\widehat{I}^m} f \, \mathrm{d}\nu \tag{3.9}$$

for every f as in (3.8). For such f, (3.1) and (2.3) yield, for $\hat{\mathbf{t}} = (\hat t_1, \dots, \hat t_m) \in \widehat{I}^m$,

$$\begin{aligned} f(\hat{\mathbf{t}}) - f(\phi_1(\hat{\mathbf{t}})) &= \bigl(f_1(\hat t_1) - f_1(\phi(\hat t_1))\bigr) g(\hat t_2, \dots, \hat t_m) \\ &= \begin{cases} 0, & \hat t_1 = s \in I, \\ -\Delta f_1(s)\, g(\hat t_2, \dots, \hat t_m), & \hat t_1 = s-. \end{cases} \end{aligned} \tag{3.10}$$

Recall that $f_1 \in C(\widehat{I}) = D([0,1])$; regard f_1 as an element of $D([0,1])$ and let $D_{f_1} \subset (0,1]$ be the countable set of discontinuities of f_1. Then, by (3.10), $f(\hat{\mathbf{t}}) - f(\phi_1(\hat{\mathbf{t}})) = 0$ unless $\hat t_1 = u-$ for some $u \in D_{f_1}$, and then $\hat{\mathbf{t}} \in E_u$. Consequently, (3.3) yields

$$\int_{\widehat{I}^m} f \, \mathrm{d}\bar\mu = \int_{\widehat{I}^m} f \, \mathrm{d}\Psi_1(\mu) = \sum_{u\in D_{f_1}} \int_{E_u} \bigl(f(\hat{\mathbf{t}}) - f(\phi_1(\hat{\mathbf{t}}))\bigr) \, \mathrm{d}\mu(\hat{\mathbf{t}}). \tag{3.11}$$

Furthermore, applying (3.3) to the function $f\mathbf{1}_{E_s}$, we also find, for any $s \in (0,1]$,

$$\int_{E_s} f \, \mathrm{d}\bar\mu = \int_{E_s} f \, \mathrm{d}\Psi_1(\mu) = \int_{E_s} \bigl(f(\hat{\mathbf{t}}) - f(\phi_1(\hat{\mathbf{t}}))\bigr) \, \mathrm{d}\mu(\hat{\mathbf{t}}). \tag{3.12}$$

This integral vanishes when $s \notin A$, because then $\bar\mu = 0$ on E_s, and also when $s \notin D_{f_1}$, because then $f(\hat{\mathbf{t}}) - f(\phi_1(\hat{\mathbf{t}})) = 0$ on E_s by (3.10). Consequently, (3.11)–(3.12) yield

$$\int_{\widehat{I}^m} f \, \mathrm{d}\bar\mu = \sum_{u\in D_{f_1}} \int_{E_u} f \, \mathrm{d}\bar\mu = \sum_{u\in A} \int_{E_u} f \, \mathrm{d}\bar\mu = \int_{\widehat{I}^m} f \, \mathrm{d}\nu, \tag{3.13}$$

which verifies (3.9) and thus $\bar\mu = \nu = \sum_{s\in A}\bar\mu_s$, which is supported on $\bigcup_{s\in A} E_s = \pi_1^{-1}(\widehat{A})$.

Finally, let $s \in (0,1]$. Again, let f be as in (3.8). Then, by (3.12) and (3.10), letting μ_{s-} denote the restriction of μ to $\{s-\}\times \widehat{I}^{m-1}$, regarded as a measure on $\widehat{I}^{m-1}$, and $\hat{\mathbf{t}}' := (\hat t_2,\dots,\hat t_m)$,

$$\begin{aligned}\int_{E_s} f\,\mathrm{d}\bar\mu &= \int_{\{s-\}\times\widehat{I}^{m-1}} (-\Delta f_1(s) g(\hat{\mathbf{t}}'))\,\mathrm{d}\mu\\ &= \bigl(f_1(s-)-f_1(s)\bigr)\int_{\widehat{I}^{m-1}} g(\hat{\mathbf{t}}')\,\mathrm{d}\mu_{s-}(\hat{\mathbf{t}}')\\ &= \int_{\widehat{I}^{m-1}} f(s-,\hat{\mathbf{t}}')\,\mathrm{d}\mu_{s-}(\hat{\mathbf{t}}') - \int_{\widehat{I}^{m-1}} f(s,\hat{\mathbf{t}}')\,\mathrm{d}\mu_{s-}(\hat{\mathbf{t}}'). \end{aligned}\tag{3.14}$$

Hence, $\Psi_1(\mu)_{s-} = \mu_{s-} = -\Psi_1(\mu)_s$. □

4. The dual space

The continuous linear functionals on $D([0,1])$ were described by Pestman [14], see also [11, §9.1]. We extend this result to several dimensions as follows.

Theorem 4.1. *Every continuous linear functional χ on $D([0,1]^m)$ has a unique representation*

$$\chi(f) = \sum_{J\subseteq[m]} \chi_J(f) \tag{4.1}$$

such that for every $J = \{j_1,\dots,j_\ell\}$, with $0 \leqslant \ell \leqslant m$, writing $J^{\mathsf{c}} := [m]\setminus J = \{j'_1,\dots,j'_{m-\ell}\}$,

$$\chi_J(f) = \sum_{t_{j_1},\dots,t_{j_\ell}\in(0,1]} \int_{I^{m-\ell}} \Delta_J f(t_1,\dots,t_m)\,\mathrm{d}\mu_{J;t_{j_1},\dots,t_{j_\ell}}(t_{j'_1},\dots,t_{j'_{m-\ell}}) \tag{4.2}$$

where each $\mu_{J;t_{j_1},\dots,t_{j_\ell}}$ is a signed Borel measure on $I^{m-\ell}$ and

$$\|\mu_J\| := \sum_{t_{j_1},\dots,t_{j_\ell}\in(0,1]} \|\mu_{J;t_{j_1},\dots,t_{j_\ell}}\| < \infty. \tag{4.3}$$

Furthermore,

$$2^{-m}\|\chi\| \leqslant \sum_{J\subseteq[m]} \|\mu_J\| \leqslant 3^m\|\chi\|. \tag{4.4}$$

Conversely, every such family of signed Borel measures $\mu_{J;t_{j_1},\dots,t_{j_\ell}}$ satisfying (4.3) defines a continuous linear functional on $D([0,1]^m)$ by (4.1)–(4.2).

Note that the sum in (4.2) formally is uncountable (when $\ell > 0$), but (4.3) implies that $\mu_{J;t_{j_1},\dots,t_{j_\ell}}$ is non-zero only for a countable set of $(t_{j_1},\dots,t_{j_\ell})$, so all sums are really countable.

Note also that for $J = \varnothing$, with $\ell = 0$, (4.2) reduces to

$$\chi_\varnothing(f) = \int_{I^m} f(t_1,\dots,t_m)\,\mathrm{d}\mu_\varnothing(t_1,\dots,t_m), \tag{4.5}$$

so this term in (4.1) is simply $\int f\,\mathrm{d}\mu_\varnothing$. For the other extreme, $\ell = m$, $\mu_{[m];t_1,\dots,t_m}$ is a signed measure on the one-point space I^0, i.e., a real number, and (4.2) is interpreted as

$$\chi_{[m]}(f) = \sum_{t_1,\dots,t_m\in(0,1]} \mu_{[m];t_1,\dots,t_m}\Delta_{[m]}f(t_1,\dots,t_m), \tag{4.6}$$

where $\|\mu_{[m]}\| := \sum_{t_1,\dots,t_m} |\mu_{[m];t_1,\dots,t_m}| < \infty$ and, again, the sums really are countable. (In other words, $\mu_{[m]} = (\mu_{[m];t_1,\dots,t_m})$ is an element of $\ell^1([0,1]^m)$.)

Proof. Since $D([0,1]^m) \cong C(\widehat{I}^m)$ (isometrically), we can use the Riesz representation theorem Proposition 2.2 and represent χ by a signed Baire measure μ on $\widehat{I}^m$. We use the projections in Section 3 and expand μ as

$$\mu = (\Phi_1 + \Psi_1)\cdots(\Phi_m + \Psi_m)\mu = \sum_{J\subseteq[m]} \tilde{\mu}_J, \tag{4.7}$$

where

$$\tilde{\mu}_J := \Bigl(\prod_{i\notin J}\Phi_i \prod_{j\in J}\Psi_j\Bigr)(\mu). \tag{4.8}$$

We define $\chi_J(f) := \int f\,\mathrm{d}\tilde{\mu}_J$; then (4.1) holds by (4.7), and we proceed to show the representation (4.2). If $j \in J$, then $\Psi_j(\tilde{\mu}_J) = \tilde{\mu}_J$, and thus by Lemma 3.2, $\tilde{\mu}_J$ is supported on the set $\pi_j^{-1}(\widehat{A}_j)$ for some countable sets $A_j \subset (0,1]$.

Suppose for notational convenience that $J = \{1,\dots,\ell\}$. Then, $\tilde{\mu}_J$ is thus supported on

$$\bigcap_{j=1}^{\ell} \pi_j^{-1}(\widehat{A}_j) = \widehat{A}_1 \times \cdots \times \widehat{A}_\ell \times \widehat{I}^{m-\ell}. \tag{4.9}$$

For $\hat{t}_1,\dots,\hat{t}_\ell \in \widehat{I}$, let $F_{\hat{t}_1,\dots,\hat{t}_\ell} := \{\hat{t}_1\}\times\cdots\times\{\hat{t}_\ell\}\times\widehat{I}^{m-\ell}$ and let $\tilde{\mu}_{J;\hat{t}_1,\dots,\hat{t}_\ell}$ be the restriction of $\tilde{\mu}_J$ to $F_{\hat{t}_1,\dots,\hat{t}_\ell}$.

Fix some $(t_1,\dots,t_\ell) \in (0,1]^\ell$, and let $\rho^* : F_{t_1,\dots,t_\ell} \to I^{m-\ell}$ be the map $\rho^*(t_1,\dots,t_\ell,\hat{t}_{\ell+1},\dots,\hat{t}_m) := (\rho(\hat{t}_{\ell+1}),\dots,\rho(\hat{t}_m))$. Let

$$\mu_{J;t_{j_1},\dots,t_{j_\ell}} := \rho^{*\sharp}(\tilde{\mu}_{J;t_{j_1},\dots,t_{j_\ell}}) \tag{4.10}$$

be the signed measure on $I^{m-\ell}$ induced by this map, noting that $\mu_{J;t_{j_1},\dots,t_{j_\ell}} = 0$ unless $(t_1,\dots,t_\ell) \in A_1\times\cdots\times A_\ell$.

If $i \in \{\ell+1,\dots,m\}$, then $\Phi_i(\tilde{\mu}_J) = \tilde{\mu}_J$ by (4.8), and thus (3.2) implies that for any bounded Baire measurable function f on $F_{t_1,\dots,t_\ell}$, recalling (3.1), that $\phi(\hat{t}) = \iota(\rho(\hat{t})) = \rho(\hat{t})$ for all $\hat{t} \in \widehat{I}$ by (2.2), and (4.10),

$$\begin{aligned}
\int_{F_{t_1,\dots,t_\ell}} f\,\mathrm{d}\tilde{\mu}_J &= \int_{\widehat{I}^m} \bigl(f\mathbf{1}_{F_{t_1,\dots,t_\ell}}\bigr)\circ\phi_{\ell+1}\circ\cdots\circ\phi_m\,\mathrm{d}\tilde{\mu}_J \\
&= \int_{F_{t_1,\dots,t_\ell}} f\bigl(t_1,\dots,t_\ell,\rho(\hat{t}_{\ell+1}),\dots,\rho(\hat{t}_m)\bigr)\,\mathrm{d}\tilde{\mu}_J \\
&= \int_{I^{m-\ell}} f(t_1,\dots,t_\ell,t_{\ell+1},\dots,t_m)\,\mathrm{d}\mu_{J;t_{j_1},\dots,t_{j_\ell}}(t_{\ell+1},\dots,t_m).
\end{aligned} \tag{4.11}$$

Consider now any sequence $(\hat{t}_i)_1^\ell$ with $\hat{t}_i \in \{t_i, t_i-\}$, $i = 1,\dots,\ell$, and let $q := |\{i : \hat{t}_i = t_i-\}|$. For $i \leqslant \ell$, (4.8) implies $\Psi_i(\tilde{\mu}_J) = \tilde{\mu}_J$, and thus by (3.6) and (4.11), for any bounded Baire measurable function f on $F_{\hat{t}_1,\dots,\hat{t}_\ell}$, writing $\mathbf{t}' = (t_{\ell+1},\dots,t_m)$,

$$\int_{F_{\hat{t}_1,\dots,\hat{t}_\ell}} f \,\mathrm{d}\tilde{\mu}_J = (-1)^q \int_{I^{m-\ell}} f(\hat{t}_1,\dots,\hat{t}_\ell,\mathbf{t}') \,\mathrm{d}\mu_{J;t_{j_1},\dots,t_{j_\ell}}(\mathbf{t}'). \tag{4.12}$$

Let $E_{t_1,\dots,t_\ell} := \widehat{\{t_1\}} \times \cdots \times \widehat{\{t_\ell\}} \times \hat{I}^{m-\ell}$. Summing (4.12) for the 2^ℓ choices of $(\hat{t}_1,\dots,\hat{t}_\ell)$, we obtain, recalling (2.6),

$$\int_{E_{t_1,\dots,t_\ell}} f \,\mathrm{d}\tilde{\mu}_J = \int_{I^{m-\ell}} \Delta_J f(t_1,\dots,t_\ell,\mathbf{t}') \,\mathrm{d}\mu_{J;t_{j_1},\dots,t_{j_\ell}}(\mathbf{t}'). \tag{4.13}$$

Furthermore, recalling that $\tilde{\mu}_J$ is supported on (4.9),

$$\chi_J(f) = \int_{\hat{I}^m} f \,\mathrm{d}\tilde{\mu}_J = \sum_{(t_1,\dots,t_\ell)\in A_1\times\cdots\times A_\ell} \int_{E_{t_1,\dots,t_\ell}} f \,\mathrm{d}\tilde{\mu}_J, \tag{4.14}$$

and (4.2) follows by summing (4.13) over all $(t_1,\dots,t_\ell) \in A_1 \times \cdots \times A_\ell$.

