Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

The space of bounded variation with infinite-dimensional codomain

Martin Heida , Robert I. A. Patterson, D. R. Michiel Renger

submitted: December 15, 2016

Weierstraß-Institut
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: martin.heida@wias-berlin.de
robert.patterson@wias-berlin.de
d.r.michiel.renger@wias-berlin.de

No. 2353
Berlin 2016

2010 *Mathematics Subject Classification.* 26A45, 26A24, 28B99, 46G05, 46G10.

Key words and phrases. Bounded variation, infinite-dimensional codomain, metric spaces, non-metric topologies, Banach spaces, vector measures, Aubin-Lions, compactness.

M.H. was supported by DFG through subproject C05 in SFB 1114 *Scaling Cascades in Complex Systems*.
D.R.M.R. was supported by DFG through subproject C08 in SFB 1114.
The authors thank Richard Kraaij and Emanuele Spadaro for valuable comments and discussions.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

We study functions of bounded variation with values in a Banach or in a metric space. We provide several equivalent notions of variations and provide the notion of a time derivative in this abstract setting. We study four distinct topologies on the space of bounded variations and provide some insight into the structure of these topologies. In particular, we study the meaning of convergence, duality and regularity for these topologies and provide some useful compactness criteria, also related to the classical Aubin-Lions theorem. We finally provide some useful applications to stochastic processes.

1 Introduction

Functions of bounded variation have a broad range of applications, including materials science, chemistry, image processing, or more generally, models that involve jumps and intervals of differentiability or even quiescence. There are even some applications to random processes, e.g. [24, 18], which was in fact one of the motivations behind the current work, see [28]. Many properties of functions of bounded variation and their corresponding topologies can be found in the standard works [17] and [4]. Considerable research has been carried out on functions of bounded variations on an infinite-dimensional domain, see e.g. [2, 12, 6, 3]. In this paper we study functions of bounded variation $\mathrm{BV}(0,T;Z)$ mapping an interval $(0,T)$ to an infinite-dimensional codomain, see for example [1, 10, 21]. In the general setting we take Z to be a metric space; of particular interest is the case where $Z = X^*$, the Banach dual of a Banach space X.

The aim is to collect a number of results which to the best of our knowledge are not yet included in the literature. First, we introduce the usual concepts as variations and time derivatives in Banach and metric spaces, and study some of their basic properties. Many of these are generalisations of the finite-dimensional theory outlined in [4]. The main ideas behind the generalisation to metric spaces is certainly not new, see for example [1], but an overview of the basic properties seems to be lacking in the literature. Secondly, we introduce a number of different topologies on the space of functions of bounded variation, and study their properties. These topological results are often subtly different from the results in the finite-dimensional setting. In particular, we find a new topology, which we call *hybrid*. Among the topological properties that we study, the characterisation theorems for compactness are of particular relevance. From the compactness result we will also derive a generalised Aubin-Lions result. Lastly, we will see that some of the topological properties imply generalised Prokhorov and Portemanteau Theorems, which are important for probabilistic applications.

In the remainder of this introduction we introduce functions of bounded variation and many of the related concepts in the setting of infinite dimensional Banach space codomains before generalising to metric spaces, which we do not require to be locally compact. We then define the different topologies on the space of functions of bounded variation, discuss applications to stochastic processes and end with a brief overview of the paper.

1.1 Functions of bounded variation with Banach codomains

To explain the main concepts we first assume that $Z = X^*$ is the Banach dual of some Banach space X; in Section 1.2 we explain how these concepts can be generalised to metric spaces. Classically, the space of functions of bounded variation is defined via the *pointwise variation* (see for example [9, 15]), that is for $f : (0,T) \to X^*$,

$$\mathrm{pvar}(f) := \sup_{0<t_0<t_1<\cdots<t_n<T} \sum_{i=1}^{n} \left\| f(t_{i-1}) - f(t_i) \right\|_{X^*}, \tag{1.1}$$

where the supremum runs over all finite partitions of the interval $(0,T)$. The pointwise variation is often used to define functions of bounded variation in fields where one is interested in the values at every single timepoint, like the field of energetic solutions for rate-independent systems and in nonsmooth mechanics, see [25] and Chapter 1 of [27], which is closely connected to our theory for the Banach-valued case. However, pointwise defined functions of bounded pointwise variation have some mathematical drawbacks, for example, one can not define weak derivatives. Therefore, functions of bounded variation are usually defined as a subset of the L^1 functions, which are strictly speaking equivalence classes of functions differing on sets of measure 0. This means that (1.1) must be extended and one introduces the *essential pointwise variation*, i.e. for $f \in L^1\big(0,T;X^*\big)$,

$$\operatorname{epvar}(f) := \inf_{g\sim f} \operatorname{pvar}(g), \tag{1.2}$$

where $g \sim f :\iff g(t) = f(t)$ for almost every $t \in (0,T)$. One can now define the Banach space of BV functions:

Definition 1.1.

$$\mathrm{BV}(0,T;X^*) := \Big\{ f \in L^1(0,T;X^*) : \|f\|_{\mathrm{BV}} := \|f\|_{L^1} + \operatorname{epvar}(f) < \infty \Big\}.$$

(Essential) Pointwise variation works very well for functions whose domain is a real interval. An equivalent notion of *variation*, which generalises better to functions with multi-dimensional domains, involves integrating against test functions $\phi \in C_0(0,T;X)$. To this end we introduce the notation

$$\langle\!\langle \phi, f \rangle\!\rangle := \int_0^T \langle \phi(t), f(t) \rangle \,\mathrm{d}t \tag{1.3}$$

where $\langle\cdot,\cdot\rangle$ is the duality pairing between the Banach space X and its Banach dual X^* (see Section A.2). It is now possible to introduce two integral-based variations for $f \in L^1(0,T;X^*)$:

$$\operatorname{var}(f) := \sup_{\substack{\phi\in C^1_0(0,T;X):\\ \|\phi\|_\infty\le 1}} -\langle\!\langle \dot\phi, f \rangle\!\rangle, \tag{1.4}$$

and

$$\operatorname{varw}(f) := \sup_{\substack{\phi\in C^1_b(0,T;X):\\ \|\phi\|_\infty\le 1}} \langle \phi(T), f(T-) \rangle - \langle \phi(0), f(0+) \rangle - \langle\!\langle \dot\phi, f \rangle\!\rangle, \tag{1.5}$$

provided the right limit $f(0+)$ at 0 and the left limit $f(T-)$ at T exist; otherwise we set the value to $+\infty$. In many settings all these notions of variation are known to be equivalent. As one of the main results of this paper, we will prove in Corollary 2.8 that this equivalence also holds in the infinite-dimensional setting, i.e. for all $f \in \mathrm{BV}(0,T;X^*)$:

$$\operatorname{epvar}(f) = \operatorname{var}(f) = \operatorname{varw}(f).$$

Note that the concept of epvar generalizes directly to the metric case, while for var and varw this is more involved; for this reason, we chose epvar in Definition 1.1.

On the other hand, the notion of $\operatorname{var}(\cdot)$ is very useful for example to introduce the notion of a derivative, in the following sense (see also Theorem 2.13). For an f with $\operatorname{var}(f) < \infty$, the mapping $\phi \mapsto -\langle\!\langle \dot\phi, f \rangle\!\rangle$ is a uniformly continuous linear functional with operator norm $\operatorname{var}(f)$ on $C^1_0(0,T;X)$ viewed as a subspace of $C_0(0,T;X)$. Therefore it has a unique, operator-norm preserving extension to an element of $C_0(0,T;X)^*$. In Theorem A.7, the space $C_0(0,T;X)^*$ will be seen to be isometrically isomorphic to $\operatorname{rca}(0,T;X^*)$, the space of regular, X^*-valued measures with the total variation norm $\|\cdot\|_{\mathrm{TV}}$, defined by (A.4). Therefore the extended operator is a weak derivative $\dot f$ in the sense that

$$-\langle\!\langle \dot\phi, f \rangle\!\rangle = \langle\!\langle \phi, \dot f \rangle\!\rangle \qquad \text{for all } \phi \in C^1_0(0,T;X) \tag{1.6}$$

and $\operatorname{var}(f) = \|\dot f\|_{\mathrm{TV}}$. Here, we extended the notation from (1.3) by setting, for $\phi \in C_0(0,T;X)$ and $\mu \in \operatorname{rca}(0,T;X^*)$,

$$\langle\!\langle \phi, \mu \rangle\!\rangle := \int_0^T \langle \phi(t), \mu(dt) \rangle\,. \tag{1.7}$$

This extension is consistent with (1.3) if one identifies elements of $L^1(0,T;X^*)$ with densities of measures.

We briefly note that there are related notions of derivatives, such as the reduced derivative [26, App. A] and Darboux-sums [25, App. B.5]. The Kurzweil Integral provides a further approach to the integrals in (1.6), see e.g. [19] and [20].

1.2 Generalisation to metric space codomains

Many of the concepts, like the notions of $\mathrm{pvar}(\cdot)$ and $\mathrm{epvar}(\cdot)$, can be easily generalised to BV-functions taking values in a metric space (Z,d). On the other, some notions, like the time derivative, the variation and the various notions of convergence described above are less straight-forward. To define these notions we will use the embedding of the metric space Z in the Banach space $\mathrm{Lip}(Z)^*$; the resulting notions are consistent with an alternative generalisation of Ambrosio [1].

We first define the metric space $\mathrm{BV}(0,T;Z)$. Fix a point $z_0 \in Z$. This point will play the role of the zero element in a Banach space; all results that we present are trivially invariant under the choice of this point. We define

$$\rho_{L^1}(f,g) := \int_0^T d\big(f(t),g(t)\big)\mathrm{d}t.$$

and write $f \sim g$ if $f = g$ a.e., that is, if $\rho_{L^1}(f,g) = 0$. Then, the space

$$L^1(0,T;Z) := \Big\{f:(0,T)\to Z \text{ for which } \textstyle\int_0^T d(f(t),z_0)\,\mathrm{d}t < \infty\Big\}\Big/\sim,$$

endowed with the metric $\rho_{L^1}(\cdot,\cdot)$ is a complete metric space whenever Z is complete. We write $f_n \xrightarrow{L^1} f$ whenever $\rho_{L^1}(f_n,f) \to 0$. The pointwise variation of a function $f:(0,T)\to Z$ is easily generalisable as follows:

$$\mathrm{pvar}(f) := \sup_{0<t_0<t_1<\cdots<t_n<T} \sum_{i=1}^n d\big(f(t_{i-1}),f(t_i)\big),$$

where the supremum runs over all finite partitions of the interval $(0,T)$. The essential pointwise variation $\mathrm{epvar}(f)$ is defined as in (1.2). Analogously to Definition 1.1, we define the space of functions of bounded variation by

Definition 1.2.

$$\mathrm{BV}(0,T;Z) := \Big\{f \in L^1(0,T;Z) : \mathrm{epvar}(f) < \infty\Big\},$$

endowed with the metric $\rho_{\mathrm{BV}}(f,g)$ *that we introduce below in* (1.9).

Observe that defining a metric ρ_{BV} is a non-trivial task since $\mathrm{epvar}(f-g)$ is not well-defined. To make progress we introduce an embedding into a larger space that does have a linear structure. Following for example [1], one can note that in the Banach valued case, every predual element $x \in X$ induces a Lipschitz functional $x^* \mapsto \langle x, x^*\rangle$ for which $\langle x, 0\rangle = 0$. Motivated by this observation, the predual space is replaced by the (potentially much larger class of) Lipschitz functions,

$$\mathrm{Lip}(Z) := \Big\{\xi : Z \to \mathbb{R} \text{ for which } \|\xi\|_{\mathrm{Lip}(Z)} < \infty \text{ and } \xi(z_0) = 0\Big\},$$

equipped with the Lipschitz constant as norm:

$$\|\xi\|_{\mathrm{Lip}(Z)} := \sup_{\substack{z_1,z_2\in Z:\\ z_1\neq z_2}} \frac{|\xi(z_2)-\xi(z_1)|}{d(z_2,z_1)}.$$

This Lipschitz norm is basically the global metric slope of ξ, see [5, Defn. 1.2.4], and is sometimes also known as the Cheeger derivative. To mimic the notation for the Banach case while emphasizing the one-sided linearity, we write

$$\langle \xi, z] := \xi(z) \qquad \text{for } \xi \in \mathrm{Lip}(Z) \text{ and } z \in Z, \qquad \text{and}$$

$$\langle\!\langle \phi, f]] := \int_0^T \langle \phi(t), f(t)]\,\mathrm{d}t \qquad \text{for } \phi \in C_0\big(0,T;\mathrm{Lip}(Z)\big) \text{ and } f \in L^1(0,T;Z).$$

We now introduce the canonical embedding

$$\begin{aligned} &\delta_\cdot : Z \to \mathrm{Lip}(Z)^*, \qquad z \mapsto \delta_z \quad \text{with} \\ &{}_{\mathrm{Lip}(Z)}\langle \xi, \delta_z \rangle_{\mathrm{Lip}(Z)^*} := \langle \xi, z] = \xi(z) \quad \forall \xi \in \mathrm{Lip}(Z). \end{aligned} \tag{1.8}$$

This extends in the obvious way to functions taking values in Z so that for $f \in \mathrm{BV}(0,T;Z)$ one has a Banach space valued function δ_f given by $\delta_f(t) = \delta_{f(t)}$ and one can try to analyse f through δ_f which now fits into the framework of Paragraph 1.1. However, the space $\mathrm{Lip}(Z)$ should be interpreted as a generalisation of the dual of Z rather than the predual; this leads to some differences between the Banach and the metric cases. For example we can only define a measure-valued derivative $\dot{\delta}_f$ that takes values in $\mathrm{Lip}(Z)^*$ rather than in Z. Moreover, if $Z = X^*$ happens to be dual Banach space and we use the embedding, it is not immediately clear whether the notions of variation in X^* coincide with the notions in $\mathrm{Lip}(X^*)^*$. In Theorem 2.10 we prove that in the metric setting $\mathrm{epvar}(f) = \mathrm{epvar}(\delta_f)$; we then find in Corollary 2.11 that for $Z = X^*$ all notions of variations coincide. This shows that the following metric on the space $\mathrm{BV}(0,T;Z)$,

$$\rho_{\mathrm{BV}}(f,g) = \rho_{L^1}(f,g) + \mathrm{epvar}(\delta_f - \delta_g), \tag{1.9}$$

generalises the distance induced by the norm $\|\cdot\|_{\mathrm{BV}}$ on $\mathrm{BV}(0,T;X^*)$.

1.3 Topologies on $\mathrm{BV}(0,T;X^*)$ and $\mathrm{BV}(0,T;Z)$

The strong topology (induced by $\|\cdot\|_{\mathrm{BV}}$) on BV is too fine for many purposes and in order to achieve convergence and compactness results one introduces coarser topologies. For related discussions see Ambrosio et al. [4, Defn. 3.11], and for a closely related approach developed from a stochastic process perspective, Jakubowski [18] and Bertini et al. [7].

We now present four distinct topologies for the space $\mathrm{BV}(0,T;X^*)$, in decreasing order of fineness. All four topologies have equivalent formulations in terms of open (semi-)balls. To simplify presentation we define these topologies by their corresponding notions of convergence. We emphasize that the definition of vague convergence given in Definition A.12 uses dual pairings with functions in $C_0(0,T;X)$ and not just $C_c(0,T;X)$. This is important because we are not simply dealing with probability measures.

It should be noted that the hybrid and weak* topologies are not necessarily first-countable, so the topologies are defined through their convergent *nets* rather than through convergent sequences, see for example [11, Section A.2].

Definition 1.3 (Topologies on $BV(0,T;X^*)$)**.** *Let X^* be a dual Banach space, and let $(f_n)_n$ be a net and f an element in* $\mathrm{BV}(0,T;X^*)$*. We say that*

f_n *converges to* f *in the* norm *or* strong *topology whenever:*

$$f_n \rightrightarrows f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{L^1} f \qquad \text{and} \qquad \big\|\dot{f}_n - \dot{f}\big\|_{\mathrm{TV}} \to 0, \tag{1.10}$$

f_n *converges to* f *in the* strict *topology whenever:*

$$f_n \overset{strict}{\rightrightarrows} f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{L^1} f \qquad \text{and} \qquad \|\dot{f}_n\|_{\mathrm{TV}} \to \|\dot{f}\|_{\mathrm{TV}}, \tag{1.11}$$

f_n *converges to* f *in the* hybrid *topology whenever:*

$$f_n \rightrightarrows f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{L^1} f \qquad \text{and} \qquad \dot{f}_n \xrightarrow{\text{vague}} \dot{f}, \tag{1.12}$$

f_n converges to f in the weak-* *topology whenever:*

$$f_n \overset{*}{\rightharpoonup} f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{\text{vague}} f \qquad \text{and} \qquad \dot{f}_n \xrightarrow{\text{vague}} \dot{f}. \tag{1.13}$$

Observe that the strong topology is induced by the norm $\|\cdot\|_{\mathrm{BV}}$. The term *strict* convergence is used in [4, Def. 3.14]. It is slightly stronger than the hybrid convergence, see Proposition A.14, and it is clearly metrisable, see (3.3). The term *weak-** convergence is appropriate since $\mathrm{BV}(0,T;X^*)$ is isometrically isomorphic to a dual space, see [4, Rem. 3.12] and Proposition 3.3. We named the convergence (1.12) *hybrid* since it is a combination of the strong convergence for the functions and weak-* convergence for the distributional time derivatives. We have not (yet) been able to determine if the hybrid topology is a *mixed* topology in the sense of Wiweger [31]; it certainly topologises the two-norm convergence of sequences, which was one of Wiweger's motivations.

