Lecture in Summer Term 2021: Functions of bounded variation and their applications

Franz Gmeineder

E-mail address: franz.gmeineder@uni-konstanz.de

Website: https://fxgmeineder.wixsite.com/meinewebsite

Version: May 18, 2021

Universität Konstanz

Abstract. The present lecture notes aim to provide a more or less thorough draft for my course *Functions of bounded variation and their applications* at the University of Konstanz in summer 2021. In this course, we give an introduction to a very broad class of functions that allows us to treat problems that require the incorporation of jumps. This is precisely the framework for *functions of bounded variation* – in brief, BV-functions.

As the course is designed as a weekly 2 hours lecture, the present notes not only comprise the material covered in class but also provide auxiliary background facts that were only mentioned in class. Whenever a statement was only mentioned but not proved in class, you directly find a short statement indicating this circumstance at the beginning of the proof. Besides, every chapter is concluded with literature references that should help you to get an idea of what further reading might be useful.

The general structure of the overall document is pretty much routine, but there are some things that are worth being pointed out:

- A list on *notation* can be found on page 1.
- *Definitions* are usually to be found within a gray (●) box.
- *Lemmas, propositions, corollaries and theorems* are usually to be found within a light blue (●) box.
- *Examples* are usually to be found within a light red (●) box.
- Paragraphs with an asterisk ($*$) have not been discussed in class – they either provide additional material or background facts which are useful for a better understanding of the main text.

As background material for my lecture, the present lecture notes are exclusively meant for attendees of my course. Certainly, these notes are not free from mathematical or grammatical errors. If you spot any such mistakes, or have suggestions on how the course and/or the lecture notes could be improved, I would be very grateful to be approached via e-mail:

`franz.gmeineder@uni-konstanz.de`

A final comment: If you feel that you lack some background knowledge that would be necessary to follow the ideas explained in class – contact me, and besides a quick meeting I will add a quick section on the topic to the notes.

Contents

The word cloud on the title page was produced using `https://www.wortwolken.com/`.

Blackbox on notation

Structures

- $|\cdot|$: Euclidean norm on $\mathbb{R}^n$
- $\langle\cdot,\cdot\rangle$: Euclidean inner product on $\mathbb{R}^n$
- $\mathrm{B}_r(x_0)$: Open ball (with respect to the euclidean norm) of radius $r > 0$ centered at $x_0 \in \mathbb{R}^n$

Measure and integration theory

- $\mathscr{L}^n$: n-dimensional Lebesgue measure
- $\mathscr{H}^{n-1}$: $(n-1)$-dimensional Hausdorff measure (practically: $(n-1)$-dimensional surface measure)
- $\mathscr{B}(\Omega)$: Borel σ-algebra
- $\mathrm{RM}(\Omega;\mathbb{R}^m)$: $\mathbb{R}^m$-valued Radon measures on Ω
- $\mathrm{RM}_{\mathrm{fin}}(\Omega;\mathbb{R}^m)$: Finite $\mathbb{R}^m$-valued Radon measures on Ω
- $\mu \llcorner A$: Restriction of μ to the set A

Function spaces

- $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$: Space of locally integrable functions $u\colon \Omega \to \mathbb{R}$ (for $\mathscr{L}^n$)
- $\mathrm{L}^p(\Omega)$: Space of p-integrable functions $u\colon \Omega \to \mathbb{R}$ (for $\mathscr{L}^n$)
- $\mathrm{W}^{1,1}_{\mathrm{loc}}(\Omega)$: Space of weakly differentiable functions $u\colon \Omega \to \mathbb{R}$
- $\mathrm{W}^{k,p}(\Omega)$: Space of k-times weakly differentiable functions $u\colon \Omega \to \mathbb{R}$ such that both u and all weak derivatives up to order k belong to $\mathrm{L}^p(\Omega)$.
- $\mathrm{C}^k_b(\Omega)$: Space of k-times continuously differentiable functions $u\colon \Omega \to \mathbb{R}$ such that u along with all partial derivatives up to order k are bounded.
- $\mathrm{C}^\infty_c(\Omega)$: Smooth functions with compact support in Ω (also termed *test functions*).

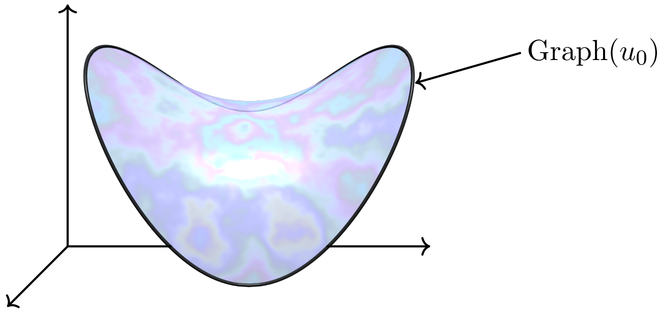

Fig. 1: The situation of Section 1.1.

1 From minimal surfaces to imaging

In this introductory section we aim to describe some variational problems that will eventually give us good motivation to study the core topic the course – the functions of *bounded variation*. In doing so, we devote ourselves to two major problems: The minimal surface problem and topics from the denoising of images.

1.1 Minimal surface-type problems

We begin with a simple variant of the (non-parametric) minimal surface problem. For a certain wire in $\mathbb{R}^3$, we aim to find a surface attached to this wire, yet having minimal surface area.

We simplify this setting even further and suppose that $\Omega \subset \mathbb{R}^2$ is an open set, $u_0 \colon \partial\Omega \to \mathbb{R}$ a function and

$$\mathrm{Graph}(u_0) := \{(x, u_0(x)) \colon\; x \in \partial\Omega\} \subset \mathbb{R}^3.$$

We take $\mathrm{Graph}(u_0)$ as a description of the aforementioned wire and only seek surfaces that can be written as graphs of functions $u \colon \Omega \to \mathbb{R}$. For sufficiently nice such functions, the surface area can be written as

$$\mathrm{Area}[u] := \int_\Omega \sqrt{1 + |\nabla u|^2}\,\mathrm{d}x.$$

Note that this functional can be established by elementary geometry and, in essence, only uses the theorem of Pythagoras. Our specific problem thus reads

$$\text{minimise} \quad \mathscr{F}[u] := \int_\Omega \sqrt{1 + |\nabla u|^2}\,\mathrm{d}x \quad \text{subject to} \;\; u|_{\partial\Omega} = u_0. \tag{1.1}$$

The condition $u|_{\partial\Omega} = u_0$ precisely asserts that the surface has to be attached to the wire, the latter being modelled as the graph of u_0.

For (1.1) to be a sensible problem, we have to specify a class of functions $\mathcal{D}$ on which we consider the minimisation problem; merely prescribing boundary

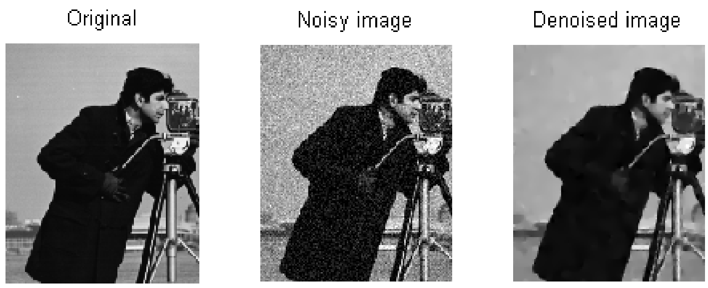

Fig. 2: On the denoising of images.

values in fact is not sufficient. A first option to come up with such a candidate for $\mathcal{D}$ are subclasses of the Sobolev space $\mathrm{W}^{1,1}(\Omega)$ (to be recalled in detail in Section 2).

However, as shall be discussed in Section 3, this space does not share good compactness properties. As we will see, this is basically due to the fact that sequences bounded in the L^1-norm might concentrate. Towards a fruitful existence theory, we thus must consider a slightly larger space – precisely, this space is given by the *functions of bounded variation* – the core topic of the present lecture. As a metaprinciple, we

will gain compactness, but enlarge our spaces substantially.

By *enlarging* we understand that functions of bounded variation comprise all $\mathrm{W}^{1,1}$-functions, but might exhibit a variety of other discontinuities. As such, the chief objectives are to

- rigorously introduce functions of bounded variation,
- understand in which sense they admit 'more' singularities than $\mathrm{W}^{1,1}$-functions, and
- how they can be applied to lead to a satisfactory existence theory for variational problems such as (1.1).

1.2 Denoising of images

Another area where BV-functions will eventually a pivotal rôle is (mathematical) *imaging*. This is an extremely rich and vast field, and we shall only be able to scratch some of the problems arising within this context.

Our first model problem is this: A picture (that we shall refer to as *original picture* **O**) was subject to some noise. This noise is here assumed to be Gaußian. We are given the noisy picture **N** and aim to

produce a **denoised** variant **D** of the noisy picture **N**.