Next, the first inequality in (4.4) follows from (4.1)–(4.3), noting $\|\Delta_J f\| \leqslant 2^{|J|}\|f\| \leqslant 2^m\|f\|$. Furthermore, $\|\mu_J\| \leqslant \|\tilde{\mu}_J\|$ by (4.3) and (4.10), and $\|\tilde{\mu}_J\| \leqslant 2^{|J|}\|\mu\|$ by (4.8) and (3.4)–(3.5). Hence, the second inequality in (4.4) follows, noting that $\sum_J 2^{|J|} = 3^m$ by the binomial theorem.

The converse, that every family satisfying (4.3) defines a continuous linear functional on $D([0,1]^m)$ by (4.1)–(4.2) is obvious. Furthermore, it is easily seen that if χ is defined in this way, then with $\tilde{\mu}_J$ defined by (4.8), $\int f \,\mathrm{d}\mu_J$ equals the summand $\chi_J(f)$ given by (4.2), since the contribution from each $\chi_{J'}$ with $J' \neq J$ vanishes by cancellations, and thus the construction above recovers the measures $\mu_{J;t_{j_1},\dots,t_{j_\ell}}$ used to define χ. In other words, the measures $\mu_{J;t_{j_1},\dots,t_{j_\ell}}$ are uniquely determined by χ. $\square$

5. Measurability and random variables in $D([0,1]^m)$

We equip $D([0,1]^m)$ with the σ-field $\mathcal{D} = \mathcal{D}_m$ generated by all point evaluations $f \mapsto f(t)$, $t \in [0,1]^m$. We sometimes mention this σ-field explicitly for emphasis, but even when no σ-field is mentioned, $\mathcal{D}$ is implicitly assumed.

A $D([0,1]^m)$-valued random variable, or equivalently a random element of $D([0,1]^m)$, is thus a function $X : \Omega \to D([0,1]^m)$, defined on some probability space $(\Omega,\mathcal{F},P)$ such that for each fixed $t \in [0,1]^m$, $X(t)$ is measurable (i.e., a random variable).

Note that the norm $f \mapsto \|f\|$ is a $\mathcal{D}$-measurable function $D([0,1]^m) \to \mathbb{R}$, since it suffices to take the supremum in (1.1) over rational t. Hence, if X is a $D([0,1]^m)$-valued random variable, then $\|X\|$ is measurable, i.e., a random variable.

Remark 5.1. $\mathcal{D}$ is *not* equal to the Borel σ-field defined by the norm topology on $D([0,1]^m)$, see e.g. [11, Example 2.5]. The same example shows also that $\mathcal{D}$ is strictly weaker than Borel σ-field defined by the weak topology. (We omit the details.)

However, in the positive direction, Corollary 5.4 below shows that $\mathcal{D}$ coincides with the σ-field generated by the continuous linear functionals.

As a consequence (or directly), if we identify $D([0,1]^m)$ and $C(\widehat{I}^m)$ as usual, then $\mathcal{D}$ is also generated by all point evaluations $f \mapsto f(\hat{t})$, $\hat{t} \in \widehat{I}^m$.

Moreover, $\mathcal{D}$ coincides also with the Borel σ-field defined by the Skorohod topology on $D([0,1]^m)$, see [13].

Theorem 5.2. *Let $m \geqslant 1$ and $\ell \geqslant 1$. Every bounded multilinear form $\Upsilon : \bigl(D([0,1]^m)\bigr)^\ell \to \mathbb{R}$ is measurable.*

For $m = 1$, this is [11, Theorem 9.19]. Instead of trying to generalize the proof in [11], we proceed through a different route, using the known case $m = 1$ and Lemma 5.12 below (proved using several preliminary lemmas).

We first record the important special case $\ell = 1$; for $m = 1$ this was proved by Pestman [14].

Corollary 5.3. *Every continuous linear functional on $D([0,1]^m)$ is measurable.* □

Corollary 5.4. *The σ-field $\mathcal{D}$ on $D([0,1]^m)$ coincides with the σ-field $\mathcal{B}_w$ generated by the continuous linear functionals.*

Proof. Corollary 5.3 implies that $\mathcal{B}_w \subseteq \mathcal{D}$. The converse follows because every point evaluation is a continuous linear functional. □

We also rephrase this in terms of $D([0,1]^m)$-valued random variables. A function X from a measure space into a Banach space B is *weakly measurable* if $\langle \chi, X \rangle$ is measurable for every $\chi \in B^*$.

Corollary 5.5. *If $X : \Omega \to D([0,1]^m)$ is a function defined on some probability space $(\Omega, \mathcal{F}, P)$, then X is $\mathcal{D}$-measurable (i.e., a random variable in $D([0,1]^m)$) if and only if it is weakly measurable.* □

We begin the proof of Theorem 5.2 by a simple observation. (See [11, Lemma 9.12] for the case $m = 1$.)

Lemma 5.6. *The evaluation map $(f, t_1, \dots, t_m) \mapsto f(t_1, \dots, t_m)$ is measurable $D([0,1]^m) \times [0,1]^m \to \mathbb{R}$.*

Proof. By right-continuity,

$$f(t_1, \dots, t_m) = \lim_{n \to \infty} f\left(\frac{\lceil nt_1 \rceil}{n}, \dots, \frac{\lceil nt_m \rceil}{n}\right), \tag{5.1}$$

where the function on the right-hand side is measurable for each fixed n. □

In the next lemmas we fix $f \in D([0,1]^m)$ and consider differences along one coordinate only; for notational convenience we consider the first coordinate and write $t = (t_1, t')$ with $t_1 \in [0,1]$ and $t' \in [0,1]^{m-1}$. Furthermore, to avoid some trivial modifications at the endpoints 0 and 1, we extend f by defining $f(t_1, t') := f(0, t')$ for $t_1 < 0$ and $f(t_1, t') := f(1, t')$ for $t_1 > 1$.

We define, recalling (2.5).

$$\Delta^*_f(t_1) := \sup_{t' \in [0,1]^{m-1}} \bigl|\Delta_1 f(t_1, t')\bigr|, \qquad t_1 \in [0,1], \tag{5.2}$$

and, for an interval J,

$$\Lambda_f(J) := \sup\bigl\{|f(t_1, t') - f(u_1, t')| : t_1, u_1 \in J,\, t' \in [0,1]^{m-1}\bigr\}. \tag{5.3}$$

Lemma 5.7. *For every $t_1 \in [0,1]$, $f(s_1,t') \to f(t_1,t')$ as $s_1 \searrow t_1$ and $f(s_1,t') \to f(t_1-,t')$ as $s_1 \nearrow t_1$, uniformly for all $t' \in [0,1]^{m-1}$.*

In other words, $\Lambda_f([t_1,t_1+\delta]) \to 0$ and $\Lambda_f([t_1-\delta,t_1)) \to 0$ as $\delta \to 0$.

Proof. A standard compactness argument. Let $\varepsilon > 0$. By the definition of $D([0,1]^m)$, for every $t = (t_1,t') \in [0,1]^m$, there exists an open ball $B_t = B(t,\delta_t)$ centred at t such that if $s \in B_t$, then $f(s)$ differs by at most $\varepsilon/4$ from the limit as $s \to t$ in the corresponding octant. It follows that if $(s_1,t') \in B_t$, then $|f(s_1,t') - f(t_1-,t')| \leqslant \varepsilon/2$ if $s_1 < t_1$ and $|f(s_1,t') - f(t_1,t')| \leqslant \varepsilon/2$ if $s_1 \geqslant t_1$.

Fix t_1. By compactness, there exists a finite set $\{t'^1,\dots t'^N\}$ such that the corresponding balls $B_{(t_1,t'^j)}$ cover $\{t_1\} \times [0,1]^{m-1}$, and furthermore, there exists $\delta > 0$ such that for every $t' \in [0,1]^{m-1}$, the ball $B((t_1,t'),\delta)$ is contained in some $B_{(t_1,t'^j)}$. It follows that for any $t' \in [0,1]^{m-1}$, if $s \in (t-\delta,t)$, then $|f(s_1,t') - f(t_1-,t')| \leqslant \varepsilon/2$, and if $s_1 \in [t_1,t_1+\delta)$, then $|f(s_1,t') - f(t_1,t')| \leqslant \varepsilon/2$. □

Fix $\varepsilon > 0$ and let

$$\Xi_{f,\varepsilon} := \{t \in [0,1] : \Delta_f^*(t) \geqslant \varepsilon\}. \tag{5.4}$$

Lemma 5.8. *The set $\Xi_{f,\varepsilon}$ is finite for every $f \in D([0,1]^m)$ and $\varepsilon > 0$.*

This lemma is essentially the same as [13, Lemma 1.3].

Proof. Let the balls B_t be as in the proof of Lemma 5.7. It follows that if $s \in B_t$ and $s_1 \neq t_1$, then $|\Delta_1 f(s)| \leqslant \varepsilon/2$.

By compactness, there exists a finite set $\{t^1,\dots,t^N\}$ such that the corresponding balls B_{t^j} cover $[0,1]^m$. It follows that $\Xi_{f,\varepsilon}$ is a subset of the finite set $\{t_1^j : j \leqslant N\}$. □

Say that an interval $J \subseteq [0,1]$ is *fat* if $\Lambda_f(J) \geqslant \varepsilon$ and *bad* if J is fat and furthermore $J \cap \Xi_{f,\varepsilon} = \varnothing$.

Lemma 5.9. *For every $f \in D([0,1]^m)$ and $\varepsilon > 0$, there exists $\eta > 0$ such that if $J \subset [0,1]$ is an interval of length $|J| < \eta$, then J is not bad.*

Proof. We claim that for every $t \in [0,1]$, there exists an open interval $O_t \ni t$ such that no interval $J \subseteq O_t$ is bad.

In order to show this claim, suppose first that $t \notin \Xi_{f,\varepsilon}$. Then $\Delta_f^*(t) < \varepsilon$, and it follows by Lemma 5.7 that we can choose $\delta > 0$ such that $\Lambda_f\big((t-\delta,t+\delta)\big) < \varepsilon$. Hence, $O_t := (t-\delta,t+\delta)$ contains no fat interval, and thus no bad interval.

On the other hand, if $t \in \Xi_{f,\varepsilon}$, we similarly see by Lemma 5.7 that we can choose $\delta > 0$ such that $\Lambda_f\big((t-\delta,t)\big) < \varepsilon$ and $\Lambda_f\big((t,t+\delta)\big) < \varepsilon$. Let $O_t := (t-\delta,t+\delta)$. Any interval $J \subseteq O_t$ either contains $t \in \Xi_{f,\varepsilon}$, or it is a subset of $(t-\delta,t)$ or $(t,t+\delta)$ and then J is not fat; in both cases J is not bad.

This proves the claim. By a standard compactness argument (Lebesgue's covering lemma), there exists $\eta > 0$ such that every interval $J \subset [0,1]$ of length $|J| < \eta$ is contained in some O_t, and thus not bad. □

Lemma 5.10. *Fix $\varepsilon > 0$, let $M := |\Xi_{f,\varepsilon}| \geqslant 0$ and write $\Xi_{f,\varepsilon} = \{\xi_1, \dots, \xi_M\}$ with $0 < \xi_1 < \dots < \xi_M \leqslant 1$. Let further $\xi_i := 0$ for $i > M$. Then M and all ξ_i, $i \geqslant 1$, are measurable functionals of $f \in D([0,1]^m)$.*

Proof. For $n \geqslant 0$ and $1 \leqslant j \leqslant 2^n$, let $J_{n,j}$ be the dyadic interval $((j-1)2^{-n}, j2^{-n}]$.

By Lemma 5.9, if n is large enough (depending on f), then $J_{n,j}$ is not bad; hence, if $J_{n,j}$ is fat, then $J_{n,j} \cap \Xi_{f,\varepsilon} \neq \varnothing$.

Conversely, if $J_{n,j} \cap \Xi_{f,\varepsilon} \neq \varnothing$, let $t \in J_{n,j} \cap \Xi_{f,\varepsilon}$; then

$$\Lambda_f(J_{n,j}) \geqslant \Delta_f^*(t) \geqslant \varepsilon. \tag{5.5}$$

Hence, for large n, $J_{n,j}$ contains some $\xi_i \in \Xi_{f,\varepsilon}$ if and only if $J_{n,j}$ is fat.

Moreover, since $\Xi_{f,\varepsilon}$ is finite by Lemma 5.8, if n is large enough, then each $J_{n,j}$ contains at most one point ξ_i.

For $n \geqslant 0$, suppose that q_n of the intervals $J_{n,j}$, $1 \leqslant j \leqslant 2^n$, are fat, and let these be J_{n,j_i}, $i = 1, \dots, q_n$, with $j_1 < \dots < j_{q_n}$. Let further

$$\xi_{ni} := \begin{cases} j_{ni}/2^n, & i \leqslant q_n, \\ 0, & i > q_n. \end{cases} \tag{5.6}$$

We have shown above that for large n, $q_n = M$. Hence, $M = \lim_{n\to\infty} q_n$. Moreover, it follows from the argument above that for each fixed $i \geqslant 1$, $\xi_{in} \to \xi_i$ as $n \to \infty$.

Since each $\Lambda_f(J_{n,j})$ is a measurable functional of f (because it suffices to take the supremum in (5.3) over rational t, u, t'), it follows that each q_n and ξ_{ni} is measurable, and thus so are their limits M and ξ_i. $\square$

If F is a finite subset of $[0,1]$, arrange the elements of $F \cup \{0,1\}$ as $0 = x_0 < x_1 < \dots < x_N = 1$ and define

$$\Lambda_f^*(F) := \max_{1\leqslant i\leqslant N} \Lambda_f\bigl([x_{i-1}, x_i)\bigr) = \max_{1\leqslant i\leqslant N} \Lambda_f\bigl((x_{i-1}, x_i)\bigr), \tag{5.7}$$

where the last equality holds by the right-continuity of f.