For finite-dimensional X, weak-* and hybrid convergence coincide whenever the net is uniformly bounded in the BV-norm. Therefore, the distinction between the two is rarely made explicit. However, this is no longer true in the infinite-dimensional setting, as the following example shows.

Example 1.4. *Suppose $X = X^*$ is a separable Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{N}_0}$, and define the sequence of constant functions $f_n(t) \equiv e_n$. This sequence has uniformly bounded norm $\|f_n\|_{\mathrm{BV}} = \|f_n\|_{L^1(0,T;X^*)} = T$, and $f_n \overset{*}{\rightharpoonup} 0$ but certainly not $f_n \rightrightarrows 0$. In particular, since $\mathrm{BV}(0,T;X^*)$ can be identified with a dual space, Banach-Alaoglu gives compactness of bounded BV-balls in the weak* topology, but not in the hybrid topology.*

Using the canonical embedding $\delta : Z \to \mathrm{Lip}(Z)^*$ and Theorem 2.10 one can easily generalise Definition 1.3 to the metric case:

Definition 1.5 (Topologies on $BV(0,T;Z)$)**.** *Let Z be a metric space and let $(f_n)_n$ be a net and f an element in $\mathrm{BV}(0,T;Z)$. We say that*

f_n converges to f in the strong *topology whenever:*

$$f_n \rightrightarrows_m f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{L^1} f \qquad \text{and} \qquad \big\|\dot{\delta}_{f_n} - \dot{\delta}_f\big\|_{\mathrm{TV}} \to 0, \tag{1.14}$$

f_n converges to f in the strict *topology whenever:*

$$f_n \overset{strict}{\rightrightarrows_m} f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{L^1} f \qquad \text{and} \qquad \|\dot{\delta}_{f_n}\|_{\mathrm{TV}} \to \|\dot{\delta}_f\|_{\mathrm{TV}}, \tag{1.15}$$

f_n converges to f in the hybrid *topology whenever:*

$$f_n \rightrightarrows_m f \qquad :\Longleftrightarrow \qquad f_n \xrightarrow{L^1} f \qquad \text{and} \qquad \dot{\delta}_{f_n} \xrightarrow{\text{vague}}_m \dot{\delta}_f, \tag{1.16}$$

f_n converges to f in the weak-* *topology whenever:*

$$f_n \overset{*}{\rightharpoonup}_m f \qquad :\Longleftrightarrow \qquad \delta_{f_n} \xrightarrow{\text{vague}} \delta_f \qquad \text{and} \qquad \dot{\delta}_{f_n} \xrightarrow{\text{vague}}_m \dot{\delta}_f. \tag{1.17}$$

Observe that the strong topology is indeed induced by the metric (1.9). In our notation, we have included the subscript $_m$, since it is not a priori clear whether these notions coincide with the notions of Definition 1.3 if $Z = X^*$ is a dual Banach space. However, we will show in Proposition 3.22 that the notions of strict convergence indeed coincide, and similarly for the strong convergence, see Remark 3.23. In Propositions 3.2 and 3.11 we show that the metric and Banach versions of the weak-$*$ and hybrid topologies agree at least on sequences; it is still an open question whether the topologies are the same, i.e. whether they agree on all nets.

Our main topological results, which also apply to our generalisations to metric space codomains, are

- that the strict and hybrid and weak-* topologies are separable (Propositions 3.26, 3.13 and 3.7),
- sufficient conditions for precompactness in the weak-*, hybrid and strict topologies (Corollary 3.8, Theorem 3.18 and Proposition 3.27).

1.4 Application to stochastic processes

Most stochastic processes in the literature have càdlàg paths. The most commonly used topology on the space of càdlàg paths is the Skohorod (J1-) topology, which makes the path space into a Polish space [16, Sec. 3.5]. This fact is very useful since general probability theory works very well in Polish spaces, in particular: *(i)* the Portemanteau Theorem, giving several equivalent formulations of narrow convergence of probability measures, *(ii)* Prohorov's theorem, which relates relative compactness to tightness of a sequence of measures, and *(iii)* the fact that tightness plus convergence of the time-slices ('finite-dimensional distributions') implies convergence in path space.

A natural question is then whether these general results are still true on the space of BV paths with any of the topologies that we introduced in Section 1.3. However, none of these topologies are Polish: the strong topology is metrisable but not separable, the strict topology is metrisable but not complete, and the hybrid and weak-$*$ topologies are not metrisable. Therefore the standard tools mentioned above may no longer work in those topologies. Although the weak-$*$ topology is probably too weak for many practical purposes, the hybrid topology turns out to be strong enough to be useful but weak and regular enough to be tractable. We will show in Section 3.2 that it is perfectly normal and completely regular. We will see in Section 4 that this is sufficient to establish the validity of the three results just mentioned. In Corollary 4.4 and Theorem 4.5 we show that the Borel σ-algebra generated by the hybrid topology coincides with the Borel σ-algebra generated by the L^1 topology restricted to BV and also coincides with the product σ-algebra. In 4.6 we show that the Portemanteau Theorem is still valid in the strong, strict and hybrid topologies. In Proposition 4.7 we show that the forward version of the Prohorov Theorem is true for the hybrid topology, and in Theorem 4.8 we show that tightness plus convergence of the finite-dimensional distributions imply convergence of the path measures.

1.5 Overview

The structure of the article is as follows. In Section 2, we study some fundamental properties of $\mathrm{BV}(0,T;Z)$-functions such as the existence of càdlàg-representatives, equivalence of the various notions of variation, and the existence of time derivatives.

In Section 3 we study the properties of the three weaker topologies that we have introduced above and provide a number of convergence, duality, regularity and compactness theorems for each of them.

In Section 4 we derive a number of important results related to stochastic processes. Most notably, we first show a number of facts about the σ-algebras corresponding to the strong, strict and hybrid topologies. Next we provide generalised versions of the Portemanteau Theorem, Prohorov's Theorem, and a criteria for convergence of measures on BV-paths.

In the appendix we recall the notions of Banach-valued measures, integrals against Banach-valued measures, regularisation, and topologies on Banach-valued measures.

2 Properties of BV-functions

We devide this section into four parts. In the first part, we show some continuity properties of BV-functions that follow from purely metric considerations. In particular, we show that every BV-function is continuous up to countably many points and that the minimizer of $\mathrm{epvar}(\cdot)$ is attained by the cadlag representative. In the second part, we show that all concepts of variation coincide in the Banach case, and we prove the same statement for metric valued functions in the third part. The fourth part deals with the existence of a time derivative and its properties.

2.1 Continuity-related properties of BV-Functions

In this section we work with pointwise variation and so we can work with functions taking values in a metric space. A crucial fact will be the existence of right and left limits, which will lead to a proof of Theorem 2.10. The first step is to show the existence of such limits for pointwise defined functions.

Proposition 2.1 (Existence of right and left limits)**.** *Let Z be a complete metric space and $g:(0,T)\to Z$ satisfy* $\mathrm{pvar}(g)<\infty$*. Then g is continuous up to a countable subset of $(0,T)$ and g has left and right sided limits:*

$$g(t-):=\lim_{\substack{s\to t\\ s<t}} g(s)\equiv\lim_{s\nearrow t} g(s),\qquad \textit{and}\qquad g(t+):=\lim_{\substack{s\to t\\ s>t}} g(s)\equiv\lim_{s\searrow t} g(s)$$

for all $t\in(0,T)$, and one sided limits at the end points.

Proof. Let us write

$$\mathrm{pvar}(g;(0,t]):=\sup_{0<t_0<t_1<\cdots<t_n\le t}\sum_{i=1}^n d\big(g(t_{i-1}),g(t_i)\big). \tag{2.1}$$

This is a monotonely increasing function of t and bounded above by $\mathrm{pvar}(g)$ so it has at most countably many jumps. Let t be a continuity point of $\mathrm{pvar}(g;(0,t])$. Then we find

$$\lim_{\tau\to t} d\left(g(t),g(\tau)\right)\le\lim_{\tau\to t}\big|\mathrm{pvar}(g;(0,t])-\mathrm{pvar}(g;(0,\tau])\big|=0\,.$$

Let $t\in[0,T)$. If there were a monotone sequence $t_n\searrow t$ such that $g(t_n)$ did not converge then the sequence cannot be Cauchy, i.e. for some $\epsilon>0$ one can pass to a subsequence such that $d\big(g(t_n),g(t_{n+1})\big)\ge\epsilon$ for all n. This would imply that $\mathrm{pvar}(g)=\infty$, which is a contradiction. Similarly, we prove existence of left limits on $(0,T]$. □

Using this fact, we can now prove that the right and left limits of a BV-equivalence class are well-defined.

Proposition 2.2 (Uniqueness of right and left limits)**.** *Let Z be a complete metric space and $f\in\mathrm{BV}(0,T;Z)$. For a $g:(0,T)\to Z$ with $g=f$ almost everywhere and $\mathrm{pvar}(g)<\infty$, define $f(t+):=g(t+)$ for all $t\in[0,T)$ and $f(t-):=g(t-)$ for all $t\in(0,T]$. Then the right and left limits of f are invariant under the choice of the representative g.*

Proof. Take two representatives g_1,g_2 of f with $g_1(t)=g_2(t)$ almost everywhere and $\mathrm{pvar}(g_1),\mathrm{pvar}(g_2)<\infty$. By Proposition 2.1 there exists a countable set $I\subset\mathbb{N}$ such that $g_1|_{[0,T]\setminus I}$ and $g_2|_{[0,T]\setminus I}$ are continuous and thus $g_1(t)=g_2(t)$ for all $t\notin I$. Using the triangle inequality we infer that

$$\mathrm{pvar}(d(g_1,g_2))=\mathrm{pvar}(\|\delta_{g_1}-\delta_{g_2}\|)\le\mathrm{pvar}(\delta_{g_1})+\mathrm{pvar}(\delta_{g_2})=\mathrm{pvar}(g_1)+\mathrm{pvar}(g_2)\,.$$

Therefore, the left and right limits of $d(g_1,g_2)=0$ exist in all $t\in(0,T)$ with one sided limits at the end points. Hence, $g_1(t-)=g_2(t-)$ for all $t\in(0,T]$ and $g_1(t+)=g_2(t+)$ for all $t\in[0,T)$. □

By Proposition 2.2 we can construct a càdlàg version of a BV-function. We prove here that this version is in fact a minimiser for essential pointwise variation. Later on in Corollary 2.19 we prove that the càdlàg version can be related to the derivative.

Proposition 2.3. *Let Z be a complete metric space and $f\in\mathrm{BV}(0,T;Z)$. Define $f_{\mathrm{cadlag}}(t):=f(t+)$. Then $f_{\mathrm{cadlag}}=f$ a.e. and $\mathrm{epvar}(f)=\mathrm{pvar}(f_{\mathrm{cadlag}})$.*

Proof. First note that $f_{\mathrm{cadlag}}(t)=f(t)$ wherever f is continuous and by Proposition 2.1 the discontinuity points are a countable set and thus of measure 0, which proves the first statement. Because of this, one has $\mathrm{epvar}(f)\le\mathrm{pvar}(f_{\mathrm{cadlag}})$. Suppose the inequality to be strict, then there exists an a $g=f$ a.e. such that $\mathrm{pvar}(g)<\mathrm{pvar}(f_{\mathrm{cadlag}})$. By Proposition 2.2 $f_{\mathrm{cadlag}}\equiv g_{\mathrm{cadlag}}$ and so $\mathrm{pvar}(g)<\mathrm{pvar}(g_{\mathrm{cadlag}})$, thus there exists an $\epsilon>0$ and a finite partition $0<t_0<t_1<\cdots<t_n<t_{n+1}=T$ such that

$$\mathrm{pvar}(g)+\epsilon\le\sum_{i=1}^n d\left(g_{\mathrm{cadlag}}(t_{i-1}),g_{\mathrm{cadlag}}(t_i)\right).$$

However we may also find $s_i \in (t_i, t_{i+1})$ such that

$$\max_{i=0,\dots,n} d\left(g(s_i), g_{\text{cadlag}}(t_i)\right) < \frac{\epsilon}{3n}$$

and thus $\text{pvar}(g) + \epsilon \le \text{pvar}(g) + \frac{2\epsilon}{3}$, which is a contradiction since $\text{pvar}(g) < \infty$. □

We end this sub-section with two results that are instrumental in proving our Compactness Theorem 3.18. For an open subinterval $I \subset (0,T)$ let

$$\text{pvar}(f; I) := \sup_{\substack{0<t_0<t_1<\dots<t_n<T:\\ t_0,t_n\in I}} \sum_{i=1}^{n} d\big(f(t_{i-1}), f(t_i)\big),$$

then the following rule for combining variation holds:

Proposition 2.4. *Let Z be a complete metric space, $0 < T_1 < T_2$ and $f \in \text{BV}(0, T_2; Z)$, then*

$$\text{pvar}(f; (0,T_2)) = \text{pvar}(f; (0,T_1)) + \text{pvar}(f; (T_1,T_2)) + \text{d}(f(T_1-), f(T_1)) + \text{d}(f(T_1), f(T_1+))$$

Lemma 2.5. *Let Z be a complete metric space, $f \in \text{BV}(0,T;Z)$, $\epsilon > 0$ and $\sigma\colon (0,T) \to (0,T)$ be measurable and satisfy $|\sigma(t) - t| \le \epsilon$ for all $t \in (0,T)$, then*

$$\int_0^T \text{d}(f(t), f(\sigma(t)+))\,\text{d}t \le 3\epsilon\,\text{epvar}(f).$$

Proof. By Proposition 2.3 one may replace $f(t)$ with $f_{\text{cadlag}}(t)$. Let n be the largest integer no greater than T/ϵ, then, implicitly intersecting domains of integration with $(0,T)$ the integral can be broken into smaller integrals:

$$\int_0^T \text{d}(f(t), f(\sigma(t)+))\,\text{d}t = \sum_{i=1}^{n+1} \int_{(i-1)\epsilon}^{i\epsilon} \text{d}(f_{\text{cadlag}}(t), f_{\text{cadlag}}(\sigma(t)))\text{d}t$$

However, for $t \in ((i-1)\epsilon, i\epsilon) \cap (0,T)$ it follows that $\sigma(t) \in ((i-2)\epsilon, (i+1)\epsilon) \cap (0,T)$ and so

$$\text{d}(f(t), f(\sigma(t)+)) \le \text{pvar}(f_{\text{cadlag}}; ((i-2)\epsilon, (i+1)\epsilon) \cap (0,T)).$$

Thus

$$\int_0^T \text{d}(f(t), f(\sigma(t)+))\,\text{d}t \le \epsilon \sum_{i=1}^{n+1} \text{pvar}(f_{\text{cadlag}}; ((i-2)\epsilon, (i+1)\epsilon) \cap (0,T))$$

and the result follows after allowing for some triple counting since pvar is subadditive by Proposition 2.4 . □

2.2 Equivalence of notions of variation for Banach-valued functions

We now prove in two parts that all notions of variations coincide.