Such a scenario is depicted in Figure 2.

Which principles should be obeyed by such a denosing process? This was precisely what Rudin, Osher & Fatemi discussed in their seminal 1992 paper [4], and the pivotal ideas are roughly as follows:

(P1) On a heuristic level, *noise* generally has a lot of *variation*. As such, if we wish to *denoise* an image, we aim to *reduce the variation* of the image.

(P2) On the other hand, the denoised image should still *remain close* to the noisy picture. This point is pretty intuitive: We want to come up with a denoised image that is not fully decoupled from our input.

(P3) Finally, a key point in images are *edges* – edges help us to distinguish between the single objects appearing on screen.

In view of this objective, we first have to make a mathematical model of the situation at our disposal. To do so, we proceed as follows: The rectangle (over which we see the picture) is modelled as a subset Ω of $\mathbb{R}^2$. We then model a *picture* or an *image* as a function $v\colon \Omega \to [0,1]$, where each number $t \in [0,1]$ represents a certain value in a gray or black/white scale. So, for instance, $t = 0$ might refer to white, whereas $t = 1$ might refer to black.

Now, the noisy picture that we received is modelled by a function $f\colon \Omega \to [0,1]$. In order to produce a *denoised picture* $u\colon \Omega \to [0,1]$ according to our three principles (P1),(P2) and (P3) from above, we start from the idea that the variation of an image is linked to the gradient of the modelling function. For a function $F\colon \mathbb{R}^n \to \mathbb{R}$ to be discussed below, we wish to minimise

$$\mathfrak{F}[u] := \underbrace{\int_\Omega F(\nabla u)\,\mathrm{d}x}_{=:\mathrm{I}} + \frac{\lambda}{2}\underbrace{\int_\Omega |u-f|^2\,\mathrm{d}x}_{=:\mathrm{II}} \tag{1.2}$$

over a suitable class of functions $u\colon \Omega \to [0,1]$. If we minimise $\mathfrak{F}$, then we will also minimise part I, and this will reduce the overall gradient – hence the variation – of u. Part II reflects (P2) in the sense that it forces u to stay close to our denoised input image f. The parameter $\lambda \geq 0$ is a free modelling parameter that influences *how close* to f the minimiser will be. Think of the extreme case $\lambda = 0$: In this case, the functional $\mathfrak{F}$ does not refer to f in any way. Hence, as a metaprinciple, the larger λ is, the better our minimiser will approximate f.

Until now, we have not incorporated principle (P3) into our discussion. Principle (P3) refers to preserving important structural properties such as edges. It is at this stage where the integrand F enters crucially.

Our subsequent discussion will make use of Sobolev spaces and test functions, and the reader is advised to see Section 2 for a recap on the underlying concepts.

A functional that might be known from an introductory course on partial differential equations or functional analysis is the *Dirichlet integral*

$$\mathrm{Dir}[u] := \frac{1}{2}\int_\Omega |\nabla u|^2\,\mathrm{d}x,$$

to be considered on a subclass of the Sobolev space $\mathrm{W}^{1,2}(\Omega)$. As such, a first choice of an integrand $F\colon \mathbb{R}^n \to \mathbb{R}$ in (1.2) would be

$$F(z) := \frac{1}{2}|z|^2, \qquad z \in \mathbb{R}^n.$$

In the following, we briefly argue *why this is not a good choice* in view of (P3). The argument uses some elliptic regularity theory; if you are not familiar with the underlying concepts, you may easily drop the following and directly jump to (1.7).

Let $f \in \mathrm{L}^2(\Omega)$ (which is certainly fulfilled for any measurable $f\colon \Omega \to [0,1]$), $\lambda > 0$ and suppose that $u \in \mathrm{W}^{1,2}(\Omega)$ is a minimiser for $\mathfrak{F}$ in the sense that

$$\mathfrak{F}[u] \le \mathfrak{F}[u+\varepsilon\varphi] \quad \text{for all } \varepsilon > 0,\ \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{1.3}$$

Considering $(-\varphi)$ instead of φ, (1.3) thus implies that the function

$$\Phi\colon \mathbb{R} \ni \varepsilon \mapsto \mathfrak{F}[u+\varepsilon\varphi] \in \mathbb{R}$$

has a minimum in $\varepsilon = 0$ for any fixed $\varphi \in \mathrm{C}_c^\infty(\Omega)$. We then record

Lemma 1.1. Let $\Omega \subset \mathbb{R}^2$ be open, $f \in \mathrm{L}^2(\Omega)$ and suppose that $u \in \mathrm{W}^{1,2}(\Omega)$ satisfies (1.3). Then there holds

$$\int_\Omega \langle \nabla u, \nabla\varphi\rangle \,\mathrm{d}x + \lambda \int_\Omega (u-f)\varphi \,\mathrm{d}x = 0 \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{1.4}$$

Proof. Let $\varphi \in \mathrm{C}_c^\infty(\Omega)$ be arbitrary but fixed. Expanding, we find

$$\begin{aligned} \Phi(\varepsilon) &= \mathfrak{F}[u+\varepsilon\varphi] \\ &= \frac{1}{2}\int_\Omega |\nabla u|^2 + 2\varepsilon\langle \nabla u, \nabla\varphi\rangle + \varepsilon^2|\nabla\varphi|^2 \,\mathrm{d}x \\ &\qquad\qquad + \frac{\lambda}{2}\int_\Omega |u-f|^2 + 2\varepsilon(u-f)\varphi + \varepsilon^2|\varphi|^2 \,\mathrm{d}x. \end{aligned} \tag{1.5}$$

This is a polynomial of degree two in ε. Thus, a necessary condition for Φ to have a minimum in $\varepsilon = 0$ is

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Big|_{\varepsilon=0} \Phi(\varepsilon) = 0. \tag{1.6}$$

Based on (1.5), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\Phi(\varepsilon) = \int_\Omega \langle \nabla u, \nabla\varphi\rangle + \varepsilon|\nabla\varphi|^2 \,\mathrm{d}x + \frac{\lambda}{2}\int_\Omega 2(u-f)\varphi + 2\varepsilon|\varphi|^2 \,\mathrm{d}x,$$

and (1.6) then implies that

$$\int_\Omega \langle \nabla u, \nabla\varphi\rangle \,\mathrm{d}x + \lambda\int_\Omega (u-f)\varphi \,\mathrm{d}x = 0.$$

This is (1.4), and the proof is complete. □

Integrating by parts, eqation (1.4) can be viewed as a *weak version* of the partial differential equation

$$-\Delta u + \lambda(u - f) = 0 \qquad \text{in } \Omega, \tag{1.7}$$

where we refer the reader to Section 2.3 for a quick discussion of this matter. Assuming $\lambda > 0$, (1.7) can be rewritten as

$$\mathscr{L}u := -\Delta u + \lambda u \stackrel{!}{=} \lambda f \qquad \text{in } \Omega. \tag{1.8}$$

The differential operator $\mathscr{L}$ is an *elliptic* operator of degree 2. As a metaprinciple, the regularity of (weak) solutions v of equations

$$\mathscr{L}v = g$$

will be two regularity degrees higher than that of g. E.g., if $g \in \mathrm{L}^2_{\mathrm{loc}}$, then $v \in \mathrm{W}^{2,2}_{\mathrm{loc}}$, or, more generally,

$$g \in \mathrm{W}^{k,2}_{\mathrm{loc}} \Longrightarrow v \in \mathrm{W}^{k+2,2}_{\mathrm{loc}}.$$

Informally, this amounts to saying that v will be **much smoother** than g.

This now has the following impact on our model (1.2): If we choose $F(z) = \frac{1}{2}|z|^2$ in (1.2), minimisers will satisfy (1.4). Even if $f \in \mathrm{L}^2(\Omega)$ (so that the picture modelled by f might in fact reveal sharp edges), u obtained by the minimisation of $\mathfrak{F}$ is much smoother than f. In particular, edges might be smeared out.

Obviously, the behaviour of F has a crucial impact on whether we have such a smoothing – we have just seen that $F(z) := \frac{1}{2}|z|^2$ leads to an undesired smoothing of edges. This motivates the following question:

In view of (P3), what is a good choice for F in (1.2)?

Throughout the course, we shall see that a *good* choice for F is a generalisation of $F(z) := |z|$. The problem, however, is slightly more subtle: As we will see in Section 2, Sobolev functions are **never allowed to jump**. In consequence, we will come up with a variant of $F(z) := |z|$ and a function space such that its elements are allowed to have *jump discontinuities*. Incorporating jump discontinuities for the functions that model our pictures, we may indeed hope for preserving edges.

As one may anticipate, this function space is precisely given by the *functions of bounded variation* – the core topic of the course.