Lemma 5.11. *For every $f \in D([0,1]^m)$, there exists a sequence $(\xi_j)_j^\infty$ in $[0,1]$ such that*

$$\Lambda_f^*\bigl(\{\xi_j\}_{j=1}^n\bigr) \to 0 \qquad \text{as } n \to \infty. \tag{5.8}$$

Moreover, these points can be chosen such that each ξ_j, $j \geqslant 1$, is a measurable functional of $f \in D([0,1]^m)$.

Proof. For each $k \geqslant 1$, let $\bar\xi_{ki}$, $i \geqslant 1$, be the numbers defined in Lemma 5.10 for $\varepsilon = 1/k$. Then each $\bar\xi_{ki}$ is a measurable functional of f. Consider all these functionals for $k \geqslant 1$ and $i \geqslant 1$, together with the constant functionals r for every rational $r \in [0,1]$, and arrange this countable collection of functionals in a sequence ξ_j, $j \geqslant 1$ (in an arbitrary but fixed way, not depending on f).

Now suppose that $f \in D([0,1]^m)$, and let $F_n := \{\xi_j\}_1^n$. Let $k \geqslant 1$, and let $\varepsilon = 1/k$. Then $M := |\Xi_{f,\varepsilon}| < \infty$ by Lemma 5.8, and thus there exists n_1 such that if $i \leqslant M$, then $\bar\xi_{ki} = \xi_j$ for some $j \leqslant n_1$. Hence, if $n \geqslant n_1$, then

$$F_n \supseteq \Xi_{f,\varepsilon}. \tag{5.9}$$

Furthermore, let η be as in Lemma 5.9, and let $L := \lfloor 1/\eta \rfloor + 1$. Since the rational numbers p/L, $0 \leqslant p \leqslant L$ all appear as some ξ_i, it follows that there exists n_2 such that if $n \geqslant n_2$ then $F_n \supseteq \{p/L\}_{p=0}^L$. Hence, if

$F_n \cup \{0,1\} = \{x_i\}_1^N$ as in (5.7), then each interval (x_{i-1}, x_i) has length $x_i - x_{i-1} \leqslant 1/L < \eta$. Consequently, by Lemma 5.9, the interval is not bad. Moreover, if also $n \geqslant n_1$, then (5.9) holds and thus $(x_{i-1}, x_i) \cap \Xi_{f,\varepsilon} = \varnothing$. Consequently, for $n \geqslant \max(n_1, n_2)$, no interval (x_{i-1}, x_i) is fat, and thus (5.7) yields $\Lambda_f^*(F_n) < \varepsilon = 1/k$.

Since k is arbitrary, this shows (5.8). □

As noted in Section 2.2, linear combinations of functions of the form $\bigotimes_1^m f_i = \prod_{i=1}^m f_i(x_i)$ with $f_i \in D([0,1])$ are dense in $D([0,1]^m)$. The next lemma shows that $f \in D([0,1]^m)$ can be approximated by such linear combinations in a measurable way.

Lemma 5.12. *For every $f \in D([0,1]^m)$, there exist functions $f_{N,k,i} \in D([0,1])$ for $N \geqslant 1$, $1 \leqslant k \leqslant N$ and $1 \leqslant i \leqslant m$ such that*

$$f_N := \sum_{k=1}^N \bigotimes_{i=1}^m f_{N,k,i} \to f \qquad \text{in } D([0,1]^m) \tag{5.10}$$

(i.e., uniformly), as $N \to \infty$. Furthermore, the functions $f_{N,k,i}$ can be chosen such that the mappings $f \mapsto f_{N,k,i}$ are measurable $D([0,1]^m) \to D([0,1])$.

Proof. We have so far considered the first coordinate. Of course, the results above hold for any coordinate. We let $\Lambda_f^i(J)$ and $\Lambda_f^{*i}(F)$ be defined as in (5.3) and (5.7), but using the i-th coordinate instead of the first. Thus Lemma 5.11 shows that for every $i \leqslant m$, there exists a sequence of measurable functionals ξ_j^i, $j \geqslant 1$, such that

$$\Lambda_f^{*i}\bigl(\{\xi_j^i\}_{j=1}^n\bigr) \to 0 \qquad \text{as } n \to \infty. \tag{5.11}$$

For $n \geqslant 0$ and $1 \leqslant i \leqslant m$, arrange $\{\xi_j^i\}_{j=1}^n \cup \{0,1\}$ in increasing order as $0 = x_0^i < \dots < x_{n_i}^i = 1$, where $n_i \leqslant n+1$. (Strict inequality is possible because there may be repetitions in $\{\xi_j^i\}_{j=1}^n \cup \{0,1\}$.) Let $J_j^i := [x_j^i, x_{j+1}^i)$ for $j < n_i$ and $J_{n_i}^i := \{1\}$. Thus $\{J_j^i\}_{j=0}^{n_i}$ is a partition of $[0,1]$. Let $h_j^i := \mathbf{1}_{J_j^i}$, the indicator function of J_j^i. (Note that J_j^i and h_j^i depend on n.)

Now define the step function g_n on $[0,1]^m$ by

$$g_n := \sum_{j_1,\dots,j_m} f(x_{j_1}^1, \dots, x_{j_m}^m) \bigotimes_{i=1}^m h_{j_i}^i, \tag{5.12}$$

i.e.,

$$g_n(t_1, \dots, t_m) := f(x_{j_1}^1, \dots, x_{j_m}^m) \qquad \text{when } t_i \in J_{j_i}^i \ (1 \leqslant i \leqslant m). \tag{5.13}$$

It follows from the definitions (5.3) and (5.7) that if $t_i \in J_{j_i}^i$ for every i, then

$$|g_n(t_1, \dots, t_m) - f(t_1, \dots, t_m)| \leqslant \sum_{i=1}^m \Lambda_f^i(J_{j_i}^i) \leqslant \sum_{i=1}^m \Lambda_f^{*i}(\{\xi_j^i\}_{j=1}^n). \tag{5.14}$$

Hence, (5.11) implies that

$$\|g_n - f\| = \sup_{t \in [0,1]^m} |g_n(t) - f(t)| \to 0 \qquad \text{as } n \to \infty, \tag{5.15}$$

i.e., $g_n \to f$ in $D([0,1]^m)$.

The rest is easy. We can write (5.12) as

$$g_n = \sum_{j_1,\dots,j_m} \bigotimes_{i=1}^{m} g_{n,j_1,\dots,j_m,i} \tag{5.16}$$

with

$$g_{n,j_1,\dots,j_m,i} := \begin{cases} f(x^1_{j_1},\dots,x^m_{j_m})h^1_{j_1}, & i=1,\\ h^i_{j_i}, & i>1. \end{cases} \tag{5.17}$$

The sum in (5.16) has $\prod_i(n_i+1) \leqslant (n+2)^m$ terms; by rearranging the terms in lexicographic order of $(j_1,\dots,j_m)$, we may write it as

$$g_n = \sum_{k=1}^{(n+2)^m} \bigotimes_{i=1}^{m} g_{n,k,i}, \tag{5.18}$$

where we, if necessary, have added terms that are 0 (with all $g_{n,k,i}=0$). Finally, we relabel again, defining for $(n+2)^m \leqslant N < (n+3)^m$

$$f_{N,k,i} := \begin{cases} g_{n,k,i}, & k \leqslant (n+2)^m,\\ 0, & k > (n+2)^m. \end{cases} \tag{5.19}$$

Then f_N defined by (5.10) satisfies $f_N = g_n$ for $(n+2)^m \leqslant N < (n+3)^m$, and thus $f_N \to f$ in $D([0,1]^m)$ as $N \to \infty$.

It is clear from the construction above that every n_i and x^i_j is a measurable functional of f; using Lemma 5.6 it follows that every $g_{n,j_1,\dots,j_m,i}$ defined by (5.17) depends measurably on f, and thus so does every $g_{n,k,i}$ and every $f_{N,k,i}$. □

Remark 5.13. The proof above yields functions $f_{N,k,i}$ of the special form $a\mathbf{1}_{[b,c)}$ or $a\mathbf{1}_{\{1\}}$, where a,b,c are measurable functionals of f. Cf. [17; 18].

Proof of Theorem 5.2. We use Lemma 5.12, with some fixed measurable choice of $f_{N,k,i}$. For every ℓ-tuple $(f^1,\dots,f^\ell)$, we apply Lemma 5.12 to each f^j and obtain, by continuity and multilinearity of Υ,

$$\begin{aligned} \Upsilon\bigl(f^1,\dots,f^\ell\bigr) &= \lim_{N\to\infty} \Upsilon\bigl(f^1_N,\dots,f^\ell_N\bigr) \\ &= \lim_{N\to\infty} \sum_{k_1,\dots,k_\ell=1}^{N} \Upsilon\Bigl(\bigotimes_{i=1}^m f^1_{N,k_1,i},\dots,\bigotimes_{i=1}^m f^\ell_{N,k_\ell,i}\Bigr). \end{aligned} \tag{5.20}$$

Define a bounded ℓm-linear form $\widetilde{\Upsilon}$ on $D([0,1])$ by

$$\widetilde{\Upsilon}\bigl(g_{11},\dots,g_{1m},\dots,g_{\ell 1},\dots,g_{\ell m}\bigr) := \Upsilon\Bigl(\bigotimes_{i=1}^m g_{1,i},\dots,\bigotimes_{i=1}^m g_{\ell,i}\Bigr). \tag{5.21}$$

Then the summand in (5.20) is $\widetilde{\Upsilon}\bigl((f^j_{N,k_j,i})_{1\leqslant j\leqslant \ell,\,1\leqslant i\leqslant m}\bigr)$. We apply the case $m=1$ of the theorem, which as said above is [11, Theorem 9.19], to $\widetilde{\Upsilon}$ (with ℓ replaced by ℓm); since each $f^j \mapsto f^j_{N,k_j,i}$ is measurable, this shows that each summand is a measurable function of $(f^1,\dots,f^\ell) \in (D([0,1]^m))^\ell$. Hence, so is their sum in (5.20), and thus by (5.20), also $\Upsilon(f^1,\dots,f^\ell)$. □

6. A Fubini theorem

Recall that a $D([0,1]^m)$-valued random variable X is a measurable function $X : (\Omega, \mathcal{F}, \mathbb{P}) \to D([0,1]^m)$ for some (usually unspecified) probability space $(\Omega, \mathcal{F}, \mathbb{P})$; hence X can be regarded as a function $X(\omega, t) : \Omega \times [0,1]^m \to \mathbb{R}$, and the measurability condition means that $X(\cdot, t)$ is measurable for each fixed $t \in [0,1]^m$. In fact, $X(\omega, t)$ is jointly measurable on $\Omega \times [0,1]^m$ as a consequence of Lemma 5.6.

Since functions in $D([0,1]^m)$ extend uniquely to $\widehat{I}^m$, yielding an natural identification $D([0,1]^m) \cong C(\widehat{I}^m)$, we can also regard the random variable X as a function $X : (\Omega, \mathcal{F}, \mathbb{P}) \to C(\widehat{I}^m)$ and thus also as a function $X(\omega, \hat{t}) : \Omega \times \widehat{I}^m \to \mathbb{R}$. This function is measurable in ω for every fixed $\hat{t} \in \widehat{I}^m$ by Corollary 5.3 (or more simply by considering a sequence $t_n \in [0,1]^m$ that converges to $\hat{t}$ in $\widehat{I}^m$); it is also a continuous function of $\hat{t}$ and thus Baire measurable for every fixed ω. In other words, the function $X(\omega, \hat{t})$ is separately measurable. However, $X(\omega, \hat{t})$ is in general *not* jointly mesurable on $\Omega \times \widehat{I}^m$. In fact, the example in [11, Remark 9.18] shows that Lemma 5.6 does not extend to the evaluation map $C(\widehat{I}^m) \times \widehat{I}^m$, and we may then choose $\Omega = C(\widehat{I}^m)$ with X the identity. (Here it does not matter whether we consider Baire or Borel measurability on $\widehat{I}^m$.)

This lack of joint measurability is a serious technical problem. A continuous linear functional on $C(\widehat{I}^m)$ is given by integration with respect to a Baire measure μ on $\widehat{I}^m$, see Proposition 2.2, and we would like to be able to use Fubini's theorem and interchange to order of integrations with respect to μ and the probability measure $\mathbb{P}$ on Ω, but the lack of joint measurability means that a straight-forward application of Fubini's theorem is not possible. However, the following theorem shows that the desired result nevertheless holds.

We say that a function f on a measure space $(S, \mathcal{S}, \mu)$ is *μ-measurable* if it is defined μ-a.e. and is μ-a.e. equal to an $\mathcal{S}$-measurable function (this is equivalent to f being measurable with respect to the μ-completion of $\mathcal{S}$). Furthermore, f is *μ-integrable* if it is μ-measurable and $\int |f| \, \mathrm{d}|\mu| < \infty$. (Recall that $|\mu|$ denotes the variation measure of μ.)

Theorem 6.1. *Suppose that X is a random variable in $D([0,1]^m) = C(\widehat{I}^m)$, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and that $\mu \in M_{\mathsf{Ba}}(\widehat{I}^m)$ is a signed Baire measure on $\widehat{I}^m$.*

(i) *If $\|X\| \leqslant C$ for some constant $C < \infty$, i.e., $|X(\omega, \hat{t})| \leqslant C$ for all $\omega \in \Omega$ and $t \in [0,1]^m$, then $\omega \mapsto \int_{\widehat{I}^m} X(\hat{t}) \, \mathrm{d}\mu(\hat{t})$ is a (bounded) measurable function on Ω, i.e., a random variable, and $\hat{t} \mapsto \mathbb{E}[X(\hat{t})]$ is an element of $C(\widehat{I}^m)$, and thus Baire measurable on $\widehat{I}^m$.*

(ii) *If $X \geqslant 0$ and μ is a positive Baire measure, then $\omega \mapsto \int_{\widehat{I}^m} X(\hat{t}) \, \mathrm{d}\mu(\hat{t})$ is a measurable function $\Omega \to [0, \infty]$, i.e., a (possibly infinite) random variable, and $\hat{t} \mapsto \mathbb{E}[X(\hat{t})]$ is a Baire measurable function $\widehat{I}^m \to [0, \infty]$.*

(iii) *If $\mathbb{E}\,|X(t)| \leqslant C$ for some constant $C < \infty$ and all $t \in [0,1]^m$, then $\int_{\widehat{I}^m} X(\hat{t}) \, \mathrm{d}\mu(\hat{t})$ exists a.s. and defines an integrable function on Ω, i.e.,* *an integrable random variable; and* $\hat t \mapsto \mathbb{E}[X(\hat t)]$ *is a bounded Baire measurable function on* $\widehat{I}^m$.