Proposition 2.6. *Let X^* be a dual Banach space and let $f \in \text{BV}(0,T;X^*)$, then*

$$\text{var}(f) = \text{epvar}(f).$$

Proof. By Proposition 2.2 we can canonically identify any $f \in \text{BV}(0,T;X^*)$ with its càdlàg version, and then by Proposition 2.3 one sees that $\text{epvar}(f) = \text{pvar}(f)$. We can choose $0 < t_0 < t_1 < \dots < t_n < T$ and $\xi_i, \zeta_i \in X$ with $\|\xi_i\|_X, \|\zeta_i\|_X = 1$ such that

$$\begin{aligned}
\text{pvar}(f) &\le \sum_{i=1}^{n} \|f(t_{i-1}) - f(t_i)\| + \epsilon \\
&\le \sum_{i=1}^{n} \Big\{ \|f(t_{i-1}) - f(t_i-)\| + \|f(t_i-) - f(t_i)\| \Big\} + \epsilon \\
&\le \sum_{i=1}^{n} \Big\{ \langle \xi_i, f(t_{i-1}) - f(t_i-) \rangle \Big\} + \sum_{i=1}^{n} \Big\{ \langle \zeta_i, f(t_i-) - f(t_i) \rangle \Big\} + 2\epsilon. \qquad (2.2)
\end{aligned}$$

We now estimate both sums separately. Since $\xi_i \circ f \in \mathrm{BV}(0,T;\mathbb{R})$, each term in the first sum is bounded by the variation $\mathrm{var}(\xi_i \circ f;(t_{i-1},t_i))$ of $\xi_i \circ f$, restricted to the interval (t_{i-1},t_i). Due to this BV-regularity, we can take $\phi_i \in C_c^1(0,T;\mathbb{R})$ with $\operatorname{supp}\phi_i \subset (t_{i-1},t_i)$, $0 \le \phi_i \le 1$ and

$$\langle \xi_i, f(t_{i-1}) - f(t_i-)\rangle \le \mathrm{var}(\xi_i \circ f;(t_{i-1},t_i)) \le \int_{t_{i-1}}^{t_i} \dot\phi_i(t)\,\langle \xi_i, f(t)\rangle \,\mathrm{d}t + \frac{\epsilon}{n}.$$

Now define $\Phi\colon (0,T) \to X$ by $\Phi(t) := \sum_{i=1}^n \phi_i(t)\xi_i$. Then $\Phi \in C_c^1\big(0,T;X\big)$ with $\|\Phi\|_\infty \le 1$ and,

$$\sum_{i=1}^n \Big\{ \langle \xi_i, f(t_{i-1}) - f(t_i-)\rangle \Big\} \le \langle\!\langle \dot\Phi, f\rangle\!\rangle + \epsilon\,.$$

For some $\delta > 0$ one has $\Phi\,|_{\cup_{i=1}^n (t_i-\delta, t_i+\delta)} \equiv 0$. We now exploit this flexibility to deal with possible jumps at the ends of the intervals in the second sum of (2.2). Define $\psi \in C_c^1(\mathbb{R};\mathbb{R})$ through

$$\dot\psi(s) = \begin{cases} -4s & \text{if } 0 \le s < \frac12 \\ 4(s-\frac12) - 2 & \text{if } \frac12 \le s < 1 \\ 0 & \text{if } s \ge 1 \end{cases}, \qquad \dot\psi(s) = -\dot\psi(-s)\,, \qquad \psi(-1) = 0\,.$$

Since $\zeta_i \circ f$ is right continuous and has left limits, we obtain that

$$\lim_{\delta\to0} \frac{1}{\delta}\int_{t_i-\delta}^{t_i+\delta} \dot\psi\Big(\frac{t-t_i}{\delta}\Big)\langle \zeta_i, f(t)\rangle\,\mathrm{d}t = \langle \zeta_i, f(t_i-) - f(t_i)\rangle\,.$$

Thus we can pick δ sufficiently small so that $\Psi(t) := \sum_{i=1}^n \psi\left(\frac{t-t_i}{\delta}\right)\zeta_i \in C_c^1\big(0,T;X\big)$ with $\|\Psi\|_\infty \le 1$ and $\mathrm{supp}\Phi \cap \mathrm{supp}\Psi = \varnothing$, and

$$\sum_{i=1}^n \Big\{ \langle \zeta_i, f(t_i-) - f(t_i)\rangle \Big\} \le \langle\!\langle \dot\Psi, f\rangle\!\rangle + \epsilon.$$

Continuing with (2.2) we find that

$$\mathrm{pvar}(f) \le \langle\!\langle \dot\Phi + \dot\Psi, f\rangle\!\rangle + 4\epsilon \tag{2.3}$$

where even $\Phi + \Psi \in C_c^1\big(0,T;X\big)$ with $\|\Phi+\Psi\|_\infty \le 1$. Since for all $\varepsilon > 0$ we can construct Φ, Ψ such that (2.3) holds, we obtain

$$\mathrm{pvar}(f) \le \sup_{\substack{\Phi\in C_c^1(0,T;X):\\ \|\Phi\|_\infty\le1}} \langle\!\langle \dot\Phi, f\rangle\!\rangle = \mathrm{var}(f).$$

For the converse it is sufficient to establish $\mathrm{var}(f) \le \mathrm{pvar}(f) = \mathrm{epvar}(f)$ since we still identify f with its càdlàg representative. For $n \in \mathbb{N}$ define $f_n \in \mathrm{BV}(0,T;X^*)$ by the piecewise constant approximation

$$f_n(t) := \sum_{i=1}^n f\Big(\tfrac{(i-1)T}{n}\Big)\mathbb{1}_{\left[\frac{(i-1)T}{n},\frac{iT}{n}\right)}(t)$$

and note that these are càdlàg by construction and satisfy$\rho_{L^1}(f_n,f) \to 0$. Further

$$\begin{aligned}\mathrm{var}(f_n) &= \sup_{\substack{\Phi\in C_0^1(0,T;X):\\ \|\Phi\|_\infty=1}} \sum_{i=1}^n \int_{\frac{(i-1)T}{n}}^{\frac{iT}{n}} \Big\langle \dot\Phi(t), f\Big(\tfrac{(i-1)T}{n}\Big)\Big\rangle \mathrm{d}t \\ &= \sup_{\substack{\Phi\in C_0^1(0,T;X):\\ \|\Phi\|_\infty=1}} \sum_{i=1}^n \Big\langle \Phi\Big(\tfrac{iT}{n}\Big) - \Phi\Big(\tfrac{(i-1)T}{n}\Big), f\Big(\tfrac{(i-1)T}{n}\Big)\Big\rangle \\ &\le \sum_{i=1}^{n-1} \Big\| f\Big(\tfrac{(i-1)T}{n}\Big) - f\Big(\tfrac{iT}{n}\Big)\Big\| \quad \le \quad \mathrm{pvar}(f).\end{aligned}$$

Since $\mathrm{var}(f)$ is the supremum over functionals continuous in the L^1 topology, it is lower semicontinuous, so that:

$$\mathrm{var}(f) \le \liminf_n \mathrm{var}(f_n) \le \mathrm{pvar}(f).$$

□

The next result can also be proved by explicit estimates in the style of the previous proof, the proof given uses material presented in the following sections and is therefore somewhat shorter.

Proposition 2.7. *Let X^* be a dual Banach space and let $f \in \mathrm{BV}(0,T;X^*)$, then*

$$\mathrm{var}(f) = \mathrm{varw}(f).$$

Proof. By definition, we find $\mathrm{var}(f) \le \mathrm{varw}(f)$. For the other direction of the equality $\mathrm{var}(f) = \mathrm{varw}(f)$, chose a sequence of functions $\psi_\eta \in C_b^1([0,1);\mathbb{R})$ such that $\psi_\eta(0) = 1$, ψ_η is non-increasing and $\psi_\eta(\eta) = 0$. For an arbitrary $\phi \in C_b^1(0,T;X)$ we write $\phi_\eta(t) := \phi(t)\,(1 - \psi_\eta(t) - \psi_\eta(T-t))$ which is now in $C_0^1(0,T;X)$. Then

$$\begin{aligned}
&\langle \phi(T), f(T-)\rangle - \langle \phi(0), f(0+)\rangle - \langle\!\langle \dot\phi, f\rangle\!\rangle \\
&\quad = \langle \phi(T), f(T-)\rangle - \langle \phi(0), f(0+)\rangle - \langle\!\langle \dot\phi_\eta, f\rangle\!\rangle + \langle\!\langle \dot\phi_\eta - \dot\phi, f\rangle\!\rangle \\
&\quad \le \langle \phi(T), f(T-)\rangle - \langle \phi(0), f(0+)\rangle + \mathrm{var}(f) - \langle\!\langle (\psi_\eta + \psi_\eta(T-\cdot))\,\dot\phi, f\rangle\!\rangle - \langle\!\langle \phi\left(\dot\psi_\eta - \dot\psi_\eta(T-\cdot)\right), f\rangle\!\rangle \\
&\quad \to \langle \phi(T), f(T-)\rangle - \langle \phi(0), f(0+)\rangle + \mathrm{var}(f) - \langle \phi(T), f(T-)\rangle + \langle \phi(0), f(0+)\rangle = \mathrm{var}(f)
\end{aligned}$$

where we used the continuity of ϕ and the existence of left and right limits of f together with $\int_0^T \dot\psi(t)\,\mathrm{d}t = 1$. Taking the supremum over $\phi \in C_b^1(0,T;X)$ proves the claim. □

To summarize the results of Propositions 2.3, 2.6 and 2.7, we now have the equivalence of all notions of variations:

Corollary 2.8. *Let X^* be a dual Banach space, then for all $f \in \mathrm{BV}(0,T;X^*)$*

$$\mathrm{epvar}(f) = \mathrm{pvar}(f_{\mathrm{cadlag}}) = \mathrm{var}(f) = \mathrm{varw}(f).$$

2.3 Equivalence of notions of variations of metric-valued functions

In this short section we investigate the relation between the variations in the metric and the Banach setting. As explained in Section 1.2, the canonical embedding $\delta : Z \to \mathrm{Lip}(Z)^*$ plays a crucial role. We note that the space $\mathrm{Lip}(Z)$ is never empty—just consider $\xi(z) := d(z, z_0)$, and that the embedding is continuous and injective, due to the following result. The following result is trivial to obtain but it can be considered the heart of the concept of δ_f.

Lemma 2.9. *For all $z_1, z_2 \in Z$*

$$\|\delta_{z_1} - \delta_{z_2}\|_{\mathrm{Lip}(Z)^*} = \sup_{\substack{\xi\in\mathrm{Lip}(Z):\\ \|\xi\|_{\mathrm{Lip}(Z)}=1}} \langle \xi, z_1] - \langle \xi, z_2] = d(z_1, z_2). \tag{2.4}$$

Proof. The inequality $\sup_{\|\xi\|_{\mathrm{Lip}(Z)}=1} \langle \xi, z_1] - \langle \xi, z_2] \le d(z_1, z_2)$ holds by definition. Equality follows for the choice $\xi(z) := d(z, z_2)$. □

This lemma guarantees that $\delta_f \in \mathrm{BV}(0,T;\mathrm{Lip}(Z)^*))$ if $f \in \mathrm{BV}(0,T;Z)$. More precisely, we have

Theorem 2.10. *Let Z be a complete metric space. For every $g : (0,T) \to Z$,*

$$\mathrm{pvar}(g) = \mathrm{pvar}(\delta_g).$$

and for every $f \in \mathrm{BV}(0,T;Z)$,

$$\mathrm{epvar}(f) = \mathrm{epvar}(\delta_f).$$

Proof. The first statement is a simple consequence of equation (2.4).

For the second statement, use $\delta_f(t) = \delta_{f(t)}$ and continuity of the mapping δ to see that $(\delta_f)_{\mathrm{cadlag}} \equiv \delta_{(f_{\mathrm{cadlag}})}$. Together with the first statement this implies that

$$\operatorname{epvar}(f) = \operatorname{pvar}(f_{\mathrm{cadlag}}) = \operatorname{pvar}\big(\delta_{(f_{\mathrm{cadlag}})}\big) = \operatorname{pvar}((\delta_f)_{\mathrm{cadlag}}) = \operatorname{epvar}(\delta_f).$$

□

Combining this result with Theorem 2.8 yields:

Corollary 2.11. *Let X^* be a dual Banach space, then for every $f \in \mathrm{BV}(0,T;X^*)$*

$$\operatorname{var}(f) = \operatorname{varw}(f) = \operatorname{epvar}(f) = \operatorname{epvar}(\delta_f) = \operatorname{varw}(\delta_f) = \operatorname{var}(\delta_f).$$

Remark 2.12. *(i) As a by-product of Corollary 2.11 it also follows that*

$$\begin{aligned}\operatorname{var}(f) &= \sup_{\phi \in C_0^1(0,T;X)} \int_0^T \big\langle \dot\phi(t), f(t)\big\rangle \,\mathrm{d}t \\ &= \sup_{\psi \in C_0^1(0,T;\mathrm{Lip}(X^*))} \int_0^T \big\langle \dot\psi(t), f(t)\big] \,\mathrm{d}t = \operatorname{var}(\delta_f),\end{aligned}$$

and so the space of test functions used in defining $\operatorname{var}(\cdot)$ *for* $\mathrm{BV}(0,T;X^*)$ *can be extended from* $C_0\big(0,T;X\big)$ *to* $C_0\big(0,T;\mathrm{Lip}(X^*)\big)$ *(which includes* $C_0\big(0,T;X^{**}\big)$*) without changing what is meant by variation.*

(ii) There is another well-established concept for the variation of a metric-valued function, which was studied by Ambrosio [1]: Given $f\colon (0,T) \to Z$, the variation of f is defined to be the smallest measure $\sigma_f \in \mathrm{rca}(0,T;\mathbb{R})$ such that

$$\forall \varphi \in \mathrm{Lip}(Z),\ B \subset (0,T) \quad \sigma_f(B) \geq |D\varphi(f)|(B)\,.$$

This definition is shown by Ambrosio to coincide with $\operatorname{epvar}(f)$*. Our approach allows us to identify σ_f with $|\dot\delta_f|$.*

2.4 Time derivatives of BV-functions

In this section we introduce the measure-valued time derivative of a function of bounded variation, and prove a number of properties related to this derivative. Firstly, such a derivative exists:

Theorem 2.13 (Existence of measure-valued derivatives)**.** *Let $f \in \mathrm{BV}(0,T;Z)$.*

(i) If $Z = X^*$ is a dual Banach space, then there exists a unique finite measure $\dot f \in \mathrm{rca}\big(0,T;X^*\big)$ with $\|\dot f\|_{\mathrm{TV}} = \operatorname{var}(f)$ and such that

$$-\langle\!\langle \dot\phi, f\rangle\!\rangle = \langle\!\langle \phi, \dot f\rangle\!\rangle \qquad \text{for all } \phi \in C_0^1(0,T;X)\,. \tag{2.5}$$

(ii) If Z is a complete metric space, then there exists a unique finite measure $\dot\delta_f \in \mathrm{rca}\big(0,T;\mathrm{Lip}(Z)^*\big)$ with $\|\dot\delta_f\|_{\mathrm{TV}} = \operatorname{var}(f)$ and such that

$$-\langle\!\langle \dot\phi, f]] = \langle\!\langle \phi, \dot\delta_f\rangle\!\rangle \qquad \text{for all } \phi \in C_0^1\big(0,T;\mathrm{Lip}(Z)\big)\,. \tag{2.6}$$

Proof. The proof is as outlined in the introduction (see (1.6)). The mapping $\phi \mapsto -\langle\!\langle \dot\phi, f\rangle\!\rangle$ is clearly linear and

$$\big|-\langle\!\langle \dot\phi, f\rangle\!\rangle\big| \leq \|\phi\|_{C_0(0,T;X)} \operatorname{var}(f) \qquad \forall \phi \in C_0^1(0,T;X)\,.$$

By denseness and by the Banach-valued Riesz-Markov-Kakutani Theorem A.7 the claim follows. The proof for the metric case is the same if we replace X by $\mathrm{Lip}(Z)$. □

Remark 2.14. *A word of warning is appropriate here: even in the case $Z = X^*$ and for differentiable f one does not in general have $\delta_{\dot{f}} = \dot{\delta}_f$ viewed as elements of* $\mathrm{rca}\,(0,T;\mathrm{Lip}(X^*)^*)$. *To see this, take a $f \in W^{1,1}(0,T;X^*)$ and a $\phi \in C_0^1\big(0,T;C_b^1(X^*)\big)$ and apply two partial integrations to get:*

$$\langle\!\langle \phi, \dot{\delta}_f \rangle\!\rangle = \int_0^T {}_{X^{**}}\big\langle \nabla_f \langle \phi(t), f(t)], \dot{f}(t) \rangle_{X^*} \,\mathrm{d}t, \qquad \textit{and}$$
$$\langle\!\langle \phi, \delta_{\dot{f}} \rangle\!\rangle = \int_0^T \langle \phi(t), \dot{f}(t)] \,\mathrm{d}t.$$

Hence in general the two only agree when integrated against test functions that are linear in f, that is $\phi \in C_0(0,T;X)$.

In many cases the measure-valued derivative can itself be identified with a function. This is captured by the following definition and result.

Definition 2.15. *Let Z be a metric space and let $f \in \mathrm{BV}(0,T;Z)$. We say that f is p-absolutely continuous, if there exists $v \in L^p(0,T)$ such that*

$$d\,(f_{\mathrm{cadlag}}(t) - f_{\mathrm{cadlag}}(\tau)) \le \int_t^\tau v(s)\,\mathrm{d}s \qquad \forall 0 \le t \le \tau \le T\,. \tag{2.7}$$

In the next lemma we show that *p-absolutely continuity* is equivalent to L^p-regularity of the derivative . This result is known in the literature for reflexive Banach spaces (see for instance [5, Rem. 1.1.3]).