2 Recap on distributions and Sobolev spaces

As discussed in the introduction, one of the main objectives of the course is to come up with a framework that lets us deal with **jump functions**. Such functions are discontinuous and hence cannot be differentiated classically. In consequence, if we wish to speak of a gradient of potentially very irregular functions, we must come up with a very robust and general notion of differentiability.

This is usually accomplished within the framework of *distributions*, and this section aims to give a very quick overview of the underlying concepts. Distibutions are – in a sense to be specified below – the continuous dual of the *test functions*, and the differentation of distributions is inspired by the integration by parts-formula for smooth functions.

Once the notion of distributional derivatives is at our disposal, we proceed as follows: Associating with each $f \in \mathrm{L}^1_{\mathrm{loc}}$ a distribution T_f, we may consider the distributional (partial) derivatives of T_f; in this way, we will obtain a notion of (partial) derivatives for all $\mathrm{L}^1_{\mathrm{loc}}$-functions. This will be accomplished in Section 2.1.

We then conclude this chapter in Section 2.2 with a discussion of Sobolev spaces. Such spaces arise frequently in the study of partial differential equations and have already been alluded to in the introduction; finally, some general ideas on weak formulations of partial differential equations are given in Section 2.3.

2.1 Distributions

Our chief objective of the present chapter is to come up with a generalisation of the usual concept of partial derivatives. To do so, we first draw some conclusions from the integration by parts-formula; this will serve as the key motivation for the differentiation of distributions later on.

To set the stage, we require some notation: For an open set $\Omega \subset \mathbb{R}^n$ and a continuous function $\varphi\colon \Omega \to \mathbb{R}$, we define its **support** by

$$\mathrm{spt}(\varphi) := \overline{\{x \in \Omega\colon\ \varphi(x) \neq 0\}},$$

i.e., as the closure of the set where φ does not equal zero. Based on the notion of support, we introduce the space of **test functions** by

$$\mathrm{C}^\infty_c(\Omega) := \left\{ \varphi\colon \Omega \to \mathbb{R}\colon \quad \begin{matrix} \varphi \in \mathrm{C}^k(\Omega) \text{ for all } k \in \mathbb{N}, \\ \mathrm{spt}(\varphi) \text{ is compactly contained in } \Omega \end{matrix} \right\}. \tag{2.1}$$

Here, as usual, we write $\varphi \in \mathrm{C}^k(\Omega)$ provided $\varphi\colon \Omega \to \mathbb{R}$ is k-times continuously differentiable.

2.1.1 Integration by parts and some consequences

From an introductory course on analysis and measure theory we recall that, for an open and bounded set $\Omega \subset \mathbb{R}^n$ with smooth boundary $\partial\Omega$, we have the

Gauß-Green formula for vector fields $v \in \mathrm{C}(\overline{\Omega};\mathbb{R}^n) \cap \mathrm{C}^1(\Omega;\mathbb{R}^n)$

$$\int_\Omega \operatorname{div}(v)\,\mathrm{d}x = \int_{\partial\Omega} \langle v,\nu\rangle\,\mathrm{d}\mathscr{H}^{n-1}, \tag{2.2}$$

where $\nu = (\nu_1,...,\nu_n)\colon \partial\Omega \to \{x \in \mathbb{R}^n\colon\ |x| = 1\}$ denotes the outward unit normal to $\partial\Omega$.

Let $u \in \mathrm{C}^1(\Omega)$ and $\varphi \in \mathrm{C}_c^\infty(\Omega)$. For $i \in \{1,...,n\}$, we consider the vector function $v = (0,...,0,\varphi u,0,...,0)$ (with φu at the i-th entry). Applying (2.2) to this choice of v, we obtain

$$\int_\Omega (\partial_{x_i}u)\varphi\,\mathrm{d}x = \int_{\partial\Omega} u\varphi\nu_i\,\mathrm{d}\mathscr{H}^{n-1} - \int_\Omega u\partial_{x_i}\varphi\,\mathrm{d}x = -\int_\Omega u\partial_{x_i}\varphi\,\mathrm{d}x. \tag{2.3}$$

Here, the boundary integral vanishes because φ has compact support in Ω and thus equals zero on $\partial\Omega$.

Even though this calculation requires u to be of class C^1, the **right hand side does not involve derivatives of u.** It is precisely this insight that will serve as a definition for distributional derivatives (see Definition 2.8 below). More precisely, for $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$, put

$$T_f(\varphi) := \int_\Omega f\varphi\,\mathrm{d}x, \qquad \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{2.4}$$

Adopting this notation, (2.3) becomes

$$T_{\partial_{x_i}u}(\varphi) = -T_u(\partial_{x_i}\varphi) \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{2.5}$$

Based on the previous identity, it is natural to *define* $\partial_{x_i}T_u$ as $(-T_u(\partial_{x_i}\varphi))$. This definition will even make sense for all $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$.

2.1.2 General theory

It is useful to embed the foregoing considerations into a broader context; this is achieved by realising that T_f as in (2.4) not only is a linear map $T_f\colon \mathrm{C}_c^\infty(\Omega) \to \mathbb{R}$ but also *continuous*. Here, continuity refers to a specific notion of convergence on $\mathrm{C}_c^\infty(\Omega)$ defined as follows:

Definition 2.1 (Convergence in $\mathscr{D}(\Omega)$)**.** Let $\varphi,\varphi_1,\varphi_2,... \in \mathrm{C}_c^\infty(\Omega)$. We say that (φ_j) **converges to φ in $\mathscr{D}(\Omega)$** if and only if

(a) $\operatorname{spt}(\varphi),\operatorname{spt}(\varphi_1),\operatorname{spt}(\varphi_2),... \subset K$ for some compact subset K of Ω, and

(b) for all $\alpha \in \mathbb{N}_0^n$ there holds $\|\partial^\alpha(\varphi-\varphi_j)\|_{\mathrm{L}^\infty(\Omega)} \to 0$ as $j \to \infty$.

It is important to note that (a) does not follow from (b) and vice versa. In particular, condition (a) ensures that the supports of $\varphi_1,\varphi_2,...$ do not approach the boundary of Ω. Moreover, as $\varphi,\varphi_1,...$ are smooth, (b) precisely means

$$\forall\alpha \in \mathbb{N}_0^n\colon\ \sup_{x\in\Omega}|\partial^\alpha(\varphi-\varphi_j)(x)| \to 0, \qquad j \to \infty,$$

and this is to say that all partial derivatives converge uniformly.

Based on the previous definition, we may now introduce the *distributions* as continuous linear maps from $\mathrm{C}_c^\infty(\Omega)$ to $\mathbb{R}$:

Definition 2.2 (Distributions)**.** Let $\Omega \subset \mathbb{R}^n$ be open. A **distribution on** Ω is a linear functional $T\colon \mathrm{C}_c^\infty(\Omega) \to \mathbb{R}$ that is continuous with respect to the convergence in $\mathscr{D}(\Omega)$, so satisfies

$$u_j \to u \quad \text{in } \mathscr{D}(\Omega) \Longrightarrow T(u_j) \to T(u) \quad \text{in } \mathbb{R}.$$

We denote $\mathscr{D}'(\Omega)$ the (linear) space of distributions on Ω.

We now return to T_f as given by (2.4) and discuss how it fits into the framework of distributions:

Example 2.3 (Regular distributions)**.** Let $\Omega \subset \mathbb{R}^n$ be open and let $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. We define a linear functional on $\mathrm{C}_c^\infty(\Omega)$ by

$$T_f(\varphi) := \int_\Omega f\varphi \,\mathrm{d}x, \qquad \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{2.6}$$

Then T is a distribution. Indeed, let $\varphi, \varphi_1, \varphi_2, \ldots \in \mathrm{C}_c^\infty(\Omega)$ be such that $\varphi_j \to \varphi$ in $\mathscr{D}(\Omega)$. By definition, there exists a compact subset $K \subset \Omega$ such that $\mathrm{spt}(\varphi), \mathrm{spt}(\varphi_1), \mathrm{spt}(\varphi_2), \ldots \subset K$. Since $f \in \mathrm{L}^1_{\mathrm{loc}}(K)$, $f|_K \in \mathrm{L}^1(K)$. Now,

$$|T_f(\varphi - \varphi_j)| \leq \int_K |f|\,|\varphi - \varphi_j| \,\mathrm{d}x \leq \|f\|_{\mathrm{L}^1(K)} \|\varphi - \varphi_j\|_{\mathrm{L}^\infty(\Omega)} \to 0,\ j \to \infty.$$

If $T \in \mathscr{D}'(\Omega)$ is such that there exists $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ with $T = T_f$, then we call T a **regular distribution**.