(iv) *If either* $\mathbb{E}\int_{\widehat{I}^m} |X(\hat t)|\,\mathrm{d}|\mu|(\hat t) < \infty$ *or* $\int_{\widehat{I}^m} \mathbb{E}[|X(\hat t)|]\,\mathrm{d}|\mu|(\hat t) < \infty$, *then* $\int_{\widehat{I}^m} X(\hat t)\,\mathrm{d}\mu(\hat t)$ *exists a.s. and defines an integrable function on* Ω, *i.e., an integrable random variable; similarly,* $\mathbb{E}[X(\hat t)]$ *exists for* μ*-a.e.* $\hat t$ *and defines a* μ*-integrable function on* $\widehat{I}^m$.

In all four cases,

$$\mathbb{E}\int_{\widehat{I}^m} X(\hat t)\,\mathrm{d}\mu(\hat t) = \int_{\widehat{I}^m} \mathbb{E}[X(\hat t)]\,\mathrm{d}\mu(\hat t). \tag{6.1}$$

We first prove a simple lemma, which is useful also in other situations.

Lemma 6.2. *Suppose that* X *is a random variable in* $D([0,1]^m)$. *Then, for every* $\hat t \in \widehat{I}^m$,

$$\mathbb{E}\,|X(\hat t)| \leqslant \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|. \tag{6.2}$$

Consequently,

$$\sup_{\hat t\in\widehat{I}^m} \mathbb{E}\,|X(\hat t)| = \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|. \tag{6.3}$$

Proof. If $\hat t \in \widehat{I}^m$, then there exists a sequence $t_n \in [0,1]^m$ such that $t_n \to \hat t$, and thus $X(t_n) \to X(\hat t)$. Hence, Fatou's lemma implies

$$\mathbb{E}\,|X(\hat t)| \leqslant \liminf_{n\to\infty} \mathbb{E}\,|X(t_n)| \leqslant \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|. \tag{6.4}$$

This shows (6.2), and (6.3) is an immediate consequence. ☐

Proof of Theorem 6.1. (i): First, $\int_{\widehat{I}^m} X(\hat t)\,\mathrm{d}\mu(\hat t)$ is measurable by Corollary 5.5, since $\chi : f \mapsto \int_{\widehat{I}^m} f\,\mathrm{d}\mu$ is a continuous linear functional on $D([0,1]^m)$.

Secondly, if $\hat t_n \to \hat t \in \widehat{I}^m$, then $X(\hat t_n) \to X(\hat t)$ by continuity, and thus $\mathbb{E}[X(\hat t_n)] \to \mathbb{E}[X(\hat t)]$ by dominated convergence. This shows that $\hat t \mapsto \mathbb{E}[X(\hat t)]$ is sequentially continuous, which is equivalent to continuity since $\widehat{I}$ is first countable (see Section 2.1). Alternatively, considering only $t \in [0,1]^m$, dominated convergence shows that $t \mapsto \mathbb{E}[X(t)]$ is a function in $D([0,1]^m)$, and that its continuous extension to $C(\widehat{I}^m)$ is given by $\mathbb{E}[X(\hat t)]$.

Finally, to show (6.1) we consider again the continuous linear functional $\chi : f \mapsto \int_{\widehat{I}^m} f\,\mathrm{d}\mu$ and use the decomposition in Theorem 4.1. Fix $J \subseteq [m]$ and suppose for notational convenience that $J = \{1,\dots,\ell\}$ for some $\ell \in \{0,\dots,m\}$. (The cases $\ell = 0$ and $\ell = m$ are somewhat special; we leave the simplifications in these cases to the reader.) Also fix $t_1,\dots,t_\ell \in (0,1]$ and consider the corresponding term in (4.2). Then $\Delta_J f(t_1,\dots,t_m)$ is a linear combination of the 2^ℓ terms $f(\hat t_1,\dots \hat t_\ell, t_{\ell+1},\dots,t_m)$ with $\hat t_j \in \{t_j, t_j-\} \subset \widehat{I}$ for $i = 1,\dots,\ell$.

Fix one such choice of $\hat t_1,\dots,\hat t_\ell$, and define for $f \in C(\widehat{I}^m)$ the function $\gamma(f)$ on $\widehat{I}^{m-\ell}$ by $\gamma(f)(\hat u_1,\dots,\hat u_{m-\ell}) := f(\hat t_1,\dots,\hat t_\ell,\hat u_1,\dots,\hat u_{m-\ell})$; in other words, $\gamma(f)$ is the restriction of f to the $(m-\ell)$-dimensional slice with the coordinates in J fixed to $\hat t' := (\hat t_1,\dots,\hat t_\ell)$, regarded as a function on $\widehat{I}^{m-\ell}$. Obviously, $\gamma(f) \in C(\widehat{I}^{m-\ell})$ for every $f \in C(\widehat{I}^m)$, so $\gamma : D([0,1]^m) = C(\widehat{I}^m) \to C(\widehat{I}^{m-\ell}) = D([0,1]^{m-\ell})$. Furthermore, for any fixed $u := (u_1, \dots, u_{m-\ell}) \in [0,1]^{m-\ell}$, the mapping $f \mapsto \gamma(f)(u) = f(\hat{t}', u)$ is $\mathcal{D}_m$-measurable on $D([0,1]^m)$, see Remark 5.1; hence, $\gamma : D([0,1]^m) \to D([0,1]^{m-\ell})$ is measurable for $\mathcal{D}_m$ and $\mathcal{D}_{m-\ell}$. Hence, $\gamma(X)$ is a $D([0,1]^{m-\ell})$-valued random variable, and applying Lemma 5.6 to $\gamma(X)$, we see that $(\omega, u) \mapsto \gamma(X)(\omega, u) = X(\omega, (\hat{t}', u))$ is jointly measurable on $\Omega \times [0,1]^{m-\ell}$. Consequently, we can use Fubini's theorem and conclude

$$\mathbb{E} \int_{I^{m-\ell}} X(\hat{t}', u) \, \mathrm{d}\mu_{J;t'}(u) = \int_{I^{m-\ell}} \mathbb{E}\, X(\hat{t}', u) \, \mathrm{d}\mu_{J;t'}(u). \tag{6.5}$$

Summing over all 2^ℓ choices of $\hat{t}'$ for a fixed $t' = (t_1, \dots, t_\ell)$, after multiplying with the correct sign, we obtain, letting $\mathbb{E}\, X$ denote the function $t \mapsto \mathbb{E}[X(t)]$ in $D([0,1]^m)$,

$$\mathbb{E} \int_{I^{m-\ell}} \Delta_J X(t', u) \, \mathrm{d}\mu_{J;t'}(u) = \int_{I^{m-\ell}} \Delta_J(\mathbb{E}\, X)(t', u) \, \mathrm{d}\mu_{J;t'}(u). \tag{6.6}$$

We now apply the decomposition (4.2) and sum (6.6) over all $t' \in (0,1]^\ell$; recall that the sum really is countable. This yields, using (4.2) twice and justifying the interchange of order of summation and expectation in the second equality below by dominated convergence, because

$$\left| \int_{I^{m-\ell}} \Delta_J X(t', u) \, \mathrm{d}\mu_{J;t'}(u) \right| \leqslant 2^J C \|\mu_{J;t'}\| \tag{6.7}$$

where the sum over t' of the right-hand sides is finite by (4.3),

$$\begin{aligned} \mathbb{E}[\chi_J(X)] &= \mathbb{E} \sum_{t' \in (0,1]^\ell} \int_{I^{m-\ell}} \Delta_J X(t', u) \, \mathrm{d}\mu_{J;t'}(u) \\ &= \sum_{t' \in (0,1]^\ell} \mathbb{E} \int_{I^{m-\ell}} \Delta_J X(t', u) \, \mathrm{d}\mu_{J;t'}(u) \\ &= \sum_{t' \in (0,1]^\ell} \int_{I^{m-\ell}} \Delta_J(\mathbb{E}\, X)(t', u) \, \mathrm{d}\mu_{J;t'}(u) \\ &= \chi_J(\mathbb{E}\, X), \end{aligned} \tag{6.8}$$

Finally, summing (6.8) over $J \subseteq [m]$ yields, by (4.1), $\mathbb{E}\, \chi(X) = \chi(\mathbb{E}\, X)$, which is (6.1).

(ii): Let $X_n(\omega, \hat{t}) := \min\bigl(X(\omega, \hat{t}), n\bigr)$, and note that X_n is a random variable in $C(\widehat{I}^m)$ for every $n \geqslant 1$. The result follows by applying (i) to each X_n and letting $n \to \infty$, using the monotone convergence theorem repeatedly.

(iii), (iv): Note first that the two alternative conditions in (iv) are equivalent by (6.1) applied to $|X|$ and $|\mu|$, which is valid by (ii). Furthermore, by Lemma 6.2, the assumption in (iii) implies

$$\mathbb{E}\, |X(\hat{t})| \leqslant C, \qquad \hat{t} \in \widehat{I}^m, \tag{6.9}$$

which in turn implies the assumption in (iv).

Decompose $X = X_+ - X_-$, where $X_+(\omega, \hat{t}) := \max\bigl(X(\omega, \hat{t}), 0\bigr)$ and $X_-(\omega, \hat{t}) := \max\bigl(-X(\omega, \hat{t}), 0\bigr)$, and note that X_+ and X_- are random variables in $C(\widehat{I}^m)$. Similarly, use the Hahn decomposition [4, Theorem 4.1.4] $\mu = \mu_+ - \mu_-$ as the difference of two positive Baire measures on $C(\widehat{I}^m)$. The result follows by applying (ii) to all combinations of $X_\pm$ and $\mu_\pm$, noting that this yields $\mathbb{E}\int_{\widehat{I}^m} X_\pm(\hat{t})\,\mathrm{d}\mu_\pm(\hat{t}) = \int_{\widehat{I}^m} \mathbb{E}[X_\pm(\hat{t})]\,\mathrm{d}\mu_\pm(\hat{t}) < \infty$, using the assumptions. For (iii), note also that (6.9) shows that $\mathbb{E}[X_\pm(\hat{t})]$ is a bounded (and thus finite) function on $\widehat{I}^m$. ☐

7. Separability

The Banach space $D([0,1]^m)$ is non-separable, which is a serious complication in various ways already for $m = 1$, see e.g. [1] and [11].

Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to B$ is a function defined on a probability space and taking values in a Banach space B. (In particular, X may be a B-valued random variable, for a given σ-field on B, but here we do not assume any measurability.) We then say, following [11, Definition 2.1], that

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

Note that in (ii), the exceptional null set may depend on x^*. (In fact, this is what makes the difference from (i): to assume $x^*(X) = 0$ outside some fixed null set for all x^* as in (ii) is equivalent to (i).)

Remark 7.1. A.s. separability is a powerful condition, which essentially reduces the study of X to the separable case. Unfortunately, it is too strong for our purposes. In the case $m = 1$, a random variable X taking values in $D([0,1])$ is a.s. separably valued if and only if there exists a fixed countable set N such that a.s. every discontinuity point of X belongs to N [11, Theorem 9.22]; we extend this to $D([0,1]^m)$ in Theorem 7.5 below. Hence, in applications to random variables in $D([0,1])$ or $D([0,1]^m)$, this condition is useful only for variables that have a fixed set of discontinuities, but not when there are discontinuities at random locations. We therefore mainly use the weaker property 'weakly a.s. separably valued' defined in (ii).

Example 7.2. 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. $X(t) = \mathbf{1}\{U \leqslant t\}$. Then, see [11, Example 2.5] for details, X is not a.s. separably valued, but X is weakly a.s. separably valued. (We can take $B_1 = C([0,1])$ in the definition (ii) above.)

We note the following simple properties.

Lemma 7.3. *Let B be a Banach space, and assume that $X_1, X_2, \dots$ are weakly a.s. separably valued functions $(\Omega, \mathcal{F}, \mathbb{P}) \to B$ for some probability space $(\Omega, \mathcal{F}, \mathbb{P})$.*

(i) *Any finite linear combination $\sum_{i=1}^N a_i X_i$ is weakly a.s. separably valued.*
(ii) *If $X_n \to X$ in B a.s., then X is weakly a.s. separably valued.*
(iii) *If $T : B \to \tilde{B}$ is a bounded linear map into another Banach space, then $T(X_1)$ is weakly a.s. separably valued in $\tilde{B}$.*

The same properties hold for a.s. separably valued functions too.

Proof. Let B_{1i} be a separable subspace of B satisfying the property in the definition for X_i, and let B_1 be the closed subspace generated by $\bigcup_i B_{1i}$.

Then B_1 is separable, and it is easily seen that this subspace verifies (i) and (ii). For (iii) we similarly use $\tilde{B}_1 := T(B_1)$. We omit the details. □

It was shown in [11, Theorems 9.24 and 9.25] that random variables in $D([0,1])$ always are weakly a.s. separably valued, and so are tensor powers of them in either the injective or projective tensor power. We extend this to $D([0,1]^m)$.

Theorem 7.4. (i) *Let X be a $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variable. Then X is weakly a.s. separably valued.*

(ii) *More generally, let $X_1,\dots,X_\ell$ be $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variables. Then, $\bigotimes_{i=1}^{\ell} X_i$ is weakly a.s. separably valued in the projective and injective tensor products $D([0,1]^m)^{\widehat{\otimes}\ell}$ and $D([0,1]^m)^{\check{\otimes}\ell}$.*

Proof. Consider first $m=1$. Then, as said above, (i) is [11, Theorem 9.24], while [11, Theorem 9.25] is the special case $X_1=\dots=X_\ell$ of (ii); moreover, it is easily checked that the proof of [11, Theorem 9.25] applies also to the case of general $X_1,\dots,X_\ell$. (The main difference in the proof is that we fix a countable set N and consider $X_1,\dots,X_\ell$ that a.s. are continuous at every fixed $t\notin N$.) Hence, the results hold for $m=1$.