Lemma 2.16 (Absolute continuity)**.** *Let Z be a complete separable metric space. Then $f \in \mathrm{BV}(0,T;Z)$ is p-absolutely continuous if and only if $\dot{\delta}_f \in L^p(0,T;\mathrm{Lip}(Z)^*)$. Let $Z = X^*$ be a dual Banach space. Then $f \in \mathrm{BV}(0,T;X^*)$ is p-absolutely continuous if and only if $\dot{f} \in L^p(0,T;X^*)$. Furthermore, $v = |\dot{\delta}_f|$ is optimal in* (2.7)*.*

Proof. We first prove the Banach case. Let $\dot{f} \in L^p(0,T;X^*)$. Then (2.7) holds for $v(t) := \|\dot{f}(t)\|_{X^*}$. On the other hand, let (2.7) hold. As explained in Section A.1, we may take the supremum over finite sub-intervals in the definition of the $\mathbb{R}$-valued measure $|\dot{f}|$. Therefore we get that for every interval $(a,b] \subset (0,T)$,

$$\begin{aligned} |\dot{f}|(a,b] = \sup_{a<t_0<\dots\le t_n \le b} \sum_{i=1}^n \big\|\dot{f}\big((t_{i-1},t_i]\big)\big\|_{X^*} &\overset{(2.8)}{=} \sup_{a<t_0<\dots\le t_n \le b} \sum_{i=1}^n \big\|f(t_i+) - f(t_{i-1}+)\big\|_{X^*} \\ &\overset{(2.7)}{\le} \sup_{a<t_0<\dots\le t_n \le b} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} v(s)\,\mathrm{d}s \le \int_{(a,b]} v(s)\,ds. \end{aligned}$$

By the classical Nikodym theorem, there exists a measurable function $w : (0,T) \to [0,1]$ such that $|\dot{f}|(\mathrm{d}s) = v(s)w(s)\,\mathrm{d}s$. By the generalized Lebesgue-Nikodym Theorem A.8, there exists $u : (0,T) \to X^*$ with $\|u(t)\|_{X^*} \le 1$ for every $t \in (0,T)$ such that for all $\phi \in C_c(0,T;X)$

$$\langle\!\langle \phi, \dot{f} \rangle\!\rangle = \int_0^T \langle \phi(t), u(t) \rangle\, |\dot{f}|(\mathrm{d}t) = \int_0^T \langle \phi(t), u(t) \rangle\, v(t) w(t)\,\mathrm{d}t\,.$$

This implies $\dot{f} = uvw$ and hence $\|\dot{f}\|_{L^p(0,T;X^*)} \le \|v\|_{L^p(0,T)}$ and $v = |\dot{f}|$ is optimal in (2.7).

The general metric case now follows immediately. □

We now show how a function of bounded variation can be reconstructed from its derivative. For this, we first need two lemmas which can be proved the same way as in [4, Example 1.75, Proposition 3.2 & Theorem 3.27].

Lemma 2.17. *Let $Z = X^*$ be a dual Banach space, $\mu \in \mathrm{rca}(0,T;X^*)$, and define $g(t) := \mu\big((0,t]\big)$ for all $t \in (0,T)$. Then $g \in \mathrm{BV}(0,T;X^*)$ and $\mu = \dot{g} = \partial g$, where ∂g is the Stieltjes measure from Theorem A.3.*

Lemma 2.18. *Let $Z = X^*$ be a dual Banach space. If $u, v \in \mathrm{BV}(0,T;X^*)$ such that $\dot{u} = \dot{v}$, then $u \equiv v + c$ for some constant $c \in X^*$.*

As a corollary of Lemmas 2.17, 2.18 and Theorem A.3 we obtain the following.

Corollary 2.19. *Let X be a Banach space with dual space X^* and let $f \in \mathrm{BV}(0,T;X^*)$. Then the càdlàg version of Proposition 2.3 can be written as*

$$f_{\mathrm{cadlag}}(t) = \dot{f}\big((0,t]\big) + f(0+). \tag{2.8}$$

Similarly, if Z is a metric space and $f \in \mathrm{BV}(0,T;Z)$ then

$$\delta_{f_{\mathrm{cadlag}}(t)} = \big(\delta_f\big)_{\mathrm{cadlag}}(t) = \dot{\delta}_f\big((0,t]\big) + \delta_{f(0+)}.$$

We now prove a number of useful results related to the derivative: a product rule, a mollification, an approximation result, and an integration by parts formula. The following two statements (Proposition 2.20 and Theorem 2.21) can be proved along the same lines as [4, Prop. 3.2, Thm. 3.9], using Lemma's A.10 and A.11 from Appendix A.3.

Proposition 2.20. *Let X be a separable Banach space and $f \in \mathrm{BV}(0,T;X^*)$.*

1. *For any Lipschitz function $\psi : (0,T) \to \mathbb{R}$ the product $f\psi \in \mathrm{BV}(0,T;X^*)$ and*

$$\frac{\mathrm{d}}{\mathrm{d}t}(f\psi) = f\dot{\psi} + \dot{f}\psi\,.$$

2. *If $\psi_\eta \in C_c^\infty(\mathbb{R})$ is a mollifier with $\operatorname{supp}\psi_\eta \subset [-\eta,\eta]$ then*

$$\frac{\mathrm{d}}{\mathrm{d}t}(\psi_\eta * f)(t) = \big(\psi_\eta * \dot{f}\big)(t) \qquad \forall t \in (\eta, T-\eta)\,.$$

Theorem 2.21 (Approximation of $\mathrm{BV}(0,T;X^*)$-functions by smooth functions.)**.** *Let X be a separable Banach space and let $f \in L^1(0,T;X^*)$. Then $f \in BV(0,T;X^*)$ if and only if there exists a sequence of functions $(f_\epsilon)_{\epsilon>0}$ with $f_\epsilon \in C^\infty(0,T;X^*)$ such that for all ϵ it holds*

$$\|f - f_\epsilon\|_{L^1(0,T;X^*)} < \epsilon\,, \qquad \|\dot{f}_\epsilon\|_{L^1(0,T;X^*)} \le \mathrm{epvar}(f) + \epsilon. \tag{2.9}$$

Finally we give an integration by parts formula for the general metric case:

Theorem 2.22 (Integration by parts)**.** *Let Z be a complete metric space, and $f \in \mathrm{BV}(0,T;Z)$ and $\phi \in W^{1,\infty}\big(0,T;\mathrm{Lip}(Z)\big)$ (ϕ is identified with its Lipschitz continuous representative). Then*

$$\langle\!\langle \dot{\phi}, f]] + \langle\!\langle \phi, \dot{\delta}_f \rangle\!\rangle = \langle \phi(T-), f(T-)] - \langle \phi(0+), f(0+)].$$

Proof. Let $k \in \mathbb{N}$ satisfy $1/k < T/4$. Choose $\chi, \rho \in C_c^2(\mathbb{R};\mathbb{R})$ such that $\mathbb{1}_{[1/k,T-1/k]} \ge \chi \ge \mathbb{1}_{[2/k,T-2/k]}$ and $\rho \ge 0$ with support contained in $[-1/2k, 1/2k]$ and $\int \rho(t)\mathrm{d}t = 1$. Define $\phi_k \in C_c^2\left(0,T;\mathrm{Lip}(Z)\right)$ by $\phi_k := (\chi\phi) * \rho$. Then, since $k > 4/T$,

$$\begin{aligned}
&\left|\langle\!\langle \dot{\phi}, f]] + \langle\!\langle \phi, \dot{\delta}_f \rangle\!\rangle - \phi(T-)\left(f(T-)\right) + \phi(0+)\left(f(0+)\right)\right| \\
&\qquad \le \left|\langle\!\langle \dot{\phi}, f]] - \langle\!\langle \dot{\phi}_k, f]] - \phi(T-)\left(f(T-)\right) + \phi(0+)\left(f(0+)\right)\right| \\
&\qquad\qquad + \left|\langle\!\langle \dot{\phi}_k, f]] + \langle\!\langle \phi_k, \dot{\delta}_f \rangle\!\rangle\right| + \left|\langle\!\langle \phi, \dot{\delta}_f \rangle\!\rangle - \langle\!\langle \phi_k, \dot{\delta}_f \rangle\!\rangle\right|.
\end{aligned}$$

Now $\left|\langle\!\langle \dot{\phi}_k, f]] + \langle\!\langle \phi_k, \dot{\delta}_f \rangle\!\rangle\right|$ is 0 by the definition of $\dot{f}$ since $\phi_k \in C_c^2\left(0,T;\mathrm{Lip}(Z)\right)$. Also for every $t \in (0,T)$ $\lim_{k\to\infty}\phi_k(t) = \phi(t)$ and $\|\phi_k\|_{L^\infty(0,T;\mathrm{Lip}(Z))} \le \|\phi\|_{L^\infty(0,T;\mathrm{Lip}(Z))}$ so by dominated convergence

$$\left|\langle\!\langle \phi, \dot{\delta}_f \rangle\!\rangle - \langle\!\langle \phi_k, \dot{\delta}_f \rangle\!\rangle\right| \le \int_0^T \|\phi(t) - \phi_k(t)\|_{\mathrm{Lip}(Z)} \left|\dot{\delta}_f\right|(\mathrm{d}t) \xrightarrow{k\to\infty} 0.$$

The remaining term can be estimated by noting $\frac{\mathrm{d}}{\mathrm{d}t}\phi_k \equiv (\chi\dot{\phi}) * \rho + (\dot{\chi}\phi) * \rho$, using Lemma 2.5 and noting that if $s, t \searrow 0$ then $\phi(s)\,(f(t)) \to \phi(0+)\,(f(0+))$ along with the analogous result for $T-$.

□

Remark 2.23. *For the case $Z = X^*$ we obtain a stronger integration-by-parts result with $\dot{\delta}_f$ replaced by $\dot{f}$.*

3 Properties of topologies on the space $\mathrm{BV}(0,T;Z)$

In this section we investigate a number of important properties of the topologies introduced in Section 1. Although the norm topology is clearly metrisable, it is not separable and it is rarely possible to establish precompactness results. Therefore we restrict our analysis to the weak*, hybrid and strict topologies. For each of these we will characterise convergence, the dual space, discuss regularity properties, and give sufficient conditions for compactness. Some of the results in this section hold when Z is a general metric space, but many others require the dual Banach space structure $Z = X^*$. These results then also hold in the metric setting after embedding $\mathrm{BV}(0,T;Z)$ into $\mathrm{BV}(0,T;\mathrm{Lip}(Z)^*)$.

3.1 The weak-$*$ topology

Recall from (1.13) and (1.17) that $f_n \overset{*}{\rightharpoonup} f$ whenever $f_n \xrightarrow{\text{vague}} f$ and $\dot{f}_n \xrightarrow{\text{vague}} \dot{f}$, and for the metric case $f_n \overset{*}{\rightharpoonup}_m f$ whenever $\delta_{f_n} \xrightarrow{\text{vague}} \delta_f$ and $\dot{\delta}_{f_n} \xrightarrow{\text{vague}}_m \dot{\delta}_f$.

Characterisation of convergence

Proposition 3.1. *Let $(f_n)_n$ be a net and f an element in $\mathrm{BV}(0,T;Z)$. If $Z = X^*$ is a dual Banach space, then*

$$f_n \overset{*}{\rightharpoonup} f \iff f_n \xrightarrow{\text{vague}} f \quad \text{and} \quad \sup_n \mathrm{var}(f_n) < \infty,$$

and equivalence holds when $(f_n)_n$ is a sequence. If Z is a complete metric space, then the same result holds if we replace $f_n \xrightarrow{\text{vague}} f$ by $\delta_{f_n} \xrightarrow{\text{vague}} \delta_f$.

Proof. Let X^* be a dual Banach space and $f_n \xrightarrow{\text{vague}} f$ a convergent net in $\mathrm{BV}(0,T;Z)$ with $\sup_n \mathrm{var}(f_n) < \infty$. Now approximate an arbitrary test function $\phi \in C_0(0,T;X)$ by a sequence $(\phi_k)_k \subset C_0^1(0,T;X)$ such that $\|\phi - \phi_k\|_\infty \to 0$. Since the variation is lower semicontinuous in the vague topology, we automatically get $\mathrm{var}(f) \le \sup_n \mathrm{var}(f_n)$. It then follows that

$$\begin{aligned}
\left|\langle\langle \phi, \dot{f}_n\rangle\rangle - \langle\langle \phi, \dot{f}\rangle\rangle\right| &\le \left|\langle\langle \phi_k, \dot{f}_n - \dot{f}\rangle\rangle\right| + \left|\langle\langle \phi - \phi_k, \dot{f}_n - \dot{f}\rangle\rangle\right| \\
&\le \left|\langle\langle \dot{\phi}_k, f_n - f\rangle\rangle\right| + 2\|\phi - \phi_k\|_\infty \sup_{\hat{n}} \mathrm{var}(f_{\hat{n}}) \\
&\underset{n}{\to} 2\|\phi - \phi_k\|_\infty \sup_{\hat{n}} \mathrm{var}(f_{\hat{n}}) \underset{k}{\to} 0,
\end{aligned}$$

which together with $f_n \xrightarrow{\text{vague}} f$ shows that $f_n \overset{*}{\rightharpoonup} f$.

On the other hand, if $(f_n)_n$ is a weak-* convergent *sequence*, then $\sup_n \langle\langle \phi, f_n\rangle\rangle < \infty$ and by Banach-Steinhaus it follows that $\sup_n \mathrm{var}(f_n) < \infty$.

The proof in the metric case is analogous once we replace f and $\dot{f}$ by δ_f and $\dot{\delta}_f$. □

We now compare the Banach-case Definition 1.3 of convergence with the metric-case Definition 1.5.

Proposition 3.2. *Let $Z = X^*$ be a dual Banach space. Let $(f_n)_n$ be a net and f an element in* $\mathrm{BV}(0,T;X^*)$. *Then*

$$f_n \rightrightarrows f \qquad \Longleftarrow \qquad f_n \rightrightarrows_m f.$$

Moreover, if $\sup_n \operatorname{epvar}(f_n) < \infty$, *for example because $(f_n)_n$ is a sequence, then the implication is in fact an equivalence.*

Proof. If $f_n \rightrightarrows_m f$, then $f_n \rightrightarrows f$ is immediate from $C_0(0,T;X) \hookrightarrow C_0(0,T;\mathrm{Lip}(X^*))$. The converse is Proposition 3.1. □

Duality

We first show that the space of functions of bounded variation can itself be regarded as a dual space. This theorem works as in the finite-dimensional case (see [4, Rem. 3.12]) and we include it here to provide the full details. To shorten notation, we introduce the spaces:

$$\Phi := C_0(0,T;X) \times C_0(0,T;X) \qquad \text{and} \qquad \Phi_{\partial t} := \big\{(\dot\phi_2,\phi_2) : \phi_2 \in C_c^\infty(0,T;X)\big\} \subset \Phi,$$

both equipped with the uniform norm $\|\phi\|_\Phi := \sup_{t\in(0,T)}\|\phi_1(t)\|_X + \sup_{t\in(0,T)}\|\phi_2(t)\|_X$.

Theorem 3.3. *Let $Z = X^*$ be a dual Banach space. Then the Banach space* $\big(\mathrm{BV}(0,T;X^*), \|\cdot\|_{\mathrm{BV}}\big)$ *is isometrically isomorphic to* $\big(\Phi/\overline{\Phi_{\partial t}}\big)^*$, *and the weak-* convergence corresponds to the convergence defined in* (1.13).

Proof. Observe that for any $f \in \mathrm{BV}(0,T;X^*)$, by Theorem 2.13 the derivative $\dot f$ is well-defined as an object in $C_0(0,T;X)^* \cong \mathrm{rca}(0,T;X^*)$. Define the map $T : \mathrm{BV}(0,T;X^*) \to \Phi^*$ by $Tf := (\hat f, \dot f)$, where $\hat f(\mathrm{d}t) := f(t)\,\mathrm{d}t$. We can then characterise the annihilator of the closure $\overline{\Phi_{\partial t}}$ as (see [29, Sec. 4.6])

$$\begin{aligned} \overline{\Phi_{\partial t}}^{\perp} &:= \big\{\mu \in \Phi^* : \langle \mu, \phi\rangle = 0 \text{ for all } \phi \in \overline{\Phi_{\partial t}}\big\} \\ &= \big\{\mu \in \Phi^* : \langle \mu, \phi\rangle = 0 \text{ for all } \phi \in \Phi_{\partial t}\big\} \\ &= \operatorname{Ran} T. \end{aligned} \tag{3.1}$$

The first equality follows immediately from the fact that $\Phi_{\partial t}$ is strongly dense in its own closure. For the second equality, the direction $\supseteq$ follows immediately from the definitions of $\dot f$ and $\Phi_{\partial t}$. For the direction $\subseteq$, pick a $\mu \in \Phi^*$ for which $\langle \mu,\phi\rangle = \langle \mu_1, \dot\phi_2\rangle + \langle \mu_2,\phi_2\rangle = 0$ for all $\phi_2 \in C_c^\infty(0,T;X)$. If we define $f(t) := \mu_2\big((0,t]\big)$, then Lemma 2.17 yields that $f \in \mathrm{BV}(0,T;X^*)$ and $\langle f, \dot\phi_2\rangle = -\langle \mu_2,\phi_2\rangle = \langle \mu_1, \dot\phi_2\rangle$ for all $\phi_2 \in C_c^\infty(0,T;X)$. We therefore find that $\mu_1(\mathrm{d}t) = f(t)\,\mathrm{d}t$, and hence indeed $\mu \in \operatorname{Ran} T$, which proves the equality (3.1).

Exploiting (3.1), by [29, Th. 4.9(b)] there exists a isometric isomorphism $\tau : \big(\Phi/\overline{\Phi_{\partial t}}\big)^* \to \overline{\Phi_{\partial t}}^{\perp} = \operatorname{Ran} T$. It is easily verified that $\|f\|_{\mathrm{BV}} = \|Tf\|_{\Phi^*}$. Therefore the map $\tau^{-1}\circ T$ is an isometric isomorphism between $\mathrm{BV}(0,T;X^*)$ and $\big(\Phi/\overline{\Phi_{\partial t}}\big)^*$.