If T is a regular distribution with $T = T_f$, f is often referred to as *the* **regular representative**. As we shall see in Corollary 2.5, this terminology makes sense indeed. As an important preparation, we require

Lemma 2.4 (Du Bois-Reymond)**.** Let $U \subset \mathbb{R}^n$ be open and suppose that $f \in \mathrm{L}^1_{\mathrm{loc}}(U)$ is such that

$$\int_U f \cdot \varphi \,\mathrm{d}x = 0 \qquad \text{holds for all } \varphi \in \mathrm{C}_c^\infty(U). \tag{2.7}$$

Then there holds $f = 0$ $\mathscr{L}^n$-a.e..

Proof. The proof has not been given in the lecture. We let $\rho \in \mathrm{C}_c^\infty(\mathbb{R}^n)$ be a non-negative, radial function that is supported in $\mathrm{B}_1(0)$ and has integral one. For $\varepsilon > 0$, we define

$$\rho_\varepsilon(x) := \frac{1}{\varepsilon^n}\rho\Big(\frac{x}{\varepsilon}\Big), \qquad x \in \mathbb{R}^n.$$

As known from a course on integration theory, for any $g \in \mathrm{L}^1(\mathbb{R}^n)$ the mollification

$$\rho_\varepsilon * g(x) := \int_{\mathbb{R}^n} \rho_\varepsilon(x-y)g(y)\,\mathrm{d}y, \qquad x \in \mathbb{R}^n,$$

satisfies $\|g - \rho_\varepsilon * g\|_{\mathrm{L}^1(\mathbb{R}^n)} \to 0$ as $\varepsilon \searrow 0$. Moreover, we have $\rho_\varepsilon * g \in \mathrm{C}^\infty(\mathbb{R}^n)$ and, if g vanishes outside a bounded set V, then $\rho_\varepsilon * g$ satisfies

$$\mathrm{spt}(\rho_\varepsilon * g) \subset V + \mathrm{B}_\varepsilon(0) := \{x + y\colon\ x \in V,\ y \in \mathrm{B}_\varepsilon(0)\}.$$

After these preparations, we embark on the actual proof. For $f \in \mathrm{L}^1_{\mathrm{loc}}(U)$ as in the lemma, we define $\mathbf{F}^{\gtrless} := \{x \in U\colon\ f(x) \gtrless 0\}$ and

$$\mathbf{F}_j^{\gtrless} := \{x \in \mathbf{F}^{\gtrless}\colon\ |x| < j,\ \mathrm{dist}(x,\partial\Omega) > \tfrac{1}{j}\}, \qquad j \in \mathbb{N}. \tag{2.8}$$

Each $\mathbf{F}_j^{\gtrless}$ is bounded, has a distance at least $\frac{1}{j}$ from the boundary $\partial\Omega$ and we have $\mathbf{F}^{\gtrless} = \bigcup_{j\in\mathbb{N}} \mathbf{F}_j^{\gtrless}$. Since $\mathbb{1}_{\mathbf{F}_j^{\gtrless}} \in \mathrm{L}^1(\mathbb{R}^n)$, $\mathbb{1}_{\mathbf{F}_j^{\gtrless}} - \rho_\varepsilon * \mathbb{1}_{\mathbf{F}_j^{\gtrless}} \to 0$ in $\mathrm{L}^1(\mathbb{R}^n)$ as $\varepsilon \searrow 0$, and so we find a sequence (ε_k) with $\varepsilon_k \searrow 0$ such that $\mathbb{1}_{\mathbf{F}_j^{\gtrless}} - \rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{\gtrless}} \to 0$ $\mathscr{L}^n$-a.e. in $\mathbb{R}^n$. Now, since $|\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{\gtrless}}| \leq 1$ everywhere and $g \in \mathrm{L}^1_{\mathrm{loc}}(U)$, we obtain by Lebesgue's theorem on dominated convergence:

$$\int_{\mathbb{R}^n} f\mathbb{1}_{\mathbf{F}_j^{\gtrless}} - f(\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{\gtrless}})\,\mathrm{d}x \to 0, \qquad k \to \infty. \tag{2.9}$$

Note carefully that, by (2.8), $\mathrm{spt}(\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{\gtrless}}) \subset U$ provided $0 < \varepsilon_k < \frac{1}{j}$. Therefore $\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{\gtrless}} \in \mathrm{C}_c^\infty(U)$, so that by (2.7),

$$\int_{\mathbb{R}^n} f(\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{\gtrless}})\,\mathrm{d}x = 0 \qquad \text{for all } k \text{ with } 0 < \varepsilon_k < \tfrac{1}{j}. \tag{2.10}$$

We thus obtain for any fixed $j \in \mathbb{N}$ and $k \in \mathbb{N}$ with $0 < \varepsilon_k < \frac{1}{j}$

$$\begin{aligned}\int_{\mathbf{F}_j^{\gtrless}} |f|\,\mathrm{d}x &= \int_{\mathbb{R}^n} |f|\mathbb{1}_{\mathbf{F}_j^{\gtrless}}\,\mathrm{d}x = \int_{\mathbb{R}^n} f\mathbb{1}_{\mathbf{F}_j^{>}}\,\mathrm{d}x - \int_{\mathbb{R}^n} f\mathbb{1}_{\mathbf{F}_j^{<}}\,\mathrm{d}x \\ &\overset{(2.10)}{=} \int_{\mathbb{R}^n} f\mathbb{1}_{\mathbf{F}_j^{>}} - f\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{>}}\,\mathrm{d}x - \int_{\mathbb{R}^n} f\mathbb{1}_{\mathbf{F}_j^{<}} - f\rho_{\varepsilon_k} * \mathbb{1}_{\mathbf{F}_j^{<}}\,\mathrm{d}x \\ &\overset{(2.9)}{\to} 0, \qquad k \to \infty.\end{aligned}$$

Thus, $\mathscr{L}^n(\mathbf{F}_j^{\gtrless}) = 0$ for all $j \in \mathbb{N}$. Since $\mathbf{F}_j^{\gtrless} = \bigcup_{j\in\mathbb{N}} \mathbf{F}_j^{\gtrless}$, $\mathscr{L}^n(\mathbf{F}^{\gtrless}) = 0$. In consequence, $f = 0$ $\mathscr{L}^n$-a.e., and the proof is complete. $\square$

We may now deal with the aforementioned uniqueness of the regular representative:

Corollary 2.5. Let $\Omega \subset \mathbb{R}^n$ be open and $T \in \mathscr{D}'(\Omega)$ be a regular distribution. If $T = T_f$ for some $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$, then f is uniquely determined $\mathscr{L}^n$-a.e..

Proof. Suppose that $T = T_f = T_g$ for some $f, g \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$. The equality $T_f = T_g$ can be restated as

$$\int_\Omega (f - g)\varphi \,\mathrm{d}x = 0 \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega),$$

and from here the claim the follows by Lemma 2.4. □

So far, we have only encountered regular distributions. As can be seen from the next example, **not every distribution is regular**:

Example 2.6 (Non-regular distributions)**.** Let $\Omega \subset \mathbb{R}^n$ be open and $x_0 \in \mathbb{R}^n$. We define a distribution on Ω by

$$T(\varphi) := \varphi(x_0), \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega).$$

This is a distribution indeed. Linearity of T can be checked directly; moreover, if $\varphi, \varphi_1, \varphi_2, ... \in \mathrm{C}_c^\infty(\Omega)$ are such that $\varphi_j \to \varphi$ in $\mathscr{D}(\Omega)$, we obtain

$$|T(\varphi) - T(\varphi_j)| = |\varphi(x_0) - \varphi_j(x_0)| \leq \sup_{x \in \Omega} |\varphi(x) - \varphi_j(x)| \to 0, \qquad j \to \infty.$$

More importantly, we claim that T is **not a regular** distribution. To see this, we suppose towards a contradiction that there exists $f \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ such that $T = T_f$, i.e.,

$$\varphi(x_0) = \int_\Omega f\varphi \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{2.11}$$

We put $U := \Omega \setminus \{x_0\}$ (which is open again). Whenever $\varphi \in \mathrm{C}_c^\infty(U)$, we have $\varphi(x_0) = 0$ since $x_0 \notin \mathrm{spt}(\varphi)$. Thus, (2.11) yields

$$0 = \varphi(x_0) = \int_U f\varphi \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(U).$$

We then deduce from Lemma 2.4 that $f = 0$ $\mathscr{L}^n$-a.e. on U, and since $\{x_0\}$ is a Lebesgue nullset, $f = 0$ $\mathscr{L}^n$-a.e. on Ω. In consequence, we obtain from (2.11) that $\varphi(x_0) = 0$ for all $\varphi \in \mathrm{C}_c^\infty(\Omega)$ – which is the desired contradiction. Hence, T is not a regular distribution.