In general, we apply Lemma 5.12 to $X_1,\dots,X_\ell$ and conclude that there are random variables $X^j_{N,k,i}$ in $D([0,1])$ such that, for every $j=1,\dots,\ell$,

$$X^j_N := \sum_{k=1}^{N} \bigotimes_{i=1}^{m} X^j_{N,k,i} \to X_j \qquad \text{in } D([0,1]^m) \tag{7.1}$$

as $N\to\infty$. Let $X_N := X^1_N\otimes\dots\otimes X^\ell_N \in D([0,1]^m)^{\widehat{\otimes}\ell}$. Then $X_N\to X := X^1\otimes\dots\otimes X^\ell$ in $D([0,1]^m)^{\widehat{\otimes}\ell}$ as $N\to\infty$. Furthermore, by (7.1),

$$X_N = \sum_{k_1,\dots,k_\ell\leqslant N} \bigotimes_{j=1}^{\ell}\bigotimes_{i=1}^{m} X^j_{N,k_j,i}. \tag{7.2}$$

By the case $m=1$ of the theorem (with ℓ replaced by ℓm), each term in this sum is a weakly a.s. separably valued random variable in $D([0,1])^{\widehat{\otimes}\ell m}$. Since the canonical inclusion $D([0,1])^{\widehat{\otimes}m}\to D([0,1])^{\check{\otimes}m}=D([0,1]^m)$ is continuous, and thus induces a continuous map $D([0,1])^{\widehat{\otimes}\ell m}=(D([0,1])^{\widehat{\otimes}m})^{\widehat{\otimes}\ell}\to D([0,1]^m)^{\widehat{\otimes}\ell}$, it follows by Lemma 7.3 that each X_N is weakly a.s. separably valued in $D([0,1]^m)^{\widehat{\otimes}\ell}$, and thus so is their limit X.

This proves the result for the projective tensor product. For the injective tensor product we use Lemma 7.3(iii) again, with the continuous inclusion $D([0,1]^m)^{\widehat{\otimes}\ell}\to D([0,1]^m)^{\check{\otimes}\ell}$. □

In contrast, and for completeness, we have the following characterization of a.s. separably valued random variables. (The case $m=1$ is [11, Theeorem 9.22].)

Theorem 7.5. *Let X be a $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variable. Then X is a.s. separably valued if and only if there exist (non-random) countable subsets $A_1,\dots,A_m$ of $[0,1]$ such that for every $i\leqslant m$, a.s. $\Delta_i X(t)=0$ for all $t=(t_1,\dots,t_m)$ with $t_i\notin A_i$.*

We consider first the deterministic case.

Lemma 7.6. *Let $f \in D([0,1]^m)$. Then there exist countable subsets $A_1, \dots, A_m$ of $[0,1]$ such that for every $i \leqslant m$, $\Delta_i f(t) = 0$ for all $t = (t_1, \dots, t_m)$ with $t_i \notin A_i$.*

Proof. Consider the first coordinate and let, recalling (5.2) and (5.4),

$$A_1 := \{t \in [0,1] : \Delta_f^*(t) > 0\} = \bigcup_k \Xi_{f,1/k}. \tag{7.3}$$

Then A_1 satisfied the claimed property by the definition (5.2), and A_1 is countable by Lemma 5.8. The same holds for $i > 1$ by relabelling the coordinates. □

Proof of Theorem 7.5. If X is a.s. separably valued, let D_1 be a separable subspace of $D([0,1]^m)$ such that $X \in D_1$ a.s. Let $\{f_n\}$ be a countable dense subset of D_1, and apply Lemma 7.6 to f_n for each n, yielding countable sets A_{in}. Define $A_i := \bigcup_n A_{in}$. Then, for every i, $\Delta_i f(t) = 0$ for all $t = (t_1, \dots, t_m)$ with $t_i \notin A_i$ and all $f \in D_1$; hence $\Delta_i X(t) = 0$ a.s. for all such t.

Conversely, suppose that such $A_1, \dots, A_m$ exist. Then, using the notation in Section 5 (with $f = X$), a.s. $\Delta_X^*(t_1) = 0$ for every $t \notin A_1$, and thus $\Xi_{X,\varepsilon} \subseteq A_1$. Hence, the construction in the proof of Lemma 5.11 yields $\xi_j \in \tilde{A}_1 := A_1 \cup (\mathbb{Q} \cap [0,1])$ a.s. for every j. Consequently, in the proof of Lemma 5.12, a.s. every $\xi_j^i \in \tilde{A}_i := A_i \cup (\mathbb{Q} \cap [0,1])$ and every $x_j^i \in \tilde{A}_i$. Let Q_i be the countable subset of $D([0,1])$ consisting of $\mathbf{1}_{[a,b)}$ with $a, b \in \tilde{A}_i$, together with $\mathbf{1}_{\{1\}}$. Then a.s. $h_j^i \in Q_i$, and thus if D_1 is the closed separable subspace of $D([0,1]^m)$ generated by the countable set $\bigotimes_{i=1}^m h_i$ with $h_i \in Q_i$, then a.s. $g_n \in Q$ for every n, and thus a.s. $X \in Q$. □

8. Moments

For a random variable X with values in some Banach space B, moments of X can be defined as $\mathbb{E}[X^{\otimes \ell}]$, see [11]. However, there are several possible interpretations of this; we may take the expectation in either the projective tensor power $B^{\hat{\otimes} \ell}$ or the injective tensor power $B^{\check{\otimes} \ell}$, and we can assume that the expectation exists in Dunford, Pettis or Bochner sense, thus giving six different cases. See [11] (and the short summary in Appendix B) for definitions and further details; we recall here only the implications for existence:

$$\text{projective} \implies \text{injective},$$
$$\text{Bochner} \implies \text{Pettis} \implies \text{Dunford},$$

and that if the moment exists in Bochner and Pettis sense, it is an element of the tensor product, but a Dunford moment is in general an element of the bidual of the tensor product.

In the special case $\ell = 1$, when we consider the mean $\mathbb{E}[X]$, there is no difference between the projective and injective case, but we can still consider the mean in Bochner, Pettis or Dunford sense.

8.1. **Bochner and Pettis moments as functions.** We consider here the different moments when $B = D([0,1]^m)$. Recall first from Section 2.2 that $D([0,1]^m)^{\check{\otimes}\ell} = D([0,1]^{\ell m})$. Hence, if the injective moment $\mathbb{E}[X^{\check{\otimes}\ell}]$ exists in Bochner or Pettis sense, then this moment is an element of $D([0,1]^{\ell m})$, and thus a function on $[0,1]^{\ell m}$.

Moreover, recall also that $D([0,1]^m)$ has the approximation property and thus the natural map $D([0,1]^m)^{\hat{\otimes}\ell} \to D([0,1]^m)^{\check{\otimes}\ell} = D([0,1]^{\ell m})$ is a continuous injection. Hence, if the projective moment $\mathbb{E}[X^{\hat{\otimes}\ell}]$ exists in Bochner or Pettis sense, then it too can be regarded as function in $D([0,1]^{\ell m})$, and it equals the corresponding injective moment. (Cf. [11, Theorem 3.3].)

It is easy to identify this function that is the moment (in any of these four senses for which the moment exists).

Theorem 8.1. *Let X be a $\mathcal{D}$-measurable random variable in $D([0,1]^m)$ and let $\ell \geqslant 1$. If X has a projective or injective moment $\mathbb{E}[X^{\otimes\ell}]$ in Bochner or Pettis sense, then this moment $\mathbb{E}[X^{\otimes\ell}]$ is the function in $D([0,1]^{\ell m})$ given by*

$$\mathbb{E}[X^{\otimes\ell}](t_1,\dots,t_\ell) = \mathbb{E}\left(\prod_{i=1}^{\ell} X(t_i)\right), \qquad t_i \in [0,1]^m. \tag{8.1}$$

In other words, the injective or projective Bochner or Pettis ℓ-th moment (when it exists) is the function describing all mixed ℓ-th moments of $X(t)$, $t \in [0,1]^m$.

Proof. As seen before the theorem, the moment can be regarded as a function in $D([0,1]^{\ell m})$. Since point evaluations are continuous linear functionals on $D([0,1]^{\ell m})$, it follows that

$$\mathbb{E}[X^{\otimes\ell}](t_1,\dots,t_\ell) = \mathbb{E}[X^{\otimes\ell}(t_1,\dots,t_\ell)] = \mathbb{E}\left(\prod_{i=1}^{\ell} X(t_i)\right), \tag{8.2}$$

showing (8.1). ◻

Remark 8.2. Theorem 8.1 does not hold for Dunford moments, since a Dunford integral in general is an element of the bidual, see Examples 8.11–8.12. However, we show a related result in Theorem 8.9, where we consider the moment function (8.1) for arguments in $\widehat{I}^m$ and not just in $[0,1]^m$.

8.2. **Existence of moments.** Since the Bochner and Pettis moments are given by (8.1) when they exist, the main problem is thus whether the different moments exist or not for a given random random $X \in D([0,1]^m)$. We give some conditions for existence, all generalizing results in [11] for the case $m = 1$.

For Bochner moments, we have a simple necessary and sufficient condition, valid for both projective and injective moments.

Theorem 8.3. *Let X be a $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variable. Then the following are equivalent.*

(i) *The projective moment $\mathbb{E}\, X^{\hat{\otimes}\ell}$ exists in Bochner sense.*
(ii) *The injective moment $\mathbb{E}\, X^{\check{\otimes}\ell}$ exists in Bochner sense.*

(iii) $\mathbb{E}\,\|X\|^\ell < \infty$ *and there exist (non-random) countable subsets* $A_1, \dots, A_m$ *of* $[0,1]$ *such that for every* $i \leqslant m$, *a.s.* $\Delta_i X(t) = 0$ *for all* $t = (t_1, \dots, t_m)$ *with* $t_i \notin A_i$.

Proof. By Theorem 7.5, (iii) is equivalent to $\mathbb{E}\,\|X\|^\ell < \infty$ and X a.s. separably valued. The equivalence now follow by [11, Theorem 3.8], since X is weakly measurable by Corollary 5.5. □

Unfortunately, the condition in Theorem 8.3(iii) shows that Bochner moments do not exist in many applications, cf. Remark 7.1. Hence the Pettis moments are more useful for applications; the following theorem gives a simple and widely applicable sufficient condition for their existence.

Theorem 8.4. *Let X be a $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variable, and suppose that $\mathbb{E}\,\|X\|^\ell < \infty$. Then the projective moment $\mathbb{E}\,X^{\hat{\otimes}\ell}$ and injective moment $\mathbb{E}\,X^{\check{\otimes}\ell}$ exist in Pettis sense.*

Proof. Recall that a bounded linear functional α on $D([0,1]^m)^{\hat{\otimes}\ell}$ is the same as a bounded ℓ-linear form $\alpha : D([0,1]^m)^\ell \to \mathbb{R}$. By Theorem 5.2, $\langle \alpha, X^{\otimes \ell}\rangle = \alpha(X, \dots, X)$ is measurable. Furthermore,

$$\left|\langle \alpha, X^{\otimes \ell}\rangle\right| \leqslant \|\alpha\| \|X\|^\ell, \tag{8.3}$$

and it follows that the family $\left\{\langle \alpha, X^{\hat{\otimes}\ell}\rangle : \|\alpha\| \leqslant 1\right\}$ is uniformly integrable. Moreover, $X^{\hat{\otimes}\ell}$ is weakly a.s. separably valued in $D([0,1]^m)^{\hat{\otimes}\ell}$ by Theorem 7.4. Hence a theorem by Huff [10], see also [11, Theorem 2.23 and Remark 2.24], shows that $\mathbb{E}\,X^{\hat{\otimes}\ell}$ exists in Pettis sense.

Since the natural inclusion $D([0,1]^m)^{\hat{\otimes}\ell} \to D([0,1]^m)^{\check{\otimes}\ell}$ is continuous, the injective moment $\mathbb{E}\,X^{\check{\otimes}\ell}$ too exists in Pettis sense. □

For injective moments, we can weaken the condition in Theorem 8.4, and obtain a necessary and sufficient condition; there is also a corresponding result for Dunford moments.

Theorem 8.5. *Suppose that X is a $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variable, and let $\ell \geqslant 1$.*

(i) $\mathbb{E}\,X^{\check{\otimes}\ell}$ *exists in Dunford sense* $\iff$ $\sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|^\ell < \infty$.

(ii) $\mathbb{E}\,X^{\check{\otimes}\ell}$ *exists in Pettis sense* $\iff$ *the family* $\{|X(t)|^\ell : t \in [0,1]^m\}$ *of random variables is uniformly integrable.*

We postpone the proof and show first two lemmas.

Lemma 8.6. *Suppose that X is a random element of $D([0,1]^m)$. Then*

$$\sup_{\chi\in D([0,1]^m)^*,\ \|\chi\|\leqslant 1} \mathbb{E}\,|\chi(X)| = \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|. \tag{8.4}$$

Proof. Denote the left and right sides of (8.4) by L and R, and note that trivially $R \leqslant L$ because each point evaluation $X \mapsto X(t)$ is a linear functional of norm 1.

For the converse, let $\chi \in D([0,1]^m)^*$ with $\|\chi\| \leqslant 1$. Note that the measurability of $\chi(X)$ follows from Corollary 5.3. By the Riesz representation theorem (Proposition 2.2), there exists a signed Baire measure μ on $\widehat{I}^m$ with

$\|\mu\| = \|\chi\| \leqslant 1$ such that $\chi(f) = \int_{\widehat{I}^m} f \,\mathrm{d}\mu$ for every $f \in D([0,1]^m)$. Consequently, applying Theorem 6.1 to $|X(\hat{t})|$ and $|\mu|$, noting that Lemma 6.2 yields $\mathbb{E}\,|X(\hat{t})| \leqslant R$ for every $\hat{t} \in \widehat{I}^m$,

$$\begin{aligned}\mathbb{E}\,|\chi(X)| = \mathbb{E}\left|\int_{\widehat{I}^m} X(\hat{t})\,\mathrm{d}\mu(\hat{t})\right| &\leqslant \mathbb{E}\int_{\widehat{I}^m} |X(\hat{t})|\,\mathrm{d}|\mu|(\hat{t}) \\ &= \int_{\widehat{I}^m} \mathbb{E}\,|X(\hat{t})|\,\mathrm{d}|\mu|(\hat{t}) \leqslant R\,|\mu|(\widehat{I}^m) = R\|\mu\| \leqslant R.\end{aligned}$$

Hence $L \leqslant R$, which completes the proof. ∎

We extend this to powers.