Finally, the desired weak* convergence is characterised by convergence against $\Phi/\overline{\Phi_{\partial t}}$. In fact, it again suffices to test against functions in $\Phi/\Phi_{\partial t}$ because of the (strong) density. Then by definition,

$$\begin{aligned} f_n \overset{*}{\rightharpoonup} f :\Longleftrightarrow\ & (\tau^{-1}\circ T)(f_n) \overset{*}{\rightharpoonup} (\tau^{-1}\circ T)(f) \\ :\Longleftrightarrow\ & {}_{\operatorname{Ran} T}\big\langle (f_n\,dx, \dot f_n), \overline{(\psi_1,\psi_2)}\big\rangle_{\Phi/\Phi_{\partial_t}} \to {}_{\operatorname{Ran} T}\big\langle (f\,dx, \dot f), \overline{(\psi_1,\psi_2)}\big\rangle_{\Phi/\Phi_{\partial_t}} \\ & \qquad\qquad \text{for all } \overline{(\psi_1,\psi_2)} = (\psi_1,\psi_2) + (\dot\phi_2,\phi_2) \in \Phi/\Phi_{\partial t} \\ \Longleftrightarrow\ & \langle f_n,\psi_1\rangle + \langle \dot f_n,\psi_2\rangle \to \langle f,\psi_1\rangle + \langle \dot f,\psi_2\rangle \quad \text{for all } \psi_1,\psi_2 \in C_0(0,T;X) \\ \Longleftrightarrow\ & f_n \rightrightarrows f. \end{aligned}$$

□

Now that we have a predual at hand, it is easy to see what the dual space for the weak-* topology is.

Corollary 3.4. *Let $Z = X^*$ be a dual Banach space. Then $\big(\mathrm{BV}(0,T;X^*), \text{weak-}*\big)^*$ is isomorphic to $\Phi/\overline{\Phi_{\partial t}}$.*

Proof. This is a general property of weak-* topologies, see for example [11, Th. V.1.3]. □

Regularity

Using Theorem 3.3 we can deduce many topological properties of the weak-* topology. In general, weak-∗ topologies are not metrisable. Nevertheless, the compact sets are metrisable under a separability assumption. For this we first state the following simple lemma.

Lemma 3.5. *Let X be a Banach space. Then $C_0(0,T;X)$ is separable if and only if X is separable.*

Proof. Let X be separable, with countable dense subset $Q \subset X$. Take a countable dense subset $\Psi \subset C_0(0,T)$. Then the countable set $\big\{ \sum_{i=1}^\infty \psi_i(t) q_i : (\psi_i)_i \subset \Psi, (q_i)_i \subset Q \big\}$ lies dense in $C_0(0,T;X)$. On the other hand, assume that $C_0(0,T;X)$ has a countable dense subset Λ. Take a function $\psi \in C_0(0,T)$ with $\psi(T/2) = 1$. Then for any arbitrary $x \in X$ there exists a sequence $(\lambda_n)_n \subset \Lambda$ such that $\lambda_n \to \psi x$. Let $\pi_{T/2} : C_0(0,T;X) \to X$ with $\pi_{T/2}[\phi] := \phi(T/2)$. By continuity of this evaluation map we get that $\pi_{T/2}[\lambda_n] \to \pi_{T/2}[\psi x] = x$. Hence the countable set $\pi_{T/2}[\Lambda]$ lies dense in X. □

From this we deduce that:

Proposition 3.6. *Let $Z = X^*$ be a dual Banach space. All weak-∗ compact sets in $\mathrm{BV}(0,T;X^*)$ are metrisable if and only if X is separable.*

Proof. By Lemma 3.5 the predual $\Phi/\overline{\Phi_{\partial t}} \subset C_0(0,T;X) \times C_0(0,T;X)$ from Theorem 3.3 is separable if and only if X is separable. The claim then follows from [9, Th. III.25]. □

Proposition 3.7. *Let $Z = X^*$ where X is a separable Banach space. Then the topological space $\big(\mathrm{BV}(0,T;X^*), \text{weak-}*\big)$ is separable.*

Proof. Again by Lemma 3.5 the predual $\Phi/\overline{\Phi_{\partial t}} \subset C_0(0,T;X) \times C_0(0,T;X)$ of $\mathrm{BV}(0,T;X^*)$ is separable. By Corollary 3.4 this space $\Phi/\overline{\Phi_{\partial t}}$ is also the dual of $\big(\mathrm{BV}(0,T;X^*), \text{weak-}*\big)$. It then follows [9, Th. III.23] that $\big(\mathrm{BV}(0,T;X^*), \text{weak-}*\big)$ is also separable. □

Compactness criteria

Again by Theorem 3.3 it is easy to get compactness:

Corollary 3.8. *Let X be a Banach space, then any set of bounded* BV*-norm is relatively compact in $\big(\mathrm{BV}(0,T;X^*), \text{weak-}*\big)$.*

Proof. By Banach-Alaoglu. □

Remark 3.9. *Again after using the embedding $\delta_z : Z \to \mathrm{Lip}(Z)^*$, the same argument applies to the case where Z is a metric space. However, the limit of a relatively compact sequence/net in $\mathrm{BV}(0,T;Z)$ might end up in the bigger space $\mathrm{BV}(0,T;\mathrm{Lip}(Z)^*)$. In an abstract sense such limit can be interpreted as a Young measure.*

3.2 The hybrid topology

Recall from (1.12) and (1.16) that $f_n \rightrightarrows f$ whenever $f_n \xrightarrow{L^1} f$ and $\dot f_n \xrightarrow{\text{vague}} \dot f$, and for the metric case $f_n \rightrightarrows_m f$ whenever $f_n \xrightarrow{L^1} f$ and $\dot\delta_{f_n} \xrightarrow{\text{vague}}_m \dot\delta_f$.

Characterisation of convergence

Proposition 3.10. *Let Z be a complete metric space, $(f_n)_n$ be a net and f an element in* $\mathrm{BV}(0,T;Z)$*, then*

$$f_n \rightrightarrows_m f \quad \Longleftarrow \quad f_n \xrightarrow{L^1} f \quad \text{and} \quad \sup_n \operatorname{var}(f_n) < \infty,$$

and equivalence holds when $(f_n)_n$ is a sequence.

Proof. The convergence $f_n \xrightarrow{L^1} f$ implies $\langle\!\langle \phi, \dot{f}_n \rangle\!\rangle \to \langle\!\langle \phi, \dot{f} \rangle\!\rangle$ for all $\phi \in C_0^1(0,T;X)$ in the case $Z = X^*$ with X Banach, and $\langle\!\langle \phi, \dot{\delta}_{f_n} \rangle\!\rangle \to \langle\!\langle \phi, \dot{\delta}_f \rangle\!\rangle$ for all $\phi \in C_0^1(0,T;\mathrm{Lip}(Z)^*)$ in the general metric case. The proof is then the same as for Proposition 3.1. □

The previous result even holds if one weakens the condition on the variations by only requiring that there is some n_0 in the index set of the net such that $\sup_{n \geq n_0} \operatorname{var}(f_n) < \infty$.

We again compare the Banach-case Definition 1.3 of convergence with the metric-case Definition 1.5.

Proposition 3.11. *Let $Z = X^*$ be a dual Banach space. Let $(f_n)_n$ be a net and f an element in* $\mathrm{BV}(0,T;X^*)$*. Then*

$$f_n \rightrightarrows f \qquad \Longleftarrow \qquad f_n \rightrightarrows_m f.$$

Moreover, if $\sup_n \operatorname{epvar}(f_n) < \infty$, for example because $(f_n)_n$ is a sequence, then the implication is in fact an equivalence.

Proof. The proof is the same as the proof of Proposition 3.2 if we replace $f_n \xrightarrow{\text{vague}} f$ by $f_n \xrightarrow{L^1} f$. □

This results extends to the case where there is an n_0 in the index set of the net such that $\sup_{n \geq n_0} \operatorname{epvar}(f_n) < \infty$, but we are not able to determine whether $f_n \rightrightarrows f$ implies $f_n \rightrightarrows_m f$ for arbitrary nets and thus whether the metric version of the hybrid topology is strictly finer than the Banach version when $Z = X^*$ a dual Banach space. We write τ_{hybrid} for the topology corresponding to $\rightrightarrows$ and $\tau_{\text{hybrid},m}$ for the topology corresponding to $\rightrightarrows_m$.

Duality

Proposition 3.12. *Let $Z = X^*$ be a dual Banach space. Suppose $l\colon \mathrm{BV}(0,T;X^*) \to \mathbb{R}$ is linear and τ_{hybrid}-continuous, then*

$$lf = \Psi f + \langle\!\langle \phi, \dot{f} \rangle\!\rangle$$

for some $\Psi \in L^1(0,T;X^*)^*$ and $\phi \in C_0(0,T;X)$. In particular, if X^{**} has the Radon-Nikodym property with respect to the Lebesgue measure on $(0,T)$, then $\Psi f = \langle\!\langle \psi, f \rangle\!\rangle$ for some $\psi \in L^\infty(0,T;X^{**})$.

Proof. By linearity we only need to consider $|lf| < 1$. Since l is hybrid-continuous, the inverse image $l^{-1}\big((-1,1)\big)$ contains a hybrid-open set from the subbase, centered around 0. Hence one can find an $a \geq 0$, $n \in \mathbb{N}$ and $(\phi_i)_{i=1}^n \subset C_0(0,T;X)$ such that

$$\Big\{ f \in \mathrm{BV}(0,T;X^*) \colon a\, \|f\|_{L^1} < 1, \big|\langle\!\langle \phi_i, \dot{f} \rangle\!\rangle\big| < 1, i = 1, \dots, n \Big\} \subset l^{-1}\big((-1,1)\big).$$

Here, we rescaled a and ϕ_i such that the (semi-) balls have radii 1. Observe that all ϕ_i may be zero; this reflects the fact that l is also L^1-continuous. Similarly a could be 0 as well. Now define $\Phi\colon \mathrm{BV}(0,T;X^*) \to \mathbb{R}^n$ by

$$\Phi f = \big(\langle\!\langle \phi_1, \dot{f} \rangle\!\rangle, \dots, \langle\!\langle \phi_n, \dot{f} \rangle\!\rangle\big)$$

and set

$$\ker(\Phi) = \bigcap_{i=1}^n \big\{ f \in \mathrm{BV}(0,T;X^*) \colon \langle\!\langle \phi_i, \dot{f} \rangle\!\rangle = 0 \big\}.$$

If $a = 0$ then $\ker(\Phi) \subset \ker(l)$ and from [29, Lemma 3.9] it follows that $lf = \langle\!\langle \phi, \dot{f} \rangle\!\rangle$ where $\phi = \sum_{i=1}^{n} \alpha_i \phi_i \in C_0(0,T;X)$ for some numbers $\alpha_i \in \mathbb{R}$.

If $a > 0$, then for $f \in \ker(\Phi)$ with $a \|f\|_{L^1} < 1$ it holds $|lf| < 1$ so $l|_{\ker(\Phi)}$ is L^1-continuous and can be extended to an L^1-continuous linear functional Ψ on all of $\mathrm{BV}(0,T;X^*)$ with norm at most a. We then have $\ker(\Phi) \subset \ker(l - \Psi)$ and so one can now find ϕ as in the case $a = 0$.

Finally, if X^{**} has the Radon-Nikodym property then $L^1(0,T;X^*)^*$ can be identified with $L^\infty(0,T;X^{**})$ [13, Ch. IV, Th. 1]. □

It is easy to see that $\Psi_1 f + \langle\!\langle \phi_1, \dot{f} \rangle\!\rangle = \Psi_2 f + \langle\!\langle \phi_2, \dot{f} \rangle\!\rangle$ for all $f \in \mathrm{BV}(0,T;X^*)$ if and only if $(\Psi_1 - \Psi_2) f = -\langle\!\langle \phi_1 - \phi_2, \dot{f} \rangle\!\rangle$. Thus if X^{**} has the Radon-Nikodym property the hybrid dual space may be identified with

$$L^\infty(0,T;X^{**}) \times C_0(0,T;X) / \left\{ \left(\phi, \dot{\phi}\right) : \phi \in W_0^{1,\infty}(0,T;X) \right\}$$

Regularity

Proposition 3.13. *Let Z be a complete, separable metric space. Then $\left(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m}\right)$ is also separable and if X^* is a dual Banach space then the result is also true for $\left(\mathrm{BV}(0,T;X^*), \tau_{\mathrm{hybrid}}\right)$.*

Proof. The proof of Proposition 3.26 applies in both cases. □

Recall that the hybrid topology is not metrisable, and not even sequential. Nonetheless, we shall see in Section 4 that it has many 'good' properties due to the fact that it is perfectly normal. In order to prove this, we first show that the space is completely regular and Souslin.

Proposition 3.14. *Let Z be a complete, separable metric space and X^* a dual Banach space, then $\left(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m}\right)$ and $\left(\mathrm{BV}(0,T;X^*), \tau_{\mathrm{hybrid}}\right)$ are both Souslin, i.e. the continuous image of a (in this case unspecified) Polish space.*

Proof. For any $n \in \mathbb{N}$ define the balls:

$$A_n := \{ f \in \mathrm{BV}(0,T;Z) : \mathrm{var}(f) \le n \}.$$

By Proposition 3.10 $\tau_{\mathrm{hybrid},m}$ restricted to A_n is the same as the L^1 topology also restricted to A_n. Moreover, the L^1 subspace topology on each A_n is Polish. Thus $\mathrm{BV}(0,T;Z) = \bigcup_n A_n$, which is Souslin by [8, Th. 6.6.6]. The proof for τ_{hybrid} is the same. □

Theorem 3.15. *Let Z be a complete, separable metric space and X^* a dual Banach space. Then $\left(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m}\right)$ and $\left(\mathrm{BV}(0,T;X^*), \tau_{\mathrm{hybrid}}\right)$ are perfectly normal topological spaces.*

Proof. The spaces are Souslin by proposition 3.14 and completely regular, because the hybrid topologies are locally convex topologies. The statement then follows from [8, Theorem 6.7.7]. □

Although the hybrid topology is not metrisable, we do have the following simple result:

Proposition 3.16. *Let Z be a complete separable metric space and X^* a dual Banach space. Then $\tau_{\mathrm{hybrid},m}$ is metrisable on its own compact subsets of $\mathrm{BV}(0,T;Z)$ and τ_{hybrid} is metrisable on its own compact subsets of $\mathrm{BV}(0,T;X^*)$.*

Proof. We mainly follow the idea of [29, Th. 3.16]

□

Compactness criteria

The next two results also hold in the case that $Z = X^*$ a dual Banach space provided one

Theorem 3.17. *Let Z be a complete metric space. If $\mathcal{F} \subset \mathrm{BV}(0,T;Z)$ is relatively compact in $\big(\mathrm{BV}(0,T;Z), \rho_{L^1}\big)$ and $\sup_{f\in\mathcal{F}} \mathrm{epvar}(f) < \infty$ then $\mathcal{F}$ is (both topologically and sequentially) relatively compact in $\big(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m}\big)$ and if $Z = X^*$ a dual Banach space the compactness results also hold with respect to τ_{hybrid}.*

Proof. For any net or sequence in $\mathcal{F}$ there exists a ρ_{L^1}-convergent subnet or subsequence respectively. Recall from Corollary 2.11 that $\mathrm{var}(\delta_f) = \mathrm{epvar}(f)$. Hence by Proposition 3.10 the subnet or subsequence is also hybrid-convergent.

□

Theorem 3.18. *Let Z be a complete metric space, and let $\mathcal{F} \subset \mathrm{BV}(0,T;Z)$ satisfy*

1. *$\sup_{f\in\mathcal{F}} \int_0^T \mathrm{d}(z_0, f(t))\mathrm{d}t + \mathrm{epvar}(f) < \infty$,*
2. *for some countable and dense $Q \subset (0,T)$ there exist compact sets $K_q \subset Z$ with $\bigcup_{f\in\mathcal{F}} \{f(q+)\} \subset K_q$ for all $q \in Q$,*

then $\mathcal{F}$ is (both topologically and sequentially) relatively compact in $\big(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m}\big)$ and if $Z = X^*$ a dual Banach space the result also holds in τ_{hybrid}.

We note that a closely related result was obtained by Mainik and Mielke [23] within the classical pvar-setting of BV-Functions.

Proof. We will establish the relative compactness of $\mathcal{F}$ in the ρ_{L^1} topology and then use Theorem 3.17. The L^1-compactness proof is an adaptation of the standard proof of the Arzela-Ascoli compactness result [15, IV.6, Th. 7] for sets of continuous functions. Since the ρ_{L^1}-topology is clearly metric it is sufficient to show that every sequence in $\mathcal{F}$ has a converging subsequence.

Take a sequence $(f_n)_n$ in $\mathcal{F}$, identify each function with its càdlàg representative. We divide the remaining proof into four steps.