It is interesting to observe that the distribution T still can be written as

$$T(\varphi) = \int_\Omega \varphi \,\mathrm{d}\delta_{x_0}, \qquad \varphi \in \mathrm{C}_c^\infty(\Omega), \tag{2.12}$$

where δ_{x_0} is the **Dirac measure** centered at x_0. For $A \subset \Omega$, this measure is defined by

$$\delta_{x_0}(A) := \begin{cases} 1 & \text{if } x_0 \in A, \\ 0 & \text{otherwise.} \end{cases}$$

To see (2.12), we use that for any measure space $(\Omega, \mathcal{A}, \mu)$ and any non-negative, measurable function $g\colon \Omega \to [0,\infty)$ we have

$$\int_\Omega g \,\mathrm{d}\mu = \int_0^\infty \mu(\{x \in \Omega\colon\; g(x) > t\}) \,\mathrm{d}t.$$

For a non-negative $\varphi \in \mathrm{C}_c^\infty(\Omega)$, we thus obtain

$$\int_\Omega \varphi \,\mathrm{d}\delta_{x_0} = \int_0^\infty \delta_{x_0}(\{x \in \Omega\colon\; \varphi(x) > t\}) \,\mathrm{d}t = \int_0^{\varphi(x_0)} \mathrm{d}t = \varphi(x_0).$$

This is (2.12) for non-negative $\varphi \in \mathrm{C}_c^\infty(\Omega)$, and the general case follows by splitting φ into its non-negative and negative parts.

Whilst δ_{x_0} is not Lebesgue measure, it is still has a very simple structure. As a far-reaching generalisation, we will now extend our theory to a very broad class of distributions that arise from certain measures.

To begin our discussion, we strengthen terminology first. Let $\Omega \subset \mathbb{R}^n$ and denote $\mathscr{B}(\Omega)$ the Borel σ-algebra[1] on Ω. A measure $\mu\colon \mathscr{B}(\Omega) \to [0,\infty]$ is called a **Radon measure** if it respects the topological structure in the following sense:

- *Local finiteness:* We have $\mu(K) < \infty$ for any *compact* set $K \subset \Omega$.
- *Inner regularity:* For any $A \in \mathscr{B}(\Omega)$ we have

$$\mu(A) = \sup\{\mu(K)\colon\; K \subset A \text{ compact}\}.$$

Similarly as $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$-functions induce regular distributions, Radon measures induce the so-called *measure regular distributions*:

[1] This is the σ-algebra generated by the open subsets of Ω, i.e., the smallest σ-algebra containing all open subsets of Ω. Here, *openness* refers to the usual euclidean topology on $\mathbb{R}^n$.

Example 2.7 (Measure regular distributions)**.** Let $\Omega \subset \mathbb{R}^n$ be open and let $\mu\colon \mathscr{B}(\Omega) \to [0,\infty]$ be a Radon measure. We define a linear functional on $\mathrm{C}_c^\infty(\Omega)$ by

$$T_\mu(\varphi) := \int_\Omega \varphi \,\mathrm{d}\mu, \qquad \varphi \in \mathrm{C}_c^\infty(\Omega) \tag{2.13}$$

and claim that this functional belongs to $\mathscr{D}'(\Omega)$. To this end, let $\varphi, \varphi_1, \varphi_2 \in \mathrm{C}_c^\infty(\Omega)$ be such that $\varphi_j \to \varphi$ in $\mathscr{D}(\Omega)$ as $j \to \infty$. Then, in particular, there exists a compact set $K \subset \Omega$ such that $\mathrm{spt}(\varphi), \mathrm{spt}(\varphi_1), \mathrm{spt}(\varphi_2), ... \subset K$ and $\sup_{x\in\Omega} |(\varphi - \varphi_j)(x)| \to \infty$ as $j \to \infty$. Since μ is a Radon measure, it is locally finite and so $\mu(K) < \infty$. Therefore,

$$\begin{aligned} |T_\mu(\varphi) - T_\mu(\varphi_j)| &= \left| \int_\Omega (\varphi - \varphi_j)\,\mathrm{d}\mu \right| \\ &= \left| \int_K (\varphi - \varphi_j)\,\mathrm{d}\mu \right| \leq (\sup_\Omega |\varphi - \varphi_j|)\mu(K) \to 0. \end{aligned}$$

As T_μ is linear, we therefore obtain $T_\mu \in \mathscr{D}(\Omega)$. If $T \in \mathscr{D}'(\Omega)$ is such that $T = T_\mu$ for some Radon measure μ on Ω, we call T a **measure regular distribution**.

After the discussion of several examples, we return to the general objective of the chapter – namely, the *differentation of distributions*. Recall that our point of departure is just (2.5), which we derived from the integration by parts-formula.

Theorem and Definition 2.8 (Differentiation of distributions)**.** Let $\Omega \subset \mathbb{R}^n$ be open and let $T \in \mathscr{D}'(\Omega)$. For each $j \in \{1, ..., n\}$,

$$(\partial_{x_j} T)(\varphi) := -T(\partial_{x_j}\varphi), \qquad \varphi \in \mathrm{C}_c^\infty(\Omega),$$

defines a distribution on Ω. We call $\partial_{x_j} T$ the j-th **distributional (partial) derivative** of T, and call

$$(\partial_{x_1} T, ..., \partial_{x_n} T) \in (\mathscr{D}'(\Omega))^n$$

the **distributional gradient** of T.

Proof. First, $\partial_{x_j} T$ is well-defined: If $\varphi \in \mathrm{C}_c^\infty(\Omega)$, then $\partial_{x_j}\varphi \in \mathrm{C}_c^\infty(\Omega)$ and so $(-T(\partial_{x_j}\varphi))$ has a clear meaning. Second, linearity is a direct consequence of the linearity of T and the (classical) partial derivatives. Now let $\varphi, \varphi_1, ... \in \mathrm{C}_c^\infty(\Omega)$ be such that $\varphi_l \to \varphi$ in $\mathscr{D}(\Omega)$ as $l \to \infty$. As a consequence of this sort of convergence, we have $\sup_{x\in\Omega}|\partial^\alpha(\varphi-\varphi_l)(x)|\to 0$ as $l\to\infty$ for all $\alpha\in\mathbb{N}$. Since this convergence holds for *all* $\alpha\in\mathbb{N}_0^n$ (and since partial derivatives commute), we have $\sup_{x\in\Omega}|\partial^\alpha(\partial_{x_j}\varphi-\partial_{x_j}\varphi_l)(x)|\to 0$. Hence, $\partial_{x_j}\varphi_l\to\partial_{x_j}\varphi$ in $\mathscr{D}(\Omega)$ as $l\to\infty$, and so we obtain

$$(\partial_{x_j}T)(\varphi)\overset{\text{Def}}{=}-T(\partial_{x_j}\varphi)=\lim_{l\to\infty}-T(\partial_{x_j}\varphi_l)\overset{\text{Def}}{=}\lim_{l\to\infty}(\partial_{x_j}T)(\varphi_l),$$

where we have used that $T\in\mathscr{D}'(\Omega)$ in the second step. Therefore, $\partial_{x_j}T\in\mathscr{D}'(\Omega)$, and the proof is complete. □

If we want to call the distributional partial derivatives a *generalisation* of the classical ones, we have to make a consistency check; this is now a direct consequence of (2.5) and the previous definition:

Remark 2.9 (Consistency)**.** Let $\Omega\subset\mathbb{R}^n$ be open and $u\in\mathrm{C}^1(\Omega)$. Denoting T_u the regular distribution associated with u, i.e.,

$$T_u(\varphi):=\int_\Omega u\varphi\,\mathrm{d}x,\qquad\varphi\in\mathrm{C}_c^\infty(\Omega),$$

we have

$$(\partial_{x_j}T_u)=T_{\partial_{x_j}u}\qquad\text{for all } j\in\{1,...,n\}.$$

Indeed, directly employing Definition 2.8 and (2.3), we have

$$\begin{aligned}(\partial_{x_i}T_u)(\varphi)&\overset{\text{Def}}{=}-T_u(\partial_{x_i}\varphi)\overset{\text{Def}}{=}-\int_\Omega u\partial_{x_i}\varphi\,\mathrm{d}x\\&\overset{u\in\mathrm{C}^1(\Omega)}{=}\int_\Omega(\partial_{x_i}u)\varphi\,\mathrm{d}x\overset{\text{Def}}{=}T_{\partial_{x_i}u}(\varphi)\qquad\text{for all }\varphi\in\mathrm{C}_c^\infty(\Omega).\end{aligned}$$

2.2 Weak differentiability and Sobolev spaces

Let $\Omega\subset\mathbb{R}^n$ be open. If $u\in\mathrm{L}^1_{\mathrm{loc}}(\Omega)$, we may consider the regular distribution T_u in the sense of Example 2.3, consider its distributional partial derivatives $\partial_{x_j}T_u$ in the sense of Definition 2.8 and ask

whether *all* $\partial_{x_j}T_u$, $j\in\{1,...,n\}$ are regular distributions again.