Lemma 8.7. *Suppose that X is a $D([0,1]^m)$-valued random variable, and $\ell \geqslant 1$.*

(i) *Then*

$$\sup_{\chi\in D([0,1]^m)^*,\ \|\chi\|\leqslant 1} \mathbb{E}\,|\chi(X)|^\ell = \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|^\ell. \tag{8.5}$$

In particular, the set $\bigl\{|\chi(X)|^\ell : \chi \in D([0,1]^m)^*,\ \|\chi\| \leqslant 1\bigr\}$ of random variables is a bounded subset of L^1 if and only if the set $\bigl\{|X(t)|^\ell : t \in [0,1]^m\bigr\}$ is.

(ii) *The set $\bigl\{|\chi(X)|^\ell : \chi \in D([0,1]^m)^*,\ \|\chi\| \leqslant 1\bigr\}$ is uniformly integrable if and only if the set $\bigl\{|X(t)|^\ell : t \in [0,1]^m\bigr\}$ is.*

Proof. (i): Let $\chi \in D([0,1]^m)^*$ with $\|\chi\| \leqslant 1$. Thus $\chi : D([0,1]^m) \to \mathbb{R}$ is a linear map, and we can take its tensor power $\chi^{\otimes\ell} : D([0,1]^m)^{\check{\otimes}\ell} \to \mathbb{R}^{\check{\otimes}\ell} = \mathbb{R}$, which is defined by

$$\chi^{\otimes\ell}(f_1 \otimes \ldots \otimes f_\ell) = \prod_{i=1}^{\ell} \chi(f_i) \tag{8.6}$$

together with linearity and continuity; $\chi^{\otimes\ell}$ is a linear functional on $D([0,1]^m)^{\check{\otimes}\ell}$ with norm $\|\chi^{\otimes\ell}\| = \|\chi\|^\ell \leqslant 1$.

Recalling that $D([0,1]^m)^{\check{\otimes}\ell} = D([0,1]^{\ell m})$, we apply Lemma 8.6 to $X^{\otimes\ell} \in D([0,1]^{\ell m})$ and the linear functional $\chi^{\otimes\ell}$ and obtain

$$\mathbb{E}\,|\chi(X)|^\ell = \mathbb{E}\,|\chi^{\otimes\ell}(X^{\otimes\ell})| \leqslant \sup_{t\in[0,1]^{\ell m}} \mathbb{E}\,|X^{\otimes\ell}(t)|. \tag{8.7}$$

Furthermore, if $t \in [0,1]^{\ell m}$, write $t = (t_1, \ldots, t_\ell)$ with $t_i \in [0,1]^m$; then by Hölder's inequality

$$\mathbb{E}\,|X^{\otimes\ell}(t)| = \mathbb{E}\,|X(t_1)\cdots X(t_\ell)| \leqslant \prod_{i=1}^{\ell} \bigl(\mathbb{E}\,|X(t_i)|^\ell\bigr)^{1/\ell} \leqslant \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|^\ell. \tag{8.8}$$

Combining (8.7) and (8.8), we see that the left-hand side of (8.5) is at most equal to the right-hand side. The converse follows again because each point evaluation $X \mapsto X(t)$ is a linear functional of norm 1.

(ii): Let $E \in \mathcal{F}$ be an arbitrary event in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and apply (8.5) to the random function $\mathbf{1}_E X$. This yields

$$\sup_{\chi\in D([0,1]^m)^*,\ \|\chi\|\leqslant 1} \mathbb{E}\bigl(\mathbf{1}_E|\chi(X)|^\ell\bigr) = \sup_{t\in[0,1]^m} \mathbb{E}\bigl(\mathbf{1}_E|X(t)|^\ell\bigr). \tag{8.9}$$

The result follows, since a collection $\{\xi_\alpha\}$ of random variables is uniformly integrable if and only if it is bounded in L^1 and $\sup_{\mathbb{P}(E)<\delta}\sup_\alpha \mathbb{E}\,|\mathbf{1}_E\xi_\alpha| \to 0$ as $\delta \to 0$ [8, Theorem 5.4.1]. $\square$

Proof of Theorem 8.5. (i): X is weakly measurable by Corollary 5.5, and weakly a.s. separably valued by Theorem 7.4. Hence [11, Theorem 3.11] shows that the injective Dunford moment exists if and only if $\big\{|\chi(X)|^\ell : \chi \in D([0,1]^m)^*,\ \|\chi\| \leqslant 1\big\}$ is a bounded subset of L^1. The proof is completed by Lemma 8.7(i).

(ii): Similar, using [11, Theorem 3.20] and Lemma 8.7(ii). $\square$

For $\ell = 1$, there is as said above no difference between projective and injective moments. For $\ell = 2$, the projective and injective moments are expectations taken in different spaces; nevertheless, the projective moments exist if and only if the injective moments do. For Bochner moments, this was shown in Theorem 8.3 (for any ℓ); for Pettis and Dunford moments this is shown in the next theorem. This theorem does not hold for $\ell \geqslant 3$; see the counterexample in [11, Example 7.27] (which is defined in $C(K)$ for another compact space K, but can be embedded in $C([0,1]) \subset D([0,1])$).

Theorem 8.8. *Let X be a $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variable.*

(i) $\mathbb{E}\,X^{\widehat{\otimes}2}$ *exists in Dunford sense* $\iff$ $\mathbb{E}\,X^{\check{\otimes}2}$ *exists in Dunford sense* $\iff$ $\sup_{t\in[0,1]^m}\mathbb{E}\,|X(t)|^2 < \infty$.

(ii) $\mathbb{E}\,X^{\widehat{\otimes}2}$ *exists in Pettis sense* $\iff$ $\mathbb{E}\,X^{\check{\otimes}2}$ *exists in Pettis sense* $\iff$ *the family* $\{|X(t)|^2 : t \in [0,1]^m\}$ *of random variables is uniformly integrable.*

Proof. The second equivalences in (i) and (ii) are the case $\ell = 2$ of Theorem 8.5. Furthermore, the existence of a projective moment always implies the existence of the corresponding injective moment. Hence it suffices to show that in both parts, the final condition implies the existence of the projective moment.

(i): Let α be a bounded bilinear form on $D([0,1]^m) = C(\widehat{I}^m)$. By Grothendieck's theorem [7], α extends to a bounded bilinear form on $L^2(\widehat{I}^m, \nu)$ for some Baire probability measure ν on $\widehat{I}^m$; furthermore, see e.g. [11, Theorem 7.20],

$$|\alpha(f,f)| \leqslant k'_G\|\alpha\| \int_{\widehat{I}^m} |f(t)|^2 \,\mathrm{d}\nu(t), \qquad f \in C(\widehat{I}^m), \tag{8.10}$$

where k'_G is a universal constant. (In this version, k'_G is at most 2 times the usual Grothendieck's constant, see [11, Remark 7.21].)

Furthermore, applying Lemma 8.6 to the random function $|X(t)|^2 \in D([0,1]^m)$ yields

$$\mathbb{E}\int_{\widehat{I}^m} |X(t)|^2 \,\mathrm{d}\nu \leqslant \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|^2. \tag{8.11}$$

Consequently, by combining (8.10) and (8.11),

$$\mathbb{E}\,|\alpha(X,X)| \leqslant k'_G\|\alpha\|\,\mathbb{E}\int_{\widehat{I}^m} |X(t)|^2 \,\mathrm{d}\nu \leqslant k'_G\|\alpha\| \sup_{t\in[0,1]^m} \mathbb{E}\,|X(t)|^2. \tag{8.12}$$

If $\sup_{t\in[0,1]^m} \mathbb{E}|X(t)|^2 < \infty$, then (8.12) shows that $\mathbb{E}|\alpha(X,X)| < \infty$ for every bounded bilinear form α on $D([0,1]^m)$, which implies the existence of the projective Dunford moment $\mathbb{E}\, X^{\hat{\otimes}2}$ by [11, Theorem 3.16].

(ii): Since the bounded linear functionals on $D([0,1]^m)^{\hat{\otimes}2}$ are identified with the bounded bilinear forms on $D([0,1]^m)$, (8.12) means, equivalently, that for every $\alpha \in (D([0,1]^m)^{\hat{\otimes}2})^*$,

$$\mathbb{E}|\alpha(X^{\hat{\otimes}2})| \leqslant k'_G\|\alpha\| \sup_{t\in[0,1]^m} \mathbb{E}|X(t)|^2. \tag{8.13}$$

Assume that the family $\{|X(t)|^2 : t \in [0,1]^m\}$ is uniformly integrable. By applying (8.13) to $\mathbf{1}_E X$ as in the proof of Lemma 8.7(ii), we obtain that the family $\big\{\alpha(X^{\hat{\otimes}2}) : \alpha \in (D([0,1]^m)^{\hat{\otimes}2})^*,\ \|\alpha\| \leqslant 1\big\}$ is uniformly integrable. Moreover, $X^{\hat{\otimes}2}$ is weakly a.s. separably valued by Theorem 7.4. Hence $\mathbb{E}\, X^{\hat{\otimes}2}$ exists in Pettis sense by Huff's theorem [10], cf. the proof of Theorem 8.4. ☐

8.3. **Dunford moments.** As said above, a Dunford moment in general is an element of the bidual of the space, and thus Theorem 8.1 does not hold for Dunford moments. Examples 8.11–8.12 below illustrate this. However, although even for $\ell = 1$, the bidual $D([0,1]^m)^{**}$ is large and unwieldy, it turns out that Dunford moments are always rather simple elements of it, and that they have a representation as functions generalising Theorem 8.1. This extends to injective moments of any order ℓ.

Theorem 8.9. *Let X be a $\mathcal{D}$-measurable random variable in $D([0,1]^m)$ and let $\ell \geqslant 1$. If X has an injective Dunford moment $\mathbb{E}[X^{\check{\otimes}\ell}]$ then this moment $\mathbb{E}[X^{\check{\otimes}\ell}]$ is represented by the bounded Baire measurable function ζ on $\widehat{I}^{\ell m}$ defined by*

$$\zeta(\hat{t}_1,\dots,\hat{t}_\ell) := \mathbb{E}\left(\prod_{i=1}^{\ell} X(\hat{t}_i)\right), \qquad \hat{t}_i \in \widehat{I}^m, \tag{8.14}$$

in the sense that if χ is any continuous linear functional on $D([0,1]^m)^{\check{\otimes}\ell} = D([0,1]^{\ell m}) = C(\widehat{I}^{\ell m})$ and χ is represented by a signed Baire measure μ on $\widehat{I}^{\ell m}$, then

$$\langle \mathbb{E}[X^{\check{\otimes}\ell}], \chi\rangle = \int_{\widehat{I}^{\ell m}} \zeta \,\mathrm{d}\mu. \tag{8.15}$$

In particular, if $\zeta \in C(\widehat{I}^{\ell m})$, then the Dunford moment $\mathbb{E}[X^{\check{\otimes}\ell}]$ is the element $\zeta \in C(\widehat{I}^{\ell m}) = D([0,1]^{\ell m})$.

Proof. Again, by considering the random variable $X^{\check{\otimes}\ell} \in D([0,1]^m)^{\check{\otimes}\ell} = D([0,1]^{\ell m})$, it suffices to consider the case $\ell = 1$.

In the case $\ell = 1$, the assumption says that $\mathbb{E}[X]$ exists as a Dunford moment; by Theorem 8.5, this implies that $\sup_{t\in[0,1]^m} \mathbb{E}|X(t)| < \infty$. It follows from Theorem 6.1(iii) that $\zeta(\hat{t}) := \mathbb{E}[X(\hat{t})]$ is a bounded Baire measurable function on $\widehat{I}^m$, and that for any continuous linear functional $\chi \in D([0,1]^m)^*$, represented by a signed Baire measure $\mu \in M_{\mathsf{Ba}}(\widehat{I}^m)$,

$$\mathbb{E}\langle \chi, X\rangle = \mathbb{E}\int_{\widehat{I}^m} X(\hat{t})\,\mathrm{d}\mu(\hat{t}) = \int_{\widehat{I}^m} \mathbb{E}[X(\hat{t})]\,\mathrm{d}\mu(\hat{t}) = \int_{\widehat{I}^m} \zeta(\hat{t})\,\mathrm{d}\mu(\hat{t}). \tag{8.16}$$

This shows, cf. Definition B.2, that the Dunford integral is given by $\langle \mathbb{E}[X], \chi\rangle = \int_{\widehat{I}^m} \zeta(\hat{t})\,\mathrm{d}\mu(\hat{t})$. ☐

Thus, similarly to the Bochner or Pettis moments in Theorem 8.1, an injective Dunford ℓ-th moment is represented by the function describing all mixed ℓ-th moments of $X(t)$, $t \in \widehat{I}^m$. However, unlike the situation for Bochner and Pettis moments, for Dunford moments we have, in general, to consider ζ as a function of $\widehat{I}^{\ell m}$ and not just on $[0,1]^{\ell m}$, see Example 8.12 below.

Remark 8.10. For projective Dunford moments $\mathbb{E}[X^{\widehat{\otimes}\ell}]$, the situation seems more complicated. We have a continuous inclusion $i : D([0,1]^m)^{\widehat{\otimes}\ell} \subseteq D([0,1]^m)^{\check{\otimes}\ell}$, which induces a continuous linear map between the biduals $i^{**} : (D([0,1]^m)^{\widehat{\otimes}\ell})^{**} \subseteq (D([0,1]^m)^{\check{\otimes}\ell})^{**}$. Thus, if a projective Dunford moment $\mathbb{E}[X^{\widehat{\otimes}\ell}]$ exists, then so does the injective Dunford moment $\mathbb{E}[X^{\check{\otimes}\ell}]$, and can by Theorem 8.9 be represented by the function ζ in (8.14). However, for $\ell \geqslant 2$, we do not know whether the map i^{**} is injective so that also the projective moment is represented by ζ.