1. Since the K_q is compact for every $q \in Q$, by a diagonal argument we can construct a subsequence, which we will also denote $(f_n)_n$, such that $f_n(q+)$ converges for each $q \in Q$, and we denote these limits $\tilde{f}(q)$.
2. For any $t \in (0,T)$ one can now define
$$\tilde{f}(t) = \lim_{q\in Q, q>t, q\to t} \tilde{f}(q).$$
To see that this limit is well defined, pick an arbitrary $t \in (0,T)$ and suppose the converse. Then as in the proof of Proposition 2.1 there would be an $\epsilon > 0$ and a sequence $(q_i)_i \subset Q$ converging monotonely to t from above such that $d\big(\tilde{f}(q_i), \tilde{f}(q_{i+1})\big) \geq \epsilon$. Then for any $N \in \mathbb{N}$ one could find an n_0 such that $d\big(f_n(q_i+), \tilde{f}(q_i)\big) < \epsilon/N$ for all $n \geq n_0$ and for all $i = 1, \ldots, N$. This would imply that $\mathrm{pvar}(f_n) = \mathrm{epvar}(f_n) \geq (N-2)\epsilon$, which is a contradiction for large N.
3. Let $\epsilon > 0$, $k \in \mathbb{N}$ and $0 < t_0 < t_1 < \cdots < t_k < T$ then by the construction of $\tilde{f}$ one can find $n \in \mathbb{N}$ and $q_i \in Q$ satisfying $0 < q_0 < q_1 < \cdots < q_k < T$ and $q_i > t_i$ such that $\max_i d\big(\tilde{f}(t_i), f_n(q_i)\big) < \epsilon/k$. Thus
$$\mathrm{epvar}\big(\tilde{f}\big) = \mathrm{pvar}\big(\tilde{f}\big) \leq \sup_{g\in\mathcal{F}} \mathrm{epvar}(g).$$
4. Now take an arbitrary $\epsilon > 0$. Since Q is dense one can find $0 < \widetilde{q}_1 < \cdots < \widetilde{q}_N < T$, all in Q such that $\max_{i=2,\ldots,N} (\widetilde{q}_i - \widetilde{q}_{i-1}) < \epsilon$ and $\widetilde{q}_1, T - \widetilde{q}_N < \epsilon$. Let
$$\sigma(t) := \widetilde{q}_1 \mathbb{1}_{(0,\widetilde{q}_1]}(t) + \sum_{i=2}^{N-1} \widetilde{q}_i \mathbb{1}_{(\widetilde{q}_{i-1},\widetilde{q}_i]}(t) + \widetilde{q}_N \mathbb{1}_{(\widetilde{q}_N,T)}(t).$$

Then, using Lemma 2.5 (the f_n and $\tilde{f}$ are càdlàg) and the pointwise convergence of the f_n on Q

$$
\begin{aligned}
&\lim_{n\to\infty} \int_0^T d\big(f_n(t), \tilde{f}(t)\big)\,\mathrm{d}t \\
&\quad \le \lim_{n\to\infty} \int_0^T d\left(f_n(t+), f_n\left(\sigma(t)+\right)\right)\mathrm{d}t + \lim_{n\to\infty} \int_0^T d\big(f_n\left(\sigma(t)+\right), \tilde{f}\left(\sigma(t)\right)\big)\,\mathrm{d}t \\
&\qquad\qquad + \int_0^T d\big(\tilde{f}\left(\sigma(t)\right), \tilde{f}(t)\big)\,\mathrm{d}t \\
&\quad \le 3\epsilon \sup_n \operatorname{epvar}(f_n) + \lim_{n\to\infty} \sum_{i=1}^N \left(\widetilde{q}_i - \widetilde{q}_{i-1}\right) d\left(f_n\left(\widetilde{q}_i+\right), g\left(\widetilde{q}_i\right)\right) + 3\epsilon \operatorname{epvar}(g) \\
&\quad \le 6\epsilon \sup_{g\in\mathcal{F}} \operatorname{epvar}(g).
\end{aligned}
$$

Since $\epsilon > 0$ is arbitrary $\lim_{n\to\infty} \rho_{L^1}(f_n, g) = 0$ and hence $\mathcal{F}$ is indeed L^1 compact.

□

We stress that the above compactness results are stated for a metric space Z without the embedding to the larger space $\mathrm{Lip}(Z)^*$. In particular, relative compactness of a set $\mathcal{F}$ in $\mathrm{BV}(0,T;Z)$ implies that the closure of $\mathcal{F}$ remains in $\mathrm{BV}(0,T;Z)$.

As an important application of Theorem 3.18, we provide the following generalization of the classical Lemma by Aubin and Lions:

Theorem 3.19. *Let $1 \le p < \infty$. Let X^* be the dual of a Banach space and let Y, Z be Banach spaces such that $Y \hookrightarrow Z$ compactly and $Z \hookrightarrow X^*$ continuously. Denote*

$$
\mathrm{BV}^p(0,T;Y,X^*) := \left\{ f \in L^p(0,T;Y) \,:\, \dot{f} \in \mathrm{rca}(0,T;X^*) \right\}.
$$

Then, the embedding $\mathrm{BV}^p(0,T;Y,X^*) \hookrightarrow L^p(0,T;Z)$ is compact.

Proof. By a contradiction argument, we easily verify that any bounded set $\mathcal{F} \subset \mathrm{BV}(0,T;X^*)$ is bounded in $L^\infty(0,T;X^*)$. Hence, due to Simon [30, Section 8, Theorem 5] we only have to show that every bounded set $\mathcal{F} \subset \mathrm{BV}^p(0,T;Y,X^*)$ satisfies

$$
\lim_{h\to 0} \sup_{f\in\mathcal{F}} \|f(h+\cdot) - f(\cdot)\|_{L^1(0,T-h;X^*)} = 0.
$$

This can be verified as follows:

$$
\int_0^{T-h} |f(t+h) - f(t)|\mathrm{d}t \le \int_0^{T-h} \int_t^{t+h} \mathrm{d}|\dot{f}|(s) \le 2h \int_0^T \mathrm{d}|\dot{f}|(s).
$$

□

Let us also mention that the classic result [4, Th. 3.23] coincides with Theorem 3.18 in case $Z = \mathbb{R}$; it is proven via the Arzela-Ascoli theorem for compactness in the space of continuous functions, by substantially the same compact containment condition and diagonal argument as the present theorem. More generally, if Z is locally compact, then Condition 2 of Theorem 3.18 is redundant, since the functions are almost everywhere in bounded Z-balls, which are automatically compact. On the other hand, Condition 2 is not necessary as the following example shows.

Example 3.20. *Suppose $Z = X = X^*$ is a separable Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{N}_0}$, and define the sequence of functions*

$$
f_n(t) := \begin{cases} \big(n(t - \tfrac{1}{2}T) + 1\big)e_n & t \in \left[\tfrac{1}{2}T - \tfrac{1}{n}, \tfrac{1}{2}T\right] \\ \big(n(\tfrac{1}{2}T - t) + 1\big)e_n & t \in \left[\tfrac{1}{2}T, \tfrac{1}{2}T + \tfrac{1}{n}\right] \\ 0 & \textit{otherwise.} \end{cases} \tag{3.2}
$$

Then clearly $\lim_{n\to\infty}\int_0^T \mathrm{d}(0,f_n(t))\,\mathrm{d}t = 0$ *and the derivative also converge vaguely to the* 0*-measure. For any* $t \in (0,T/2)\cup(T/2,T)$*, the value* $f_n(t)$ *lies in the compact set* $\{0\}$ *for* n *large enough, but how large* n *should be depends on* t*. Therefore Condition 2 cannot be satisfied even though the sequence is hybrid convergent.*

3.3 The strict topology

Recall from (1.11) and (1.15) that $f_n \overset{\text{strict}}{\rightrightarrows} f$ whenever $f_n \xrightarrow{L^1} f$ and $\|\dot{f}_n\|_{\mathrm{TV}} \to \|\dot{f}\|_{\mathrm{TV}}$, and for the metric case $f_n \overset{\text{strict}}{\rightrightarrows}_m f$ whenever $f_n \xrightarrow{L^1} f$ and $\|\dot{\delta}_{f_n}\|_{\mathrm{TV}} \to \|\dot{\delta}_f\|_{\mathrm{TV}}$.

Characterisation of convergence

Let Z be a metric space. By its definition, The strict topology on $\mathrm{BV}(0,T;Z)$ is metrizable by

$$\rho_{\text{strict}}(f,g) := \rho_{L^1}(f,g) + \big|\|\dot{\delta}_g\|_{\mathrm{TV}} - \|\dot{\delta}_f\|_{\mathrm{TV}}\big|. \tag{3.3}$$

The following Proposition is stated for convergent sequences and may actually fail for convergent nets, see the discussion in Section A.4. However, since the strict topology is metrisable, it is fully characterised by its convergent sequences.

Proposition 3.21. *Let* Z *be a metric space, and let* $(f_n)_n$ *be a sequence and* f *an element in* $\mathrm{BV}(0,T;Z)$*. If* $Z = X^*$ *is a dual Banach space then*

$$f_n \overset{\text{strict}}{\rightrightarrows} f \quad\Longrightarrow\quad f_n \xrightarrow{L^1} f \quad \text{and} \quad \dot{f}_n \xrightarrow{\text{narrow}} \dot{f}.$$

If Z *is a complete metric space, then*

$$f_n \overset{\text{strict}}{\rightrightarrows}_m f \quad\Longrightarrow\quad f_n \xrightarrow{L^1} f \quad \text{and} \quad \dot{\delta}_{f_n} \xrightarrow{\text{narrow}} \dot{\delta}_f.$$

Proof. For a strictly convergent sequence $f_n \overset{\text{strict}}{\rightrightarrows} f$ we have $\sup_{n\ge N}\mathrm{var}(f_n) < \infty$. By Proposition 3.10 we thus have in particular $\dot{f}_n \xrightarrow{\text{vague}} \dot{f}$. Using Proposition A.14, this implies together with $\|\dot{f}_n\|_{\mathrm{TV}} \to \|\dot{f}\|_{\mathrm{TV}}$ that $\dot{f}_n \xrightarrow{\text{narrow}} \dot{f}$. □

It turns out that if $Z = X^*$, then the topologies induced by $\overset{\text{strict}}{\rightrightarrows}$ and $\overset{\text{strict}}{\rightrightarrows}_m$ coincide:

Proposition 3.22. *Let* $Z = X^*$ *be a dual Banach space. Let* $(f_n)_n$ *be a net and* f *an element in* $\mathrm{BV}(0,T;X^*)$*. Then*

$$f_n \overset{\text{strict}}{\rightrightarrows} f \qquad\Longleftrightarrow\qquad f_n \overset{\text{strict}}{\rightrightarrows}_m f.$$

Proof. Observe that the L^1 norm and metric are equivalent, and that Corollary 2.11 and Theorem 2.13 imply for $Z = X^*$ that $\|\dot{f}\|_{\mathrm{TV}} = \mathrm{epvar}(f) = \mathrm{epvar}(\delta_f) = \|\dot{\delta}_f\|_{\mathrm{TV}}$ and similarly $\|\dot{f}_n\|_{\mathrm{TV}} = \|\dot{\delta}_{f_n}\|_{\mathrm{TV}}$. □

Remark 3.23. *The same argument shows that if* $Z = X^*$*, then strong topologies induced by* $\rightrightarrows$ *and* $\rightrightarrows_m$ *also coincide.*

Remark 3.24. *It is still an open question whether the converse of Proposition 3.21 is also true; since the converse of of Proposition A.14 does not hold (see Remark A.15), the proof of this statement seems to be more involved than in the case* $X = \mathbb{R}$*.*

Duality

Proposition 3.25. *Let $Z = X^*$ be a dual Banach space. Suppose $l: \mathrm{BV}(0,T;X^*) \to \mathbb{R}$ is linear and strictly continuous, then*

$$lf = \Psi f + \langle\langle \phi, \dot{f} \rangle\rangle$$

for some $\Psi \in L^1(0,T;X^*)^*$ and $\phi \in C_b(0,T;X)$. In particular, if X^{**} has the Radon-Nikodym property with respect to the Lebesgue measure on $(0,T)$, then $\Psi f = \langle\langle \psi, f \rangle\rangle$ for some $\psi \in L^\infty(0,T;X^{**})$.

Proof. This exactly follows the proof of Proposition 3.12. □

Regularity

Proposition 3.26. *Let Z be complete and separable (in its metric topology), then $(\mathrm{BV}(0,T;Z), \text{strict})$ is separable.*

Proof. Let $f \in \mathrm{BV}(0,T;Z)$. For $n \in \mathbb{N}$ define the piecewise constant function $g_n \in \mathrm{BV}(0,T;Z)$ by

$$g_n(t) \equiv f\left(\frac{(i-1)T}{n}+\right) \qquad \text{for } t \in \left[\frac{(i-1)T}{n}, \frac{iT}{n}\right), \quad i = 1, \cdots, n.$$

Note firstly that $\rho_{L^1}(g_n, f) \le \mathrm{var}(f)T/n \to 0$ and further that

$$\mathrm{pvar}(f_n) = \sum_{i=1}^n d\Big(f\big(\tfrac{(i-1)T}{n}+\big), f\big(\tfrac{iT}{n}+\big)\Big) \le \mathrm{pvar}(f_{cadlag}) = \mathrm{var}(f).$$

Since Z is separable one can take z_n^i from a countable subset such that

$$d\Big(z_n^i, f\big(\tfrac{iT}{n}+\big)\Big) \le \frac{1}{2n^2} \quad \forall n, i \le n.$$

Define the piecewise constant function $f_n \in \mathrm{BV}(0,T;Z)$ by

$$f_n(t) \equiv z_n^i \qquad \text{for } t \in \left[\frac{(i-1)T}{n}, \frac{iT}{n}\right), \quad i = 1, \cdots, n.$$

Then $\rho_{L^1}(f_n, f) \le \rho_{L^1}(f_n, g_n) + \rho_{L^1}(g_n, f) \to 0$ and $\mathrm{var}(f_n) \le \mathrm{var}(f) + 1/n$ so by the L^1 lower semi-continuity of the variation $\mathrm{var}(f_n) \to \mathrm{var}(f)$. □

The strict metric is not a complete metric: consider $f_n \in \mathrm{BV}(0,1;\mathbb{R})$, $f_n(t) = \mathbb{1}_{(0,1/n)}(t)$. Then $\mathrm{var}(f_n) = 1$ for all n so $\rho_{\text{strict}}(f_n, f_m) = |n-m|/nm$. The sequence f_n is thus Cauchy for ρ_{strict}, but $\rho_{L^1}(f_n, 0) \to 0$ so it cannot converge in the strict metric.

Compactness criteria

As mentioned in Remark 3.24, it is still unclear whether the strict topology is the same as the topology characterised by the convergence

$$f_n \xrightarrow{L^1} f \quad \text{and} \quad \dot{f}_n \xrightarrow{\text{narrow}} \dot{f}, \quad \text{or in the metric case:} \quad f_n \xrightarrow{L^1} f \quad \text{and} \quad \dot{\delta}_{f_n} \xrightarrow{\text{narrow}} \dot{\delta}_f. \tag{3.4}$$

However, it is the latter topology for which we state a compactness result:

Proposition 3.27. *Let Z be a metric space or a dual Banach space, let $\mathcal{F} \subset \mathrm{BV}(0,T;Z)$ be compact in the hybrid topology and suppose additionally that the set $\{|\dot{\delta}_f| : f \in \mathcal{F}\} \subset \mathrm{rca}(0,T)$ is tight. Then $\mathcal{F}$ is compact in the topology characterised by (3.4).*

Proof. By Lemma A.13 every hybridly convergent subnet of an arbitrary net is also convergent in the sense of (3.4). □

4 Application to stochastic processes

We now prove a number of results that provide a basis for considering stochastic processes with bounded variation paths. Because measure theory is based on countable numbers of operations, many results will require the metric topology on Z to be separable.

Firstly we consider possible σ-algebras. As usual, the Borel σ-algebra generated by a topology is the smallest σ-algebra containing all open sets, the Baire σ-algebra is the smallest σ-algebra making all continuous functions measurable. Let $\pi_t : \mathrm{BV}(0,T;Z) \to Z$ with $\pi_t(f) = f(t+)$ for $t \in [0,T)$ and note that by Proposition 2.1 this is well defined on L^1-equivalence classes. The product σ-algebra is defined to be the smallest σ-algebra making all the π_t measurable.

Most results in this section hold true for $\mathrm{BV}(0,T;Z)$ with the norm, strict and hybrid topology. The key fact that makes this work is the following:

Proposition 4.1. *Let Z be a complete metric space, and let $\mathrm{BV}(0,T;Z)$ be equipped with the strong or the strict topology; alternatively, let Z also be separable, and let $\mathrm{BV}(0,T;Z)$ be equipped with the hybrid topology. Then in each of these three cases the Baire and the Borel σ-algebras coincide.*

Proof. The norm and the strict topologies are metrisable and therefore perfectly normal. If Z is separable then by Proposition 3.15 the hybrid topology is perfectly normal. By [8, Prop. 6.3.4] the Baire and Borel σ-algebras of a perfectly normal topological spaces coincide. □

For brevity we will write σ_{L^1} for the Borel σ-algebra generated by the ρ_{L^1} topology on $\mathrm{BV}(0,T;Z)$.