If so, we say that u is **weakly differentiable** and belongs to the **(local) Sobolev space** $\mathrm{W}^{1,1}_{\mathrm{loc}}(\Omega)$.

If $u\in\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ is weakly differentiable – so $\partial_{x_j}T_u=T_{u_j}$ for $u_j\in\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ – u_j is uniquely determined by Corollary 2.5. With slight abuse of notation, yet in accordance with Remark 2.9, we then define

$$\partial_{x_j}u:=u_j,\qquad j\in\{1,...,n\}.$$

In this context, we also refer to $\partial_{x_j} u$ as j-th **weak partial derivative**.

Until now, we only worked in the $\mathrm{L}^1_{\mathrm{loc}}$-framework. **Sobolev spaces** arise when we impose L^p-size conditions on weak derivatives as follows:

Definition 2.10 (Sobolev spaces)**.** Let $\Omega \subset \mathbb{R}^n$ be open and let $1 \leq p \leq \infty$. The **Sobolev space** $\mathrm{W}^{1,p}(\Omega)$ is defined as the linear space of all

(a) $u \in \mathrm{L}^p(\Omega)$ which are

(b) weakly differentiable, and

(c) all of their weak partial derivatives belong to $\mathrm{L}^p(\Omega)$, in brief

$$\|u\|_{\mathrm{W}^{1,p}(\Omega)} := (\|u\|^p_{\mathrm{L}^p(\Omega)} + \|\partial_{x_1} u\|^p_{\mathrm{L}^p(\Omega)} + \ldots + \|\partial_{x_n} u\|^p_{\mathrm{L}^p(\Omega)})^{\frac{1}{p}} < \infty$$

if $1 \leq p < \infty$, and, if $p = \infty$,

$$\|u\|_{\mathrm{W}^{1,\infty}(\Omega)} := \|u\|_{\mathrm{L}^\infty(\Omega)} + \|\partial_{x_1} u\|_{\mathrm{L}^\infty(\Omega)} + \ldots + \|\partial_{x_n} u\|_{\mathrm{L}^\infty(\Omega)} < \infty.$$

The higher order variants of weak differentiability and hereafter higher order Sobolev spaces are defined inductively: For instance, we say that $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ is twice weakly differentiable provided all first weak partial derivatives are weakly differentiable. In this context, it is also customary to denote $\partial^\alpha u$ the α-th weak partial derivative (provided it exists). Analogously, the Sobolev space $\mathrm{W}^{k,p}(\Omega)$ then is defined as the collection of all k-times weakly differentiable functions $u\colon \Omega \to \mathbb{R}$ such that

$$\|u\|_{\mathrm{W}^{k,p}(\Omega)} := \Big(\sum_{|\alpha| \leq k} \|\partial^\alpha u\|^p_{\mathrm{L}^p(\Omega)} \Big)^{\frac{1}{p}} < \infty \tag{2.14}$$

with the usual modifications if $p = \infty$.

We now collect some basic facts on Sobolev functions. These facts are stated without proof, and we refer the reader to [3, Chpt. 5] for more detail. Throughout, let $\Omega \subset \mathbb{R}^n$ be open.

- **Completeness.** For any $k \in \mathbb{N}$ and all $1 \leq p \leq \infty$, $(\mathrm{W}^{k,p}(\Omega), \|\cdot\|_{\mathrm{W}^{k,p}(\Omega)})$ is a Banach space.
- **Smooth approximation.** For any $k \in \mathbb{N}$ and all $1 \leq p < \infty$, the subspace $(\mathrm{C}^\infty \cap \mathrm{W}^{k,p})(\Omega)$ is dense in $\mathrm{W}^{k,p}(\Omega)$ for the norm topology[2].
- **Lipschitz functions.** The space $\mathrm{W}^{1,\infty}(\Omega)$ consists of functions which have a (locally) Lipschitz continuous representative. Recall that a function $f\colon \Omega \to \mathbb{R}$ is called **locally Lipschitz** provided for any $x_0 \in \Omega$ and

[2] We will revisit and modify the underlying argument when studying smooth approximation for functions of bounded variation.

$R > 0$ such that $B_R(x_0) \Subset \Omega$ there exists $L > 0$ such that $|f(x) - f(y)| < L|x - y|$ holds for all $x, y \in B_R(x_0)$. It is called **globally Lipschitz** provided the ultimate inequality holds for all $x, y \in \Omega$.

Note that, if $\Omega \subset \mathbb{R}^n$ not only is open but also bounded with smooth boundary, then $W^{1,\infty}(\Omega)$ precisely consists of global Lipschitz functions.

Remark 2.11 (Smooth approximation)**.** If $p = \infty$, we recall from the theory of L^p-spaces that $C^\infty \cap L^\infty$ is not dense in L^∞. Similarly, the Sobolev spaces $W^{k,\infty}(\Omega)$ do **not** admit smooth approximation in the sense that $(W^{k,\infty} \cap C^\infty)(\Omega)$ is not dense in $W^{k,\infty}(\Omega)$. This is seen by considering $k = 1$ as follows.

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with smooth boundary. Suppose towards a contradiction that $(C^\infty \cap W^{1,\infty})(\Omega)$ is dense in $W^{1,\infty}(\Omega)$ for the norm topology on $W^{1,\infty}(\Omega)$. Pick some $u\colon \Omega \to \mathbb{R}$ which is bounded and globally Lipschitz (and thus belongs to $W^{1,\infty}(\Omega)$ by the third item from above), but does not belong to $C_b^1(\Omega)$. Here, $C_b^1(\Omega)$ denotes the continuously differentiable functions which, along with their partial derivatives of order one, are bounded. For the following, we recall that $C_b^1(\Omega)$ is a Banach space with respect to the norm

$$\|v\|_{C_b^1(\Omega)} := \|v\|_{L^\infty(\Omega)} + \|\partial_{x_1} v\|_{L^\infty(\Omega)} + ... + \|\partial_{x_n} v\|_{L^\infty(\Omega)}.$$

Since we assumed smooth approximation, we consequently find a sequence $(u_j) \subset (C^\infty \cap W^{1,\infty})(\Omega)$ such that $\|u_j - u\|_{W^{1,\infty}(\Omega)} \to 0$ as $j \to \infty$. As a consequence, (u_j) is a Cauchy sequence for the $W^{1,\infty}$-norm. However, since the $W^{1,\infty}$-norm coincides with the C_b^1-norm on smooth functions, it is a Cauchy sequence in $C_b^1(\Omega)$. As the latter is Banach, it will converge to some $v \in C_b^1(\Omega)$. As necessarily $u = v$, this would imply that every Lipschitz function is continuously differentiable – which is obviously wrong.

Before we come to the discussion of *jumps*, let us note that the smooth approximation for $W^{k,p}$-functions motivates to define a subspace of functions *vanishing at the boundary*:

Definition 2.12. Let $\Omega \subset \mathbb{R}^n$, $1 \le p < \infty$ and $k \in \mathbb{N}$. We define the Sobolev space of order (k, p) with **zero boundary values** by

$$W_0^{k,p}(\Omega) := \overline{C_c^\infty(\Omega)}^{\|\cdot\|_{W^{k,p}(\Omega)}}$$

As a key problem of the present lecture, we aim to provide a framework within which we may allow for *jumps*. Until now, we merely have a heuristic

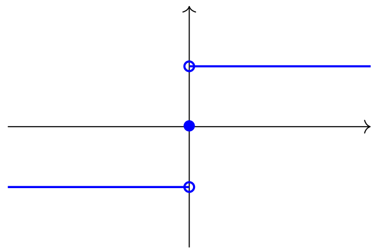

Fig. 3: The sign function as a prototypical example of a function that *jumps*.

understanding of what *jumps* should mean. The easiest example of a function which displays our heuristic understanding of *jumps* is the sign function, see Figure 3. As the next example establishes, this function does **not** belong to $\mathrm{W}^{1,1}_{\mathrm{loc}}(\Omega)$ – in other words, it is not weakly differentiable. In consequence, Sobolev spaces do **not** provide the right framework to deal with jump functions.

Example 2.13. As alluded to above, let $u\colon (-1,1) \to \{-1,0,1\}$ be the sign-function:

$$u(x) := \begin{cases} -1 & \text{if } -1 < x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } 0 < x < 1. \end{cases}$$

We now compute the distributional derivative of T_u. For this, let $\varphi \in \mathrm{C}^\infty_c((-1,1))$ be arbitrary. Using the fundamental theorem of calculus, we then find

$$\begin{aligned} \partial_x T_u(\varphi) &\overset{\text{Def}}{=} -T_u(\partial_x\varphi) = -\int_{(-1,1)} u\,\partial_x\varphi\,\mathrm{d}x \\ &= -\int_{-1}^{0} u\partial_x\varphi\,\mathrm{d}x - \int_0^1 u\partial_x\varphi\,\mathrm{d}x = \int_{-1}^{0}\partial_x\varphi\,\mathrm{d}x - \int_0^1 \partial_x\varphi\,\mathrm{d}x \\ &= (\varphi(0)-\varphi(-1)) - (\varphi(1)-\varphi(0)) = 2\varphi(0) = T_{2\delta_0}(\varphi). \end{aligned}$$

By Example 2.7, $T_{2\delta_0}$ is a measure regular distribution which is *not* regular. Therefore, $u \notin \mathrm{W}^{1,1}_{\mathrm{loc}}(\Omega)$.