We give two simple examples of Dunford moments, showing some bad behaviour that may occur. We take $m = 1$ and $\ell = 1$, i.e., we consider the mean $\mathbb{E}\,X$ of random variables X in $D([0,1])$.

Example 8.11. Let $X = 2^n \mathbf{1}_{[2^{-n-1},2^{-n})}$ with probability 2^{-n}, $n \geqslant 1$. Then $\mathbb{E}\,|X(t)| = \mathbb{E}\,X(t) = \mathbf{1}_{(0,1/2)}(t) \leqslant 1$, for every $t \in [0,1]$, and thus $\mathbb{E}\,X$ exists in Dunford sense by Theorem 8.5(i). However, the function $\zeta(t) := \mathbb{E}[X(t)]$ is not right-continuous at 0, so it does not belong to $D([0,1])$; hence this function does not represent $\mathbb{E}\,X$ in the sense of Theorem 8.1. In fact, it follows that $\mathbb{E}\,X \in D([0,1])^{**} \setminus D([0,1])$.

Nevertheless, Theorem 8.9 shows that $\mathbb{E}\,X$ is represented by ζ, regarded as a function on $\widehat{I}$. It is easily seen that $\zeta(\hat{t}) := \mathbb{E}[X(\hat{t})] = \mathbf{1}_{(0,1/2)}(\hat{t})$ for all $\hat{t} \in \widehat{I}$, and thus the Dunford mean $\mathbb{E}\,X$ is given by this function $\mathbf{1}_{(0,1/2)}$ on $\widehat{I}$; this function is bounded and Baire measurable (as guaranteed by Theorem 8.9), but it is not continuous, and thus does not correspond to an element of $D([0,1])$.

By Theorem 8.5, or Theorem 8.1, $\mathbb{E}\,X$ does not exist in Pettis (or Bochner) sense.

Example 8.12. Let $X = 2^n \mathbf{1}_{[1-2^{-n},1-2^{-n-1})}$ with probability 2^{-n}, $n \geqslant 1$. Then $\mathbb{E}\,|X(t)| = \mathbb{E}\,X(t) = \mathbf{1}_{[1/2,1)}(t) \leqslant 1$ for every $t \in [0,1]$, and thus $\mathbb{E}\,X$ exists in Dunford sense by Theorem 8.5(i). In this case, the function $\zeta(t) := \mathbb{E}[X(t)] = \mathbf{1}_{[1/2,1)}(t)$ is a function in $D([0,1])$. Nevertheless, the Dunford moment $\mathbb{E}\,X \in D([0,1])^{**}$ cannot be identified with the function $\zeta = \mathbf{1}_{[1/2,1)} \in D([0,1])$.

To see this, we consider $\hat{t} \in \widehat{I}$, as prescribed by Theorem 8.9. We have $X(1-) = 0$ a.s., and thus (8.14) yields $\zeta(1-) := \mathbb{E}[X(1-)] = 0$. Hence, if $t_n \nearrow 1-$, with $t_n \in (1/2,1)$, then $\zeta(t_n) = 1$ does not converge to $\zeta(1-) = 0$, and thus ζ is not continuous on $\widehat{I}$ at $1-$.

Consequently, we see that also in this example, $\mathbb{E} X \in D([0,1])^{**} \setminus D([0,1])$. Nevertheless, Theorem 8.9 shows that $\mathbb{E} X$ is represented by the function $\zeta(\hat{t}) = \mathbf{1}_{[1/2,1-)}(\hat{t})$ on $\widehat{I}$, and that (8.15) holds.

By Theorem 8.5, $\mathbb{E} X$ does not exist in Pettis (or Bochner) sense.

This example shows that it is necessary to consider the function ζ given by (8.14) as defined on $\widehat{I}^{\ell m}$, and not just on $[0,1]^{\ell m}$. In the present example, $\zeta \notin C(\widehat{I})$, but its restriction to $[0,1]$ is an element of $D([0,1])$, and thus the restriction of another function $\zeta' \in C(\widehat{I})$. Theorem 8.9 shows that the mean $\mathbb{E} X$ is represented by ζ, which is interpreted as an element of $C(\widehat{I})^{**} \setminus C(\widehat{I})$ by (8.15), and not by $\zeta' \in C(\widehat{I})$.

If a Pettis moment exists, then the corresponding Dunford moment exists and is equal to the Pettis moment. Theorems 8.1 and 8.9 then yield two versions of the same representation; obviously (8.1) is the restriction of (8.14) to $[0,1]^{\ell m}$; we state a simple result showing the consistency of the extensions to $\widehat{I}^{\ell m}$.

Theorem 8.13. *Let X be a $\mathcal{D}$-measurable random variable in $D([0,1]^m)$ and let $\ell \geqslant 1$. If X has an injective Pettis moment $\mathbb{E}[X^{\check{\otimes}\ell}]$, then the function ζ in* (8.14) *is continuous on $\widehat{I}^{\ell m}$, and this element of $C(\widehat{I}^{\ell m}) = D([0,1]^{\ell m}) = D([0,1]^m)^{\check{\otimes}\ell}$ equals the moment $\mathbb{E}[X^{\check{\otimes}\ell}]$.*

Proof. The Pettis moment $\mathbb{E}[X^{\check{\otimes}\ell}] \in D([0,1]^{\ell m}) = C(\widehat{I}^{\ell m})$, and this function on $\widehat{I}^{\ell m}$ equals ζ in (8.14) by the calculation (8.2) extended to $\hat{t}_1, \dots, \hat{t}_\ell \in \widehat{I}^m$, see also [11, Theorem 7.10]. ☐

There exists no general converse to this; even if the function ζ in (8.14) is continuous on $\widehat{I}^{\ell m}$, the Pettis moment $\mathbb{E}[X^{\check{\otimes}\ell}]$ does not have to exist, as shown by the trivial Example 8.14 below. However, Theorem 8.15 shows that the implication holds in some cases.

Example 8.14. Take again $\ell = m = 1$. Let $Y = \eta X$, where X is as in Example 8.11 and $\eta = \pm 1$, with $\mathbb{P}(\eta = 1) = \mathbb{P}(\eta = -1) = 1/2$, with X and η independent. Then $\zeta(\hat{t}) := \mathbb{E}[X(\hat{t})] = 0$ for every $\hat{t} \in \widehat{I}$, so $\zeta \in C(\widehat{I})$; nevertheless $\mathbb{E} Y$ does not exist in Pettis sense by Theorem 8.5 (or by the definition (B.3), taking $E := \{\eta = 1\}$).

Theorem 8.15. *Let X be a $\mathcal{D}$-measurable random variable in $D([0,1]^m)$ and let $\ell \geqslant 1$. Suppose further that either*

(a) *$X(t) \geqslant 0$ a.s. for every $t \in [0,1]$, or*
(b) *ℓ is even.*

Then the following are equivalent.

(i) *X has an injective Pettis moment $\mathbb{E}[X^{\check{\otimes}\ell}]$.*
(ii) *The function ζ in* (8.14) *exists everywhere and is continuous on $\widehat{I}^{\ell m}$.*
(iii) *$\mathbb{E}|X(\hat{t})|^\ell < \infty$ for every $\hat{t} \in \widehat{I}^m$, and the function $\hat{t} \mapsto \zeta(\hat{t}, \dots, \hat{t}) := \mathbb{E}[X(\hat{t})^\ell]$ is continuous on $\widehat{I}^m$.*

Proof. (i) $\Longrightarrow$ (ii): By Theorem 8.13.

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

(iii) $\Longrightarrow$ (i): In both cases we have $X(\hat{t})^\ell \geqslant 0$ and thus $|X(\hat{t})|^\ell = X(\hat{t})^\ell$ a.s. The argument in the proof of the similar [11, Theorem 7.19] shows that the family $\{|X(\hat{t})|^\ell : \hat{t} \in \widehat{I}^m\}$ of random variables is uniformly integrable. This proof in [11] is stated for $C(K)$ when K is a metrizable compact, but in the part of the proof used here, metrizability is used only to show that a sequentially continuous function on K is continuous, and this holds for every first countable compact K [6, Theorem 1.6.14 and Proposition 1.6.15], and thus for $\widehat{I}^m$.

The result (i) now follows from Theorem 8.5. $\square$

9. An application to Zolotarev distances

9.1. **Equal moments.** As a corollary of the results on moment above, we obtain the following results on equality of moments of two different $D([0,1]^m)$-valued random variables.

Theorem 9.1. *Let X and Y be $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variables, and let $\ell \geqslant 1$. Suppose that $\mathbb{E}\,\|X\|^\ell, \mathbb{E}\,\|Y\|^\ell < \infty$. Then the moments in* (iii) *and* (iv) *below exist in Pettis sense, and the following are equivalent.*

(i) *For every $t_1, \dots, t_\ell \in [0,1]^m$,*

$$\mathbb{E}\left(\prod_{i=1}^{\ell} X(t_i)\right) = \mathbb{E}\left(\prod_{i=1}^{\ell} Y(t_i)\right). \tag{9.1}$$

(ii) *For every continuous ℓ-linear form α on $D([0,1]^m)$,*

$$\mathbb{E}\,\alpha(X, \dots, X) = \mathbb{E}\,\alpha(Y, \dots, Y). \tag{9.2}$$

(iii) $\mathbb{E}[X^{\widehat{\otimes}\ell}] = \mathbb{E}[Y^{\widehat{\otimes}\ell}]$.

(iv) $\mathbb{E}[X^{\check{\otimes}\ell}] = \mathbb{E}[Y^{\check{\otimes}\ell}]$.

Note that (i) is a special case of (ii); the converse implication (i) $\Longrightarrow$ (ii) is far from trivial and is the main content of this theorem.

Proof. First, the Pettis moments in (iii) and (iv) exist by Theorem 8.4. The assumptions imply also that the expectations in (9.1) and (9.2) are finite, see (8.3).

(i) $\iff$ (iii) $\iff$ (iv): By Theorem 8.1.

(ii) $\iff$ (iii): The equality $\mathbb{E}[X^{\widehat{\otimes}\ell}] = \mathbb{E}[Y^{\widehat{\otimes}\ell}]$ holds if and only if we have $\langle \alpha, \mathbb{E}[X^{\widehat{\otimes}\ell}] \rangle = \langle \alpha, \mathbb{E}[Y^{\widehat{\otimes}\ell}] \rangle$ for every continuous linear functional α on $D([0,1]^m)^{\widehat{\otimes}\ell}$; these α can be identified with the continuous ℓ-linear forms on $D([0,1]^m)$, and the result follows since

$$\langle \alpha, \mathbb{E}\, X^{\widehat{\otimes}\ell} \rangle = \mathbb{E}\langle \alpha, X^{\widehat{\otimes}\ell} \rangle = \mathbb{E}\langle \alpha(X, \dots, X) \rangle \tag{9.3}$$

and similarly for Y. $\square$

9.2. **Zolotarev distance.** The *Zolotarev distance* $\zeta_s(X, Y)$ between two random variables X and Y with values in a Banach space, or more precisely between their distributions $\mathcal{L}(X)$ and $\mathcal{L}(Y)$, with $s > 0$ a real parameter, was defined by Zolotarev [19]; we refer to [19] or to [11, Appendix B] and the further references there for the definition and basic properties.

The Zolotarev distance is a useful tool to show approximation and convergence of distributions. In order to apply the Zolotarev distance to a problem, the first step is to show that the distance $\zeta_s(X,Y)$ between two given random variables is finite. It was shown in [11, Lemma B.2] that, assuming that X and Y are weakly measurable and that $\mathbb{E}\|X\|^s, \mathbb{E}\|Y\|^s < \infty$, this holds if and only if the projective (Dunford) moments $\mathbb{E}\,X^{\widehat{\otimes}\ell}$ and $\mathbb{E}\,Y^{\widehat{\otimes}\ell}$ are equal for every positive integer $\ell < s$. (This condition is vacuous if $0 < s \leqslant 1$.)

For the case of random variables in $D([0,1]^m)$, this and the results above yield the following simple criterion, which extends results for the case $m = 1$ in [11].

Theorem 9.2. *Let X and Y be $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variables, and let $s > 0$. Suppose that $\mathbb{E}\|X\|^s, \mathbb{E}\|Y\|^s < \infty$. Then the following are equivalent.*

(i) *The Zolotarev distance $\zeta_s(X,Y) < \infty$.*
(ii) *For every positive integer $\ell < s$ and every $t_1,\dots,t_\ell \in [0,1]^m$,*

$$\mathbb{E}\left(\prod_{i=1}^{\ell} X(t_i)\right) = \mathbb{E}\left(\prod_{i=1}^{\ell} Y(t_i)\right). \tag{9.4}$$

Proof. By Corollary 5.5, X and Y are weakly measurable, and thus [11, Lemma B.2] applies and shows, as said above, that (i) $\iff$ $\mathbb{E}[X^{\widehat{\otimes}\ell}] = \mathbb{E}[Y^{\widehat{\otimes}\ell}]$, which is equivalent to (ii) by Theorem 9.1. ∎

9.3. **Further results and comments.** For injective moments, we can weaken the moment assumption in Theorem 9.1.

Theorem 9.3. *Let X and Y be $\mathcal{D}$-measurable $D([0,1]^m)$-valued random variables, and let $\ell \geqslant 1$. Suppose further that $\sup_{t\in[0,1]^m}\mathbb{E}|X(t)|^\ell < \infty$ and $\sup_{t\in[0,1]^m}\mathbb{E}|Y(t)|^\ell < \infty$. Then the injective moments in* (ii) *below exist in Dunford sense, and the following are equivalent.*

(i) *For every $\hat{t}_1,\dots,\hat{t}_\ell \in \widehat{I}^m$,*

$$\mathbb{E}\left(\prod_{i=1}^{\ell} X(\hat{t}_i)\right) = \mathbb{E}\left(\prod_{i=1}^{\ell} Y(\hat{t}_i)\right). \tag{9.5}$$

(ii) $\mathbb{E}[X^{\check{\otimes}\ell}] = \mathbb{E}[Y^{\check{\otimes}\ell}]$.