Proposition 4.2. *Let Z be a complete, separable metric space, then the functions π are measurable with respect to σ_{L^1}.*

Proof. The metric topology on Z is separable so every open set can be written as a countable union of open balls $U_{z,\delta} := \{y \in Z : d(y,z) < \delta\}$ for $z \in Z$ and $\delta > 0$. Hence it suffices to prove $\pi_t^{-1}(U_{z,\delta}) \in \sigma_{L^1}$ for arbitrary z and δ.

Now note that

$$f(t+) \in U_{z,\delta} \iff \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \int_t^{t+\epsilon} d(f(s), z)\, \mathrm{d}s < \delta.$$

Further, for fixed ϵ, t, z the function $f \mapsto \frac{1}{\epsilon}\int_t^{t+\epsilon} d(f(s),z)\,\mathrm{d}s$ is L^1-continuous and thus σ_{L^1} measurable, so that the limit as $\epsilon \searrow 0$ is also σ_{L^1} measurable. □

Proposition 4.3. *Let Z be a complete, separable metric space. Then the function $f \mapsto \mathrm{var}(f)$ is measurable with respect to σ_{L^1} on $\mathrm{BV}(0,T;Z)$.*

Proof. Use Proposition 4.2 and note that the variation can be written as the supremum over rational partitions. □

From Proposition 4.3 one sees that $A_n := \{f \in \mathrm{BV}(0,T;Z) : \mathrm{var}(f) \le n\} \in \sigma_{L^1}$. By Proposition 3.10 the hybrid- and L^1-topologies coincide on each A_n and thus the associated Borel σ-algebras are equal on each A_n. In particular, any hybrid open set, U, can be written as

$$U = \bigcup_{n\in\mathbb{N}} U \cap A_n = \bigcup_{n\in\mathbb{N}} V_n \cap A_n \tag{4.1}$$

for some $V_n \in \sigma_{L^1}$. It then follows that:

Corollary 4.4. *The Borel σ-algebras generated by the L^1 and hybrid topologies on $\mathrm{BV}(0,T;Z)$ are identical.*

Theorem 4.5. *Let Z be a complete, separable metric space. Then the Borel σ-algebras of the topological spaces $(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m})$ and $(\mathrm{BV}(0,T;Z), \rho_{L^1})$ are both equal to the product σ-algebra.*

Proof. Using Propositions 4.2&4.4 it is sufficient to prove that all L^1-measurable functions are product measurable. To this end note that

$$\rho_{L^1}(f,g) = \int_0^T d\big(f(t), g(t)\big)\,\mathrm{d}t = \int_0^T d\big(f(t+), g(t+)\big)\,\mathrm{d}t = \lim_{n\to\infty} \tfrac{T}{n}\sum_{i=1}^n d\big(f(\tfrac{iT}{n}+), g(\tfrac{iT}{n}+)\big),$$

and that for fixed $n \in \mathbb{N}$ and $f \in \mathrm{BV}(0,T;Z)$ the mapping $g \mapsto \sum_{i=1}^n d\big(f(\frac{iT}{n}+), g(\frac{iT}{n}+)\big)$ is measurable in the product σ-algebra (defined above as making the evaluations π_t measurable). Since countable limits of measurable functions are measurable, the balls $\{g \in \mathrm{BV}(0,T;Z) : \rho_{L^1}(f,g) < \epsilon\}$ are therefore also measurable in the product σ-algebra. Finally, Z and $\mathrm{BV}(0,T;Z)$ are separable so that every open set can be written as a countable union of such balls. □

We now study the space of Borel probability measures $\mathcal{P}\big(\mathrm{BV}(0,T;Z)\big)$ on the space $\mathrm{BV}(0,T;Z)$ equipped with the norm, strict or hybrid topology. Narrow convergence of probability measures in this space is characterised by the following Portemanteau theorem:

Proposition 4.6 (Portemanteau Theorem, [8, Th. 8.2.10])**.** *Let Z be a metric space, and let $\mathrm{BV}(0,T;Z)$ be equipped with the norm or the strict topology; alternatively, let Z be a separable metric space, and let $\mathrm{BV}(0,T;Z)$ be equipped with the hybrid topology. Let $(\nu_n)_n$ be a net and ν be an element of the Borel probability measures $\mathcal{P}\big(\mathrm{BV}(0,T;Z)\big)$. Then the following statements are equivalent:*

(i) $\nu_n \rightharpoonup \nu$, *that is,* $\int \Phi\, d\nu_n \to \int \Phi\, d\nu$ *for all* $\Phi \in C_b\big(\mathrm{BV}(0,T;Z)\big)$,

(ii) $\limsup_n \nu_n(\mathcal{F}) \le \nu(\mathcal{F})$ *for all closed sets* $\mathcal{F} \subset \mathrm{BV}(0,T;Z)$,

(iii) $\liminf_n \nu_n(\mathcal{U}) \le \nu(\mathcal{U})$ *for all open sets* $\mathcal{U} \subset \mathrm{BV}(0,T;Z)$,

(iv) $\lim_n \nu_n(\mathcal{C}) \le \nu(\mathcal{C})$ *for all continuity sets* $\mathcal{C} \subset \mathrm{BV}(0,T;Z)$, *i.e.* $\nu\big(\overline{\mathcal{C}} \backslash \mathring{\mathcal{C}}\big) = 0$.

Proof. This is again a consequence of the fact that the topologies perfectly normal, see [8, Th. 8.2.10]. □

The forward part of Prohorov's theorem also holds:

Proposition 4.7 (Generalised Prohorov)**.** *Let Z be a complete separable metric space. Then a tight collection of probability measures on $(\mathrm{BV}(0,T;Z), \tau_{\mathrm{hybrid},m})$ is topologically and sequentially compact. If $Z = X^*$ a dual Banach space, then the result also holds for probability measures on $(\mathrm{BV}(0,T;X^*), \tau_{\mathrm{hybrid}})$.*

Proof. Both BV spaces are Souslin by Proposition 3.14 so [8, Theorem 7.4.3] shows that the probability measures are Radon measures (by Theorem 4.5 it is no restriction to assume that they are Borel). Hybrid-compact sets are metrisable by Proposition 3.16 and so the result follows from [8, Theorem 8.6.7]. □

Proposition 4.8. *Let Z be a complete separable metric space, and let $\mathrm{BV}(0,T;Z)$ be equipped with $\tau_{\mathrm{hybrid},m}$, or let $Z = X^*$ be a dual Banach space, and let $\mathrm{BV}(0,T;X^*)$ be equipped with τ_{hybrid}. Let $(\nu_n)_n$ be a sequence or net and ν be an element of the Borel probability measures $\mathcal{P}\big(\mathrm{BV}(0,T;Z)\big)$. For each $t \in (0,T)$ define the finite-dimensional distributions by $\pi_t\#\nu_n(\mathcal{Z}) \coloneqq \nu_n\big(\pi_t^{-1}(\mathcal{Z})\big)$ and similarly for ν. Assume:*

(i) the sequence ν_n is tight;

(ii) the finite-dimensional distributions $\pi_t\#\nu_n \rightharpoonup \pi_t\#\nu$ convergence narrowly for each $t \in (0,T)$.

Then the sequence converges narrowly $\nu_n \rightharpoonup \nu$.

Proof. By Proposition 4.7 the net (or sequence) $(\nu_n)_n$ has a narrowly convergent subnet (or subsequence). Since the finite-dimensional distributions convergence, any cluster point agrees with ν on the finite-dimensional distributions. Because of Theorem 4.5, a measure is uniquely characterised by its finite-dimensional distributions and hence ν is the unique limit. □

A Preliminaries on Banach-valued measures and integration

We summarize the main concepts of the theory of Banach-valued measures. For a deeper insight into this subject, we refer the reader to the classical books [13, 14] and [15, § IV.10] and the recent monograph by Ma [22]. In what follows, we mostly stick to the presentation in [22]. While the Bochner theory of Banach-valued functions has become very popular among analysts, this seems not to be the case for Banach-valued measures. Banach-valued measures are defined similarly to classical measures and one often can prove the intuitive analogues of classical results from ($\mathbb{R}$- or $\mathbb{C}$-) measure theory. The theory of Banach-valued measures is useful to understand the time derivative of a X^*-valued function of bounded variation.

A.1 Banach-valued measures

For $T > 0$, let $\mathcal{B}$ denote the set of all Borel-sets of the interval $(0,T)$ and let X be a Banach space. For a set function $\mu : \mathcal{B} \to X$ we define the set function $|\mu| : \mathcal{B} \to \mathbb{R}$ through ([22, Paragraph 17-4.1])

$$|\mu|(A) := \sup_{P \in \mathcal{Q}(0,T)} \sum_{D \in P} \|\mu(A \cap D)\|_X \qquad \forall A \in \mathcal{B}\,. \tag{A.1}$$

where the supremum is taken over all finite families P of disjoint subsets of $(0,T)$. Observe that it is not a priori clear whether it suffices to take the supremum over intervals $0 < t_1, \cdots < t_n < T$, like in Definition 1.1 of the pointwise variation. However, in the case where A is the full interval $(0,T)$ and $\mu \in \mathrm{rca}(0,T;X^*)$, we get by Theorem A.7 that $|\mu|(0,T) = \|\mu\|_{\mathrm{TV}}$, the total (not bounded) variation norm defined in (A.4).

Definition A.1. *[14, 22, 15] Let X be a Banach space and $\mathcal{B}$ be the Borel σ-algebra on $(0,T)$.*

- *A set function $\mu : \mathcal{B} \to X$ is called an X-valued measure on $(0,T)$ if*
  1. *it is of finite total variation, i.e. $|\mu|(0,T) < \infty$ and*
  2. *it is countably additive, i.e. for every disjoint union $A = \bigcup_{j\in\mathbb{N}} B_j$, where $A \in \mathcal{B}$ and $B_j \in \mathcal{B}$ for all $j \in \mathbb{N}$*
  $$\mu(A) = \sum_{j\in\mathbb{N}} \mu(B_j)\,.$$
- *An X-valued measure μ is called* regular *if for every $A \in \mathcal{B}$ and every $\varepsilon > 0$, there exists a compact (some authors only require closed, but here closed and compact are equivalent) set $K \subset A$ and an open set $G \supset A$ such that for all $A' \in \mathcal{B}$ with $K \subseteq A' \subseteq G$ it holds $\|\mu(A') - \mu(A)\|_X \le \frac{\varepsilon}{2}$.*
- *We denote*
  $$\mathrm{rca}(0,T;X) := \{\mu : \mathcal{B} \to X \quad \textit{regular } X\textit{-valued measure}\},$$
  which is a Banach space with norm $\mu \mapsto |\mu|(0,T) = \|\mu\|_{\mathrm{TV}}$ by Theorem A.7 or [15, Chapter III §7].

We make a few observations about this definition. First of all, this definition only allows for finite measures, that is, measures of finite total variation. This is related to the fact that the measures take values in a Banach space, where each element $x \in X$ is of finite norm (see e.g. [15, IV.10, Cor. 2]). Secondly, we note that the countable additivity is equivalent to $\mu(A_j) \to \mu(A)$ whenever $A_j \supset A_{j+1}$ for all j and $\bigcap A_j = A$ [22, 17-5.4]. Thirdly, we point out that every X-valued regular Borel-measure of finite total variation has the direct sum property of [14, Def. 10-7.1].

Theorem A.2. *Suppose $\mu \in \mathrm{rca}(0,T;X)$ then $|\mu| \in \mathrm{rca}(0,T;\mathbb{R})$.*

Proof. That $|\mu|$ is a measure (a countably additive set function) is the content of [15, III.4.7], however all (real-valued) Borel measures on a metric space are regular by [8, Theorem 7.1.7]. An extensive discussion leads up to the statement of this result as [14, Chapter 3 §15 Proposition 21]. □

For a map $g:(0,T)\to X$, we define the set function ∂g through the sets:

$$\partial g\big((a,b]\big) := g(b) - g(a).$$

Recall the function $t\mapsto \mathrm{pvar}(g;(0,t])$ from (2.1). We then find the following important theorem.

Theorem A.3 ([22, Ths. 17-7.4 and 17-7.9])**.** *Let X be a Banach space and let $g:(0,T)\to X$ be of finite pointwise variation. Then ∂g is a Banach-valued measure on $\mathcal{B}$ if and only if g is right-continuous. In this case, $|\partial g| = \partial_t \mathrm{pvar}_{(0,t]}(g)$ and ∂g is called the Stieltjes' measure induced by g.*

A.2 Integration theory for Banach-valued measures

Let X,Y,Z be Banach spaces with a bilinear continuous mapping $p_{XY}: X\times Y\to Z$, $\mathcal{B}$ the Borel algebra on $(0,T)$ and let $\mu:\mathcal{B}\to Y$ be a Y-valued measure. In this section we will introduce the Z-valued integral $\int_0^T p_{XY}\big(\phi(t),\mu(dt)\big)$. Recall from Appendix A the definition of the $\mathbb{R}$-valued measure $|\mu|$. With this we define the space

$$L^p_\mu(0,T;X,Y) := \Big\{\phi:\,(0,T)\to X\;:\|\phi\|^p_{L^p_\mu(0,T;X,Y)} := \int_0^T \|\phi(t)\|_X^p\,|\mu|(\mathrm{d}t)<\infty\Big\}.$$

A $\mathcal{B}$-step X-map $\phi:(0,T)\to X$ (or simple function) is of the form $\sum_{j=1}^N \alpha_j\chi_{A_j}$, where $\alpha_j\in X$ and χ_{A_j} is the indicator function for the set $A_j\in\mathcal{B}$. We define the step-integral

$$I_{p_{XY}}(\phi) := \sum_{j=1}^N p_{XY}\left(\alpha_j,\mu(A_j)\right) \quad \text{for } \phi = \sum_{j=1}^N \alpha_j\chi_{A_j}. \tag{A.2}$$

This map can be extended to an integral in the following way (see [22, Sec. 21]). First, one can show that for every $\phi\in L^p_\mu(0,T;X,Y)$ there exists a sequence of $\mathcal{B}$-step X-maps ϕ_n, such that $\|\phi_n(t)\|_X\uparrow\|\phi(t)\|_X$ and $\phi_n(t)\to\phi(t)$ for $|\mu|$-almost every $t\in(0,T)$. Then, for such approximating sequences $(\phi_n)_n$, one can show that the limit $\lim_{n\to\infty} I_{p_{XY}}(\phi_n)$ is independent of the choice of the sequence $(\phi_n)_n$. This defines the integral $\int_0^T p_{XY}\big(\phi(t),\mu(\mathrm{d}t)\big)$.

For $1\le p<\infty$ one can show that the set of $\mathcal{B}$-step X-maps is dense in $L^p_\mu(0,T;X,Y)$. This in turn implies that the functions $C_c(0,T;X)$ are dense in $L^p_\mu(0,T;X,Y)$. From [22, Th. 21-2.11], we obtain that for every integrable map $\phi\in L^1_\mu(0,T;X,Y)$, we have

$$\left\|\int_0^T p_{XY}(\phi,\mathrm{d}\mu)\right\|_Z \le \|p_{XY}\|\int_0^T\|\phi\|_X\,\mathrm{d}|\mu|. \tag{A.3}$$

Remark A.4. *The above definition of the integral is very general. Let us mention four possible settings here:*

(A) Let $Y=\mathbb{R}$, $Z=X$ is a Banach space, and $p_{XY}(x,y):=xy$. Then the theory in [22] is the commonly used Bochner theory.

(B) The case $Y=Z$ and $X=\mathbb{R}$ is a further connection to Bochner theory.

(C) Let X and Z be Banach spaces, and let $Y=L(X;Z)$ with $p_{XY}(x,y):=y(x)$.

(D) Let X be a Banach space, $Y=X^*$ and $Z=\mathbb{R}$ with $p_{XY}(x,y):=\langle x,y\rangle$. This is the setting of the main content of this paper.

The following generalisations of the classical Riesz-Markov-Kakutuni result on the duality between $C_c(0,T;\mathbb{R})^*$ and $\mathrm{rca}(0,T;\mathbb{R})$ will turn out to be very useful in the proof of Theorem 2.13. For this we first define:

Definition A.5. *Let X and Z be Banach spaces. A linear mapping $U : C_c(0,T;X) \to Z$ is called dominated if there exists a regular positive Borel measure ν such that*

$$\|U(\phi)\|_Z \le \int_0^T \|\phi(t)\|_X \,\nu(\mathrm{d}t) \qquad \forall \phi \in C_c(0,T;X)\,.$$

Proposition A.6 ([14, §19, Prop. 2 and Th. 3])**.** *Let $U : C_c(0,T;X) \to \mathbb{R}$ be linear. Then U is dominated if and only if $\|U\|_{C_c(0,T;X)^*} < \infty$.*

The previous result will enable us to apply the generalised Riesz-Markov-Kakutani result to the settings (C) and (D) from Remark A.4.