2.3 Weak formulations of partial differential equations*

This section – which has not been addressed in class – serves to recall some concepts from the weak theory of partial differential equations. *Weak solutions* are a more general notion of solution than the usual classical solutions. By virtue of the higher generality of the concept of solutions, weak solutions are easier to find by means of functional analytic methods. As a drawback, the true difficulty then often lies in establishing regularity properties of such weak solutions, sometimes identifying them as classical solutions.

In general, *weak* formulations involve only lower order derivatives than actually required by the underlying differential operator; usually, they are obtained by succesive application of the integration by parts formula (2.3). As will become obvious from the discussion below, this is reflected by weak formulations being defined via *testing the equation against* $\mathrm{C}_c^\infty(\Omega)$*-functions*. This is also why $\mathrm{C}_c^\infty(\Omega)$ are often referred to as *test functions*.

Here is an example: Given an open set $\Omega \subset \mathbb{R}^n$, suppose that we wish to solve the partial differential equation

$$-\Delta u = f \qquad \text{in } \Omega \tag{2.15}$$

where $f \in \mathrm{L}^\infty(\Omega)$. This equation cannot be stated classically. To obtain a weak formulation, we let $\varphi \in \mathrm{C}_c^\infty(\Omega)$ be arbitrary. We multiply both sides of the equation with φ and integrate over Ω to find

$$\int_\Omega (-\Delta u)\varphi \,\mathrm{d}x = \int_\Omega f\varphi \,\mathrm{d}x.$$

Using that $\varphi \in \mathrm{C}_c^\infty(\Omega)$, we integrate by parts and obtain

$$\int_\Omega \langle \nabla u, \nabla \varphi \rangle \,\mathrm{d}x = \int_\Omega f\varphi \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega). \tag{2.16}$$

This equation only involves zero and first order derivatives on $u \in \mathrm{W}^{1,2}_{\mathrm{loc}}(\Omega)$, and thus might be regarded as a *weak version* of the partial differential equation (2.15). Also note that it is crucial for (2.16) to hold for *all* C_c^∞-functions – which is a key aspect of weak formulations.

Lecture 1, April 16, 2021

Lecture 2, April 16, 2021

3 The direct method of the Calculus of Variations

In the introductory Section 1 we have encountered two minimisation problems – the minimal surface problem (1.1) and the minimisation of Rudin-Osher-Fatemi-type functionals (1.2). In this chapter we discuss a method that – subject to some assumptions on the functionals – allows us to conclude the existence of minimisers. This is the **direct method of the Calculus of Variations**.

By the very nature of the method, we shall see that the minimisation of a large class of functionals can be tackled succesfully in Sobolev spaces. As we will discover, however, the functional analytic setup for the model problems from the introduction is **not** directly implementable in Sobolev spaces. This will give us a solid motivation to study BV-functions from the viewpoint of functional analysis.

3.1 An abstract version of the direct method

In this section we provide a scheme to establish the existence of minimisers of functionals. Our setting is not the most general one, but it suffices for all of what follows.

Let $(X, \|\cdot\|)$ be a normed space, $D \subset X$ a non-empty subset and let

$$\mathcal{F}\colon X \to \mathbb{R} \tag{3.1}$$

be a functional. Our aim is to establish the existence of a **minimiser** of $\mathcal{F}$ over D, i.e., an element $x \in D$ such that

$$\mathcal{F}[x] = \inf_{y\in D} \mathcal{F}[y]. \tag{3.2}$$

In the sequel, we present several assumptions on X, D and $\mathcal{F}$ that turn out crucial for this objective. Let us hereafter *assume* that

(a) $\mathcal{F}$ is **bounded below** on D: There exists $\widetilde{m} \in \mathbb{R}$ such that $\mathcal{F}[x] \geq \widetilde{m}$ holds for all $x \in X$. In consequence,

$$m := \inf_{y\in D} \mathcal{F}[y]$$

exists and is finite. Therefore, by the very definition of the infimum, we find a **minimising sequence** $(x_j) \subset D$ such that

$$\mathcal{F}[x_j] \to m = \inf_{y\in D} \mathcal{F}[y]. \tag{3.3}$$

Our overall aim now is to establish that (x_j) possesses a subsequence which, in a suitable sense, converges to a minimiser. This is a **compactness feature**. Compactness relies on boundedness, and so we further assume that

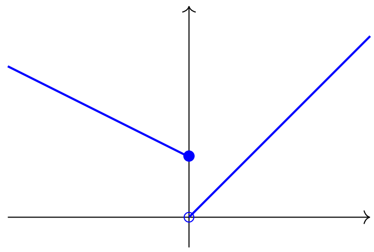

Fig. 4: The graph of a function $f\colon \mathbb{R} \to \mathbb{R}$ that is not lower semicontinuous at the origin. One directly sees that f does not admit a minimiser.

(b) $\mathcal{F}$ is **coercive**: If $(y_j) \subset X$ is a sequence with $\|y_j\| \to \infty$, then there holds $\lim_{j\to\infty} \mathcal{F}[y_j] = +\infty$.

For us, this has the following implication: By (3.3), $\sup_{j\in\mathbb{N}} \mathcal{F}[x_j] < \infty$. Thus, the assumed coercivity of $\mathcal{F}$ yields that (x_j) is bounded: There exists $R > 0$ such that $\|x_j\| \leq R$ holds for all $j \in \mathbb{N}$.

Toward compactness, we now make another assumption:

(c) **Compactness:** There exists a notion of convergence '$\rightsquigarrow$' with the following property: If $(y_j) \subset X$ is a sequence which is bounded with respect to $\|\cdot\|$, then there exists $y \in X$ and a subsequence $(y_{j(k)}) \subset (y_j)$ such that $y_{j(k)} \rightsquigarrow y$ as $k \to \infty$.

Then the minimising sequence (x_j) (which is bounded because of (b)) possesses a subsequence $(x_{j(k)})$ such that $x_{j(k)} \rightsquigarrow x$ as $k \to \infty$.

The element x provided in the ultimate item will turn out the right candidate for a minimiser. In order to identify it as a minimiser, we need two final ingredients:

(d) **Compatibility:** The set D is closed with respect to the convergence '$\rightsquigarrow$'. By (c), we have $x_{j(k)} \in D$ for all $k \in \mathbb{N}$, and $x_{j(k)} \rightsquigarrow x$ will consequently imply that $x \in D$. Thus x is admissible for the minimisation problem (3.2).

(e) **Lower semicontinuity:** We finally assume that $\mathcal{F}$ is lower semicontinuous with respect to the convergence '$\rightsquigarrow$': If $(y_j) \subset X$ and $y \in X$ are such that $y_j \rightsquigarrow y$, then there holds

$$\mathcal{F}[y] \leq \liminf_{j\to\infty} \mathcal{F}[y_j]. \tag{3.4}$$

The importance of lower semicontinuity for the existence of minima is depicted in Figure 4.

If all of (a)–(e) are in action, we conclude the existence of a minimiser as follows: By (c), $x_{j(k)} \rightsquigarrow x$ for some subsequence $(x_{j(k)}) \subset (x_j)$. By (d), we have $x \in D$ and finally, by (e),

$$\inf_{y\in D} \mathcal{F}[y] \overset{x\in D}{\leq} \mathcal{F}[x] \overset{(3.4)}{\leq} \liminf_{k\to\infty} \mathcal{F}[x_{j(k)}] \overset{(3.3)}{=} \inf_{y\in D} \mathcal{F}[y].$$

In consequence, $x \in D$ is a minimiser for $\mathcal{F}$ over D. We have thus proved the following abstract version of the *direct method*:

The direct method of CoV: Let $(X, \|\cdot\|)$ be a normed space, $D \subset X$ a non-empty subset and $\mathcal{F}\colon X \to \mathbb{R}$ be a function such that the following hold:

(a) $\mathcal{F}$ is **bounded below**.

(b) $\mathcal{F}$ is **coercive**: If $(y_j) \subset X$ satisfies $\lim_{j\to\infty} \|y_j\| = \infty$, then $\lim_{j\to\infty} \mathcal{F}[y_j] = \infty$.