Proof. The injective Dunford moments in (ii) exist by Theorem 8.5.
The equivalence (i) $\iff$ (ii) follows by Theorem 8.9. ∎

Note that we consider arbitrary $\hat{t}_i \in \widehat{I}$ in (9.5), unlike in (9.1); this is necessary as is shown by the following example.

Example 9.4. Take $\ell = m = 1$. Let X be as in Example 8.12, and let Y be the deterministic function $\mathbf{1}_{[1/2,1)} \in D([0,1])$. Then, as shown in Example 8.12, $\mathbb{E}[X(t)] = \mathbb{E}[Y(t)]$ for every $t \in [0,1]$, so (9.1) holds, but $\mathbb{E}[X(1-)] = 0 \neq 1 = \mathbb{E}[Y(1-)]$, and (9.5) fails; consequently, neither (i) nor (ii) in Theorem 9.3 holds.

For $\ell = 1$, there is no difference between injective and projective moments, and thus Theorem 9.3 applies to projective moments as well.

For $\ell = 2$, Theorem 8.8 shows that the assumptions of Theorem 9.3 imply also existence of the projective Dunford moments $\mathbb{E}\, X^{\hat{\otimes} 2}$ and $\mathbb{E}\, Y^{\hat{\otimes} 2}$. However, we do not know whether they always are equal when the injective moments are, see also Remark 8.10.

Problem 9.5. Assume that the assumptions of Theorem 9.3 hold with $\ell = 2$. Are (i) and (ii) equivalent also to $\mathbb{E}[X^{\hat{\otimes}\ell}] = \mathbb{E}[Y^{\hat{\otimes}\ell}]$?

Appendix A. Baire and Borel sets in $\widehat{I}^m$

We show here the claims in Example 2.4, and give some further results. The results are presumably known, but we have not found a reference and give proofs for completeness.

A.1. **The case $m = 1$.** Recall from Lemma 2.1(ii) that the Baire and Borel σ-fields coincide for every metrizable compact space; in particular $\mathsf{Ba}(I) = \mathcal{B}(I)$. The space $\widehat{I}$ is compact but not metrizable; nevertheless, as shown below, the Baire and Borel σ-field coincide there too.

Recall also that $\rho : \widehat{I} \to I$ is the natural projection.

Proposition A.1.

$$\mathsf{Ba}(\widehat{I}) = \mathcal{B}(\widehat{I}) = \bigl\{\rho^{-1}(A) \triangle N : A \in \mathcal{B}(I),\ N \subset \widehat{I} \textit{ with } |N| \leqslant \aleph_0\bigr\}. \quad \text{(A.1)}$$

In other words, the Borel (or Baire) sets in $\widehat{I}$ are obtained from the Borel sets in I in the natural way (by identifying t and $t-$), except that there may be a countable number of t such that the set contains only one of t and $t-$.

Proof. Let $\mathcal{G} := \bigl\{\rho^{-1}(A) \triangle N : A \in \mathcal{B}(I),\ N \subset \widehat{I} \text{ with } |N| \leqslant \aleph_0\bigr\}$; $\mathcal{G}$ is easily seen to be a σ-field. We prove three inclusions separately.

(i) $\mathsf{Ba}(\widehat{I}) \subseteq \mathcal{B}(\widehat{I})$. Trivial.

(ii) $\mathcal{B}(\widehat{I}) \subseteq \mathcal{G}$. Note first that an open interval in $\widehat{I}$ always is either of the form $\rho^{-1}\bigl((a,b)\bigr)$ for some open interval $(a,b) \subset (0,1)$, or of this form with one or two of the endpoints a and $b-$ added; if $b = 1$ we may also add 1. Let $\widehat{U} \subseteq \widehat{I}$ be open; then $\widehat{U}$ is a union of a (possibly uncountable) set of open intervals $\widehat{U}_\alpha \subseteq \widehat{I}$. For each $\widehat{U}_\alpha$, let $V_\alpha = (a_\alpha, b_\alpha)$ be the corresponding open interval in I; thus $\widehat{U}_\alpha \supseteq \rho^{-1}(V_\alpha)$ and $\widehat{U}_\alpha \setminus \rho^{-1}(V_\alpha)$ consists of at most the two endpoints and 1. Let $V := \bigcup_\alpha V_\alpha$; this is an open subset of $(0,1)$. Consequently, $V = \bigcup_j W_j$ for some countable collection of open disjoint intervals $W_j = (c_j, d_j) \subset (0,1)$.

Consider one of the intervals $V_\alpha = (a_\alpha, b_\alpha)$. If $a_\alpha \in V$, then $a_\alpha \in \rho^{-1}(V)$, and if $a_\alpha \notin V$, then a_α equals one of the endpoints c_j. Similarly, either $b_\alpha- \in \rho^{-1}(V)$ or b_α equals some endpoint d_j. Consequently, $\widehat{U} = \bigcup_\alpha \widehat{U}_\alpha \supseteq \rho^{-1}(V)$, and $\widehat{U} \setminus \rho^{-1}(V)$ is a subset of the countable set $\{c_j, d_j-, 1\}$. Consequently, $\widehat{U} \in \mathcal{G}$.

This shows that $\mathcal{G}$ contains every open subset of $\widehat{I}$, and thus the Borel σ-field $\mathcal{B}(\widehat{I})$.

(iii) $\mathcal{G} \subseteq \mathsf{Ba}(\widehat{I})$: By Lemma 3.1(ii), the mapping ρ is continuous and thus Baire measurable; hence $\rho^{-1}(A) \in \mathsf{Ba}(\widehat{I})$ for every $A \in \mathcal{B}(I) = \mathsf{Ba}(I)$.

If $\hat{t} \in \widehat{I}$, then the singleton $\{\hat{t}\}$ is closed, and a G_δ set; hence $\{\hat{t}\}$ is a Baire set. Consequently, every countable subset N of $\widehat{I}$ is a Baire set.

It follows that $\mathcal{G} \subseteq \mathsf{Ba}(\widehat{I})$. □

Corollary A.2. *Every closed subset of $\widehat{I}$ is a G_δ and every open subset of $\widehat{I}$ is an F_σ.*

Proof. The two parts are obviously equivalent. If F is a closed, and thus compact, subset, then F is a Borel set and thus by the proposition a Baire set. By [9, Theorem 51D], every compact Baire set is a G_δ. □

Since the Baire σ-field is generated by the compact G_δ sets, the corollary is equivalent to the proposition.

Corollary A.3. *Every finite Borel measure on $\widehat{I}$ is regular.*

Proof. Every finite Baire measure is regular [9, Theorem 52G]. □

A.2. **The case** $m \geqslant 2$**.** The equality of the Baire and Borel σ-fields in Proposition A.1 does not extend to $\widehat{I}^m$ for $m > 1$. We begin with the case $m = 2$.

Proposition A.4. $\mathcal{B}(\widehat{I}) \times \mathcal{B}(\widehat{I}) = \mathsf{Ba}(\widehat{I}) \times \mathsf{Ba}(\widehat{I}) = \mathsf{Ba}(\widehat{I}^2) \subsetneq \mathcal{B}(\widehat{I}^2)$.

Proof. The first equality follows by Proposition A.1, and the second by Lemma 2.1(iii). The final inclusion is trivial, and it remains to show that it is strict.

Let $h : I \to \widehat{I}^2$ be given by $h(t) = (t, 1-t)$. Thus, if we write $h = (h_1, h_2)$, then h_1 is the inclusion ι in Section 2.4, and $h_2(t) = \iota(1-t)$. By Lemma 3.1, ι is measurable $(I, \mathcal{B}) = (I, \mathsf{Ba}) \to (\widehat{I}, \mathsf{Ba})$, and thus both h_1 and h_2 are measurable $(I, \mathcal{B}) \to (\widehat{I}, \mathsf{Ba})$; Hence, if $E \subset \mathsf{Ba}(\widehat{I}^2) = \mathsf{Ba}(\widehat{I}) \times \mathsf{Ba}(\widehat{I})$, then $h^{-1}(E)$ is a Borel set in I.

On the other hand, for any $t \in [0,1]$, $[t, 1]$ and $[1-t, 1]$ are open intervals in $\widehat{I}$. Now let $A \subseteq I$ be arbitrary, and define a subset of $\widehat{I}^2$ by

$$E_A := \bigcup_{t \in A} \bigl([t, 1] \times [1-t, 1]\bigr). \tag{A.2}$$

This is an open subset of $\widehat{I}^2$, and thus a Borel set, i.e., $E_A \in \mathcal{B}(\widehat{I}^2)$.

However, $h^{-1}(E_A) = A$. Hence, if we take a set A that is not a Borel set, then $h^{-1}(E_A)$ is not a Borel set, and thus, by the first part of the proof, $E_A \notin \mathsf{Ba}(\widehat{I}^2)$. □

It follows easily from Proposition A.4 that $\mathsf{Ba}(\widehat{I}^m) \subsetneq \mathcal{B}(\widehat{I}^m)$ for every $m > 2$ too; we omit the details.

Note also that Proposition A.4 implies that there exists a closed set in $\widehat{I}^2$ that is not G_δ; cf. Corollary A.2.

Remark A.5. Proposition A.4 gives one proof that $\widehat{I}$ is not metrizable, since if it were, then $\widehat{I}^2$ would be as well, which would contradict Lemma 2.1(ii). (Another proof uses that the proof above shows that the set $\{(t, 1-t) : t \in I\}$ is an uncountable discrete subset of $\widehat{I}^2$; this is impossible for a compact metric space.)

Appendix B. Integration in Banach spaces

Let f be a function defined on a measure space $(S, \mathcal{S}, \mu)$ with values in a Banach space B. Then there are (at least) three different ways to define the integral $\int_S f \, d\mu$; the three definitions apply to different classes of functions f, but when two or all three definitions apply to a function f, then the integrals coincide. We use all three integrals in Section 8 in the case when $(S, \mathcal{S}, \mu)$ is a probability space and the integrals can be seen as expectations.

We give here a brief summary, and refer to [11] and the reference given there for further details.

B.1. **Bochner integral.**

The Bochner integral is a straight-forward generalization of the Lebesgue integral to Banach-space valued functions.

Theorem B.1. *A function f is Bochner integrable if and only if f is Borel measurable, a.s. separably valued, and $\int_S \|f\| \, d\mu < \infty$.* □

The Bochner integral $\int_S f \, d\mu$ then is an element of B.

Unfortunately, as discussed in Section 8, the condition of a.s. separably valued makes the Bochner integral unapplicable in many interesting examples of $D([0,1]^m)$-valued random variables.

B.2. **Dunford integral.**

The Dunford integral is the most general of our integrals.

Definition B.2. A function $f : S \to B$ is Dunford integrable if $x \mapsto \langle \chi, f(x) \rangle$ is integrable (and in particular measurable) on S for every continuous linear functional $\chi \in B^*$. In this case, as a consequence of the closed graph theorem, there exists a (unique) element $\int_S f \, d\mu \in B^{**}$ (the *Dunford integral*) such that

$$\int_S \langle \chi, f(x) \rangle \, d\mu = \langle \int_S f \, d\mu, \chi \rangle, \qquad \chi \in B^*. \tag{B.1}$$

Note that the Dunford element is defined as an element of the bidual B^{**}; in general, it is not an element of B. (See Example 8.11.)

B.3. **Pettis integral.**

A Pettis integral is a Dunford integral that has its value in B; furthermore, the following is required (in order to have useful properties). Note that if f is Dunford integrable over S, then f is always Dunford integrable over every measurable subset $E \subseteq S$.

Definition B.3. A function $f : S \to B$ is Pettis integrable if f is Dunford integrable with $\int_S f \, d\mu \in B$, and, moreover, $\int_E f \, d\mu \in B$ for every measurable subset $E \subseteq S$.

By definition, a Pettis integrable function is also Dunford integrable, and the two integrals coincide. Similarly, it is easy to see that a Bochner integrable function is Pettis integrable (and thus also Dunford integrable) and that the integrals coincide. The converses do not hold, in general.

References

[1] Patrick Billingsley. *Convergence of Probability Measures.* 2nd ed., Wiley, New York, 1999.

[2] V. I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[3] Nicolas Broutin and Henning Sulzbach. Partial match limit processes in higher dimensions. In preparation.

[4] Donald L. Cohn. *Measure Theory*, Birkhäuser, Boston, 1980.

[5] G. A. Edgar. Measurability in a Banach space. II. *Indiana Univ. Math. J.* **28** (1979), no. 4, 559–579.

[6] Ryszard Engelking. *General Topology.* 2nd ed., Heldermann Verlag, Berlin, 1989.

[7] Alexandre Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo* **8** (1953), 1–79.

[8] Allan Gut. *Probability: A Graduate Course*, 2nd ed., Springer, New York, 2013.

[9] Paul R. Halmos. *Measure Theory.* Van Nostrand, Princeton, NJ, 1950.

[10] R. Huff. Remarks on Pettis integrability. *Proc. Amer. Math. Soc.* **96** (1986), no. 3, 402–404.

[11] Svante Janson and Sten Kaijser. *Higher moments of Banach Space Valued Random variables. Memoirs Amer. Math. Soc.* **238** (2015), no. 1127.

[12] Ralph Neininger and Henning Sulzbach. On a functional contraction method. *Ann. Probab.* **43** (2015), no. 4, 1777–1822.

[13] Georg Neuhaus. On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42** (1971), 1285–1295.

[14] Wiebe R. Pestman. Measurability of linear operators in the Skorokhod topology. *Bull. Belg. Math. Soc. Simon Stevin* **2** (1995), no. 4, 381–388.

[15] Raymond A. Ryan. *Introduction to Tensor Products of Banach Spaces.* Springer-Verlag, London, 2002.

[16] Miron L. Straf. Weak convergence of stochastic processes with several parameters. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pp. 187–221. Univ. California Press, Berkeley, Calif., 1972.

[17] Michael J. Wichura. On the Weak Convergence of non-Borel Probabilities on a Metric Space. Ph.D. Thesis, Columbia University, 1968.

[18] Michael J. Wichura. Inequalities with applications to the weak convergence of random processes with multi-dimensional time parameters. *Ann. Math. Statist.* **40** (1969), 681–687.

[19] 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).

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

Email address: `svante.janson@math.uu.se`

URL: `http://www.math.uu.se/svante-janson`