Theorem A.7 ([14, §19, Th. 2])**.** *(i) Assume X and Z are Banach spaces, $Y = L(X;Z)$ with $p_{XY}(x,y) := y(x)$. Then there exists an isomorphism between the dominated linear operators $U : C_c(0,T;X) \to Z$ and $\mathrm{rca}\big(0,T;L(X;Z)\big)$, given by the equality*

$$U(\phi) = \int_0^T p_{XY}(\phi, \mathrm{d}\mu)\,.$$

(ii) Assume X is a Banach space, $Y = X^*$ and $Z = \mathbb{R}$ with $p_{XY}(x,y) := \langle x,y\rangle$. Then there exists an isomorphism between $C_c(0,T;X)^* = C_0(0,T;X)^*$ and $\mathrm{rca}(0,T;X^*)$, with $|\mu|(0,T) = \|U\|_{C_c(0,T;X)^*} = \|\mu\|_{\mathrm{TV}}$, where

$$\|\mu\|_{\mathrm{TV}} := \sup_{\substack{\phi\in C_0(0,T;X):\\ \|\phi\|_\infty\le 1}} \langle\!\langle \phi, \mu\rangle\!\rangle. \tag{A.4}$$

We finally cite the following Lebesgue-Nikodym Theorem. Recall in this context, every X^*-valued Borel-measure μ of finite variation has the direct sum property of [14].

Theorem A.8 (General Lebesgue-Nikodym Theorem, [14, §13, Th. 4])**.** *Let $\mu : \mathcal{B} \to X^*$ be a measure of finite total variation. There exists a function $u : (0,T) \to X^*$ such that $\|u(t)\|_{X^*} = 1$ for $|\mu|$-almost every $t \in (0,T)$ and*

$$\int_0^T \langle \phi, d\mu\rangle = \int_0^T \langle \phi, u\rangle \, d|\mu| \qquad \forall \phi \in C_c(0,T;X)\,.$$

Theorem A.8 is less general than the classical Radon-Nikodym theorem as it only postulates the existence of a density for μ with respect to $|\mu|$ instead of a density with respect to a general real measure ν. If in the statement of Theorem A.8, we would want to replace $|\mu|$ by a general real measure ν, the space X^* would need to satisfy the Radon-Nikodym property. This holds for example if X is reflexive or if X^* is separable, see [13].

Remark A.9. *Using the Lebesgue-Nikodym Theorem, we can decompose any $\mu \in \mathrm{rca}(0,T;X^*)$ into an absolute continuous part μ^c, an atomic part μ^a and a diffuse singular part μ^d (without atoms). This can be seen as follows: The measure $|\mu|$ can be decomposed into absolute continuous, atomic and diffuse singular parts $|\mu| = |\mu|^c + |\mu|^a + |\mu|^d$, see [4] (3.26). By Theorem A.8, there exists an $u : (0,T) \to X^*$ such that $\mu = u|\mu|$ and thus, we can set $\mu^c = u|\mu|^c$, $\mu^a = u|\mu|^a$ and $\mu^d = u|\mu|^d$.*

A.3 Regularisation

In this section we recall some standard regularisation results that are needed to prove Proposition 2.20 and Theorem 2.21.

Given $\mu \in \mathrm{rca}(0,T;X^*)$ and $\psi \in C_c(\mathbb{R})$, we define the convolution through

$$\psi * \mu(t) := \int_{\mathbb{R}} \varphi(t-s)\,\mu(\mathrm{d}s)\,.$$

Of particular interest are measures of the form $\mu(\mathrm{d}t) = f(t)\,\mathrm{d}t$ for some $f \in L^p(0,T;X^*)$. Recall the definition of $L^p_\mu(0,T;X,Y)$ from Appendix A.2. We now have the following lemma:

Lemma A.10. *Let $(\psi_\eta)_{\eta>0} \subset C_c(\mathbb{R})$ be a Dirac-sequence and let $1 \le p < \infty$. For all $\phi \in L^p(0,T;X,Y,\mu)$ we have $\psi_\eta * \phi \to \phi$ in $L^p_\mu(0,T;X,Y)$ as $\eta \to 0$.*

Proof. If $(\psi_\eta)_{\eta>0} \subset C_c^\infty(\mathbb{R})$ is a Dirac-sequence, one can use the denseness of $C_c(0,T;X)$ in $L^p(0,T;X,Y,\mu)$ and the uniform convergence of $\psi_\eta * \phi \to \phi$ for $\phi \in C_c(0,T;X)$. □

Lemma A.11. *Let $\psi \in C_c^\infty(\mathbb{R})$ be non-negative, symmetric, with support in $(-1,1)$ and with total mass $\int_{\mathbb{R}}\psi(t)\,dt = 1$, and define the family of mollifiers $\psi_\eta(t) := \eta^{-1}\psi(t/\eta)$. For any $\mu \in \mathrm{rca}(0,T;X^*)$ and $\eta > 0$, the functions $\psi_\eta * \mu$ belong to $C^\infty(0,T;X^*)$ and $\frac{\mathrm{d}}{\mathrm{d}t}(\psi_\eta * \mu)(t) = (\dot{\psi}_\eta * \mu)(t)$. Moreover, the measures $\psi_\eta * \mu$ converge weakly-$*$ to μ as $\eta \to 0$ and the following estimate holds for all Borel-sets $I \subset (0,T)$:*

$$\int_I |\psi_\eta * \mu|(t)dt \le |\mu|\big(\cup_{t\in I}(t-\eta,t+\eta)\big)\,,$$

Proof. The proof follows the lines of [4, Theorem 2.2]. □

A.4 Topologies on Banach-valued measures

In this section we recall the most relevant topologies on the space $\mathrm{rca}(0,T;X^*)$, where X is a Banach space. Although $\mathrm{rca}(0,T;X^*)$ is a Banach space with norm $\|\cdot\|_{\mathrm{TV}}$, its norm topology is too strong for many practical purposes. Instead, we mostly work with the weak-* topology, which is sometimes called the *vague* topology. On the other hand, motivated by the duality between $C_b(0,T;X)$ and the space of finite, finitely additive regular signed Borel set functions (see [15, Th. IV.6.2] for the finite-dimensional version), one often also works with the topology induced by duality with $C_b(0,T;X)$, which is sometimes called the *narrow* or *weak* topology. To avoid confusion, we will avoid calling these topologies weak or weak-*, and stick to vague and narrow instead. To be more precise, we define:

Definition A.12. *Let $(\mu_n)_n$ be a net and μ an element in $\mathrm{rca}(0,T;X^*)$. We say that*

μ_n converges to μ in the vague *topology whenever:*

$$\mu_n \xrightarrow{\text{vague}} \mu \qquad :\Longleftrightarrow \qquad \langle\!\langle \phi,\mu_n\rangle\!\rangle \to \langle\!\langle \phi,\mu\rangle\!\rangle \qquad \text{for all } \phi \in C_0(0,T;X), \tag{A.5}$$

μ_n converges to μ in the narrow *topology whenever:*

$$\mu_n \xrightarrow{\text{narrow}} \mu \qquad :\Longleftrightarrow \qquad \langle\!\langle \phi,\mu_n\rangle\!\rangle \to \langle\!\langle \phi,\mu\rangle\!\rangle \qquad \text{for all } \phi \in C_b(0,T;X). \tag{A.6}$$

Moreover, we will say that a net $(f_n)_n \subset L^1(0,T;X^*)$ converges to an element $f \in L^1(0,T;X^*)$ in the vague or narrow topology whenever the measures $(f_n(t)\,\mathrm{d}t)_n$ converge to $f(t)\,\mathrm{d}t$ in the vague or narrow topology respectively.

The vague topology is not metrisable, since it is really a weak-* topology; in particular it cannot be fully characterised through its convergent sequences. It should be noted that vague convergence is often also defined as convergence against compactly supported test functions $\phi \in C_c(0,T;X)$, in which case it would be metrisable. For measures of uniformly bounded finite total variation (e.g. probabililty measures) the two notions coincide; this is not the case in the present work and we work with test functions in $C_0(0,T;X)$.

Clearly the narrow topology is stronger then the vague topology. As in the case of real-valued measures, vague convergence can be strengthened by a tightness argument.

Lemma A.13. *Let $(\mu_n)_n$ be a net and μ an element in* $\mathrm{rca}(0,T;X^*)$. *If $\mu_n \xrightarrow{\text{vague}} \mu$ and the sequence $(|\mu_n|)_n \subset \mathrm{rca}(0,T)$ is tight, then $\mu_n \xrightarrow{\text{narrow}} \mu$.*

Proof. Take an arbitrary test function $\phi \in C_b(0,T;X)$ and an arbitrary $\epsilon > 0$. By the tightness there exists a compact set $K_\epsilon \subset (0,T)$ for which $|\mu_n|(K_\epsilon^c) \le \epsilon$ for all n and without loss of generality we can assume that $|\mu|(K_\epsilon^c) < \epsilon$ since $|\mu|$ is regular. Take a test function $\psi \in C_0(0,T;X)$ such that $\phi|_{K_\epsilon} \equiv \psi|_{K_\epsilon}$. Then

$$\begin{aligned}\left|\langle\!\langle \phi, \mu_n - \mu\rangle\!\rangle\right| &\le \left|\langle\!\langle \psi, \mu_n - \mu\rangle\!\rangle\right| + \big(\|\psi\|_\infty + \|\phi\|_\infty\big)\big(|\mu_n|(K_\epsilon^c) + |\mu|(K_\epsilon^c)\big) \\ &< \left|\langle\!\langle \psi, \mu_n - \mu\rangle\!\rangle\right| + 2\epsilon\big(\|\psi\|_\infty + \|\phi\|_\infty\big) \to 2\epsilon\big(\|\psi\|_\infty + \|\phi\|_\infty\big),\end{aligned}$$

which proves the statement as ϵ was arbitrary. □

We can also strengthen vague convergence if one knows that the total variations converge. The proof requires sequences; the argument breaks down for general nets since a convergent net does not necessarily form a compact set.

Proposition A.14. *Let $(\mu_n)_n$ be a sequence and μ an element in* $\mathrm{rca}(0,T;X^*)$. *If $\mu_n \xrightarrow{\text{vague}} \mu$ and $\|\mu_n\|_{\mathrm{TV}} \to \|\mu\|_{\mathrm{TV}}$ then $\mu_n \xrightarrow{\text{narrow}} \mu$.*

Proof. First we show by standard approximation arguments that the $|\mu_n|$ converge narrowly, which implies the tightness of the variation measures. Then we exploit the tightness to strengthen the vague convergence to the narrow convergence.

Recall from Theorem A.7 that

$$\big\||\mu_n|\big\|_{\mathrm{TV}} = \|\mu_n\|_{\mathrm{TV}} \to \|\mu\|_{\mathrm{TV}} = \big\||\mu|\big\|_{\mathrm{TV}}. \tag{A.7}$$

Therefore the variation measures $|\mu_n|$ are bounded and a subsequence converges vaguely to some finite, positive measure $\nu \in \mathrm{rca}(0,T)$. Because of Theorem A.8, for any $\phi \in C_0(0,T;X)$ and along the convergent subsequence,

$$\langle\!\langle \phi, \mu\rangle\!\rangle \leftarrow \langle\!\langle \phi, \mu_n\rangle\!\rangle \le \int_0^T \|\phi(t)\|_X \,|\mu_n|(\mathrm{d}t) \to \int_0^T \|\phi(t)\|_X \,\nu(\mathrm{d}t).$$

Taking the supremum over test function yields the inequality,

$$\|\mu\|_{\mathrm{TV}} = \sup_{\phi \in C_0(0,T;X)} \langle\!\langle \phi, \mu\rangle\!\rangle \le \sup_{\phi \in C_0(0,T;X)} \int_0^T \|\phi(t)\|_X \,\nu(\mathrm{d}t) \le \sup_{\psi \in C_0(0,T)} \int_0^T \psi(t)\,\nu(\mathrm{d}t) = \|\nu\|_{\mathrm{TV}}.$$

However, the other direction is immediately from the vague lower semicontinuity of the total variation and so $\|\nu\|_{\mathrm{TV}} = \|\mu\|_{\mathrm{TV}}$. Together with (A.7) and $|\mu_n| \xrightarrow{\text{vague}} \nu$, this implies that $|\mu_n| \xrightarrow{\text{narrow}} \nu$. Hence by the (real-valued) Prohorov Theorem the measures $|\mu_n|$ are tight, and by Lemma A.13 the measures μ_n convergence narrowly to μ. □

Remark A.15. *The converse statement is, in general, wrong. To see this let X be a separable Hilbert space with basis $(e_n)_{n\in\mathbb{N}}$ and set $\mu_n = \mathrm{e}_n\,\mathcal{L}$. Then*

$$\forall f \in C_b(0,T;X) \qquad \int_0^T \langle f, d\mu_n\rangle = \langle \int_0^T f\,dt, \mathrm{e}_n\rangle \to 0$$

but $\|\mu_n\|_{\mathrm{TV}} = 1$ for every n.

References

[1] L. Ambrosio. Metric space valued functions of bounded variation. *Annali della Scuola Normale Superiore di Pisa - Classe di Scienze, 4e série*, 17(3):439–478, 1990.

[2] L. Ambrosio and S. Di Marino. Equivalent definitions of BV space and of total variation on metric measure spaces. *Journal of Functional Analysis*, 266(7):4150–4188, 2014.

[3] L. Ambrosio and R. Ghezzi. Sobolev and bounded variation functions on metric measure spaces (lecture notes). `http://cvgmt.sns.it/paper/2738/`, 2016.

[4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, NY, USA, 2006.

[5] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics. ETH Zürich. Birkhauser, Basel, Switzerland, 2nd edition, 2008.

[6] L. Ambrosio, R. Ghezzi, and V. Magnani. BV functions and sets of finite perimeter in sub-Riemannian manifolds. *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire*, 32 (3):489–517, 2015.

[7] L. Bertini, A. Faggionato, and D. Gabrielli. Large deviations of the empirical flow for continuous time markov chains. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques*, 51(3):867–900, 2015.

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

[9] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, volume 5 of *Math. Studies*. North-Holland, Amsterdam, 1973.

[10] D. Chiron. On the definitions of Sobolev and BV spaces into singular spaces and the trace problem. *Communications in Contemporary Mathematics*, 9(4):473, 2007.

[11] J. Conway. *A course in functional analysis*. Springer, New York, N.Y., USA, 2nd edition, 2007.

[12] S. Di Marino. Sobolev and BV spaces on metric measure spaces via derivations and integration by parts. `http://cvgmt.sns.it/paper/2521`, 2014.

[13] J. Diestel and J. J. J. Uhl. *Vector Measures: Mathematical Surveys 15*, volume 95. American Mathematical Society, 1967.

[14] N. Dinculeanu. *Vector Measures: International Series of Monographs in Pure and Applied Mathematics Vol. 95*, volume 95. Pergamon press / Deutscher Verlag der Wissenschaften, 1967.

[15] N. Dunford and J. Schwartz. *Linear operators, part one: general theory*. Interscience, New York, NY, USA, 1957.

[16] S. Ethier and T. Kurtz. *Markov processes – characterization and convergence*. John Wiley & sons, Inc., Hoboken, NJ, USA, 2005.

[17] L. Evans and R. Gariepy. *Measure theory and fine properties of functions*. CRC Press, Boca Raton, Fl, USA, 1992.

[18] A. Jakubowski. A non-Skorohod topology on the Skorohod space. *Electronic Journal of Probability*, 2(4):1–21, 1997.

[19] P. Krejcı. Regulated evolution quasivariational inequalities. *Notes to Lectures held at the University of Pavia*, 2003.

[20] P. Krejcí. The Kurzweil integral and hysteresis. In *Journal of Physics: Conference Series*, volume 55, page 144. IOP Publishing, 2006.

[21] P. Logaritsch and E. Spadaro. A representation formula for the p-energy of metric space-valued Sobolev maps. *Communications in Contemporary Mathematics*, 14(6):1250043, 2012.

[22] T.-W. Ma. *Banach-Hilbert Spaces, Vector Measures and Group Representations.* World Scientific, Singapore, 2002.

[23] A. Mainik and A. Mielke. Existence results for energetic models for rate-independent systems. *Calculus of Variations and Partial Differential Equations*, 22(1):73–99, 2005.

[24] P. Meyer and W. Zheng. Tightness criteria for laws of semimartingales. *Annales de l'I.H.P., section B*, 20(4):353–372, 1984.

[25] A. Mielke and T. Roubicek. *Rate-Independent Systems.* Springer, 2015.

[26] A. Mielke, F. Theil, and I. V. Levitas. A variational formulation of rate-independent phase transformations using an extremum principle. *Archive for Rational Mechanics and Analysis*, 162(2):137–177, 2012. ISSN 1432-0673. doi: 10.1007/s002050200194. URL `http://dx.doi.org/10.1007/s002050200194`.

[27] J. J. Moreau, P. D. Panagiotopoulos, and G. Strang. *Topics in nonsmooth mechanics.* Birkhäuser, 1988.

[28] I. A. Patterson and D. R. M. Renger. Dynamical large deviations of countable reaction networks under a weak reversibility condition. WIAS Preprint No. 2273, 2016.

[29] W. Rudin. *Functional Analysis.* Tata McGraw-Hill, New Delhi, India, tmh edition, 1974.

[30] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Annali di Matematica pura ed applicata*, 146(1):65–96, 1986.

[31] A. Wiweger. Linear spaces with mixed topology. *Studia Mathematica*, 20(1):47–68, 1961.