(c) There exists a notion of convergence '$\rightsquigarrow$' with the following properties:

- **Compactness:** If $(y_j) \subset X$ is a sequence that is bounded for $\|\cdot\|$, then there exists $y \in X$ and a subsequence $(y_{j(k)}) \subset (y_j)$ such that

$$y_{j(k)} \rightsquigarrow y, \qquad k \to \infty.$$

- **Lower semicontinuity:** $\mathcal{F}$ is lower semicontinuous with respect to '$\rightsquigarrow$': Whenever $y_j \rightsquigarrow x$ as $j \to \infty$, then

$$\mathcal{F}[y] \leq \liminf_{j\to\infty} \mathcal{F}[y_j].$$

- **Compatibility with D:** D is closed with respect to '$\rightsquigarrow$': If $(y_j) \subset D$ is a sequence that converges to $y \in X$ with respect to '$\rightsquigarrow$', then $y \in D$.

Then there exists a minimiser $x \in D$ of $\mathcal{F}$ over D.

We conclude with a remark:

Remark 3.1 (On compactness and lower semicontinuity)**.**

- Note that **compactness** and **lower semicontinuity** are two **competing properties** for the functional $\mathcal{F}$: The weaker our notion of convergence '$\rightsquigarrow$', the more sequences will converge and so it is harder for $\mathcal{F}$ to qualify as a lower semicontinuous functional. As such, we need to strike a balance: The convergence '$\rightsquigarrow$' must be weak enough to ensure sufficient compactness properties, but strong enough to not destroy the lower semicontinuity of $\mathcal{F}$.

- The usual norm convergence is **not a good candidate**. To see this, recall that in any infinite dimensional space there is at least *one* sequence that is bounded but does not possess any convergent subsequence (for the norm topology). The standard example is this: Consider, for $1 \leq p < \infty$, the sequence spaces $\ell^p(\mathbb{N})$. For $j \in \mathbb{N}$, define a sequence $e_j := (\delta_{ij})_{i\in\mathbb{N}}$. Then $\|e_j\|_{\ell^p(\mathbb{N})} = 1$ for any $1 \leq p < \infty$, so (e_j) is bounded. Now, if $e_j \to x = (x_1, x_2, ...)$ for the norm topology in $\ell^p(\mathbb{N})$, we conclude that $x_i = 0$ for all $i \in \mathbb{N}$. However, then we have $\|x - e_j\|_{\ell^p(\mathbb{N})} = \|e_j\|_{\ell^p(\mathbb{N})} = 1$ for all $j \in \mathbb{N}$ and so there is no subsequence of (e_j) that converges to x for the norm topology.

3.2 A class of model functionals

In this section, we come up with a unifying framework that allows to deal with the functionals considered so far. We will consequently exhibit scenarious where the direct method of the previous section can be applied.

Let $\Omega \subset \mathbb{R}^n$ be open and bounded with boundary of class C^1. For $1 \leq p < \infty$, we moreover suppose that

- $F\colon \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable[3] and convex function[4] which satisfies the **growth bounds**

$$c_1|z|^p - c_2 \leq F(z) \leq c_3(1 + |z|^p) \qquad \text{for all } z \in \mathbb{R}^n \tag{3.5}$$

  for fixed constants $c_i > 0$, $i \in \{1, 2, 3\}$. Let us recall that F is called **convex** provided

$$F(\lambda\xi + (1-\lambda)\eta) \leq \lambda F(\xi) + (1-\lambda)F(\eta)$$

  holds for all $\xi, \eta \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

- $q \geq p$ is an exponent for which there exists $c > 0$ such that there holds

$$\|u\|_{\mathrm{L}^q(\Omega)} \leq c\|u\|_{\mathrm{W}^{1,p}(\Omega)} \qquad \text{for all } u \in \mathrm{W}^{1,p}(\Omega). \tag{3.6}$$

  We will see later that we may always take some q that is *strictly larger* than p. This is the content of the so-called *Sobolev embedding theorem*.

Subject to these assumptions, we then consider the functionals (where $\lambda \geq 0$)

$$\mathscr{F}[u] := \int_\Omega F(\nabla u)\,\mathrm{d}x + \frac{\lambda}{2}\int_\Omega |u - f|^q\,\mathrm{d}x \tag{3.7}$$

in two different settings – one linked to the minimal surface functional, one to the Rudin-Osher-Fatemi functional. In any case, note that subject to the above assumptions, the functional $\mathscr{F}$ is well-defined on the entire Sobolev space $\mathrm{W}^{1,p}(\Omega)$ and so in particular on any of its subsets.

The relevant subsets D that we are interested in most are

[3] This assumption appears for simplicity only; mere continuity would do as well.

[4] As convexity implies continuity in this context, this assumption is slightly redundant.

- **Dirichlet classes:** For a given $u_0 \in \mathrm{W}^{1,p}(\Omega)$, the Dirichlet class corresponding to u_0 is given by

$$D_1 := u_0 + \mathrm{W}_0^{1,p}(\Omega).$$

  The terminology is natural: Recalling Definition 2.12, elements of $\mathrm{W}_0^{1,p}(\Omega)$ are interpreted to vanish at the boundary $\partial\Omega$, and so elements of D_1 will precisely have the boundary values of u_0 along $\partial\Omega$.

- **Neumann**[5] **classes:** Here we put

$$D_2 := \mathrm{W}^{1,p}(\Omega) \cap \left\{ v \in \mathrm{L}^1(\Omega) \colon \int_\Omega v \,\mathrm{d}x = 0 \right\}.$$

Subject to the above assumptions, we aim to establish the existence of a minimiser, so a solution of the variational principle

$$\text{minimise } \mathscr{F}[u] \text{ over } u \in D_1 \text{ or } u \in D_2. \tag{3.8}$$

Our plan is to implement the direct method as discussed in the previous section. Towards this aim, we need some background from functional analysis and Sobolev spaces to be recorded in the subsequent section[6]. This, in particular, will canonically lead to the distinction of the cases $1 < p < \infty$ and $p = 1$.

3.2.1 Auxiliary facts from functional analysis and Sobolev spaces

In view of the direct method, we must come up with a notion of convergence that yields some sort of compactness. That is, if $(u_j) \subset \mathrm{W}^{1,p}(\Omega)$ is bounded, we wish to conclude $u_{j(k)} \rightsquigarrow u$ for some $u \in \mathrm{W}^{1,p}(\Omega)$ and some sort of convergence '$\rightsquigarrow$'.

For a variety of spaces, compactness still can be saved when passing to a weaker sort of convergence than norm convergence. Let $(X, \|\cdot\|)$ be a normed space. We say that $(x_j) \subset X$ **converges weakly** to $x \in X$ provided

$$x'(x_j) \to x'(x) \qquad \text{for all } x' \in X',$$

and write $x_j \rightharpoonup u$. Here, as usual, X' denotes the continuous dual of X. On the other hand, a sequence $(x'_j) \subset X'$ **converges in the weak*-sense** to $x' \in X'$ provided

$$x'_j(x) \to x'(x) \qquad \text{for all } x \in X,$$

and write $x'_j \overset{*}{\rightharpoonup} x'$.

The fundamental compactness principle we shall rely on is this:

---

[5] In class, we did not call these classes 'Neumann' classes. The reason will become clear later, and we accept this terminology for the time being.

[6] For more detail, see the lecture notes of Robert Denk (in particular, chapters 3b), 5c) and 9).

Theorem 3.2 (Banach-Alaoglu-Bourbaki)**.** Let $(X, \|\cdot\|_X)$ be a **separable Banach space**. Whenever $(x'_j) \subset X'$ is a sequence which is bounded for $\|\cdot\|_{X'}$, there exists $x' \in X'$ and a subsequence $(x'_{j(k)}) \subset (x'_j)$ such that

$$x'_{j(k)} \overset{*}{\rightharpoonup} x' \qquad \text{as } k \to \infty.$$

In the form as given in Theorem 3.2, the Banach-Alaoglu-Bourbaki theorem yields a compactness result on *dual spaces*; note that not every Banach space arises as a dual space[7]. For a very vast class of Banach spaces it is still possible to obtain a compactness result for the spaces themselves. These are the reflexive spaces to be discussed next.

As usual, we denote $X'' := (X')'$ the bidual of a normed space $(X, \|\cdot\|_X)$. To motivate the concept of reflexivity, note that elements of X'' are continuous linear functionals on X'. The probably easiest example of such an object is, for a given $x \in X$, the map

$$\iota(x) \colon X' \ni x' \mapsto x'(x).$$

This is nothing but the **evaluation functional** that evaluates $x' \in X'$ at $x \in X$.

By means of $\iota\colon X \ni x \mapsto \iota(x) \in X''$, we obtain a linear isometry from X to X'' (this step uses the Hahn-Banach theorem). We may hereafter say that X embeds into X''. If every $x'' \in X''$ arises in this way – that is to say, $\iota\colon X \to X''$ is surjective – then we call X **reflexive**. Now we have

Corollary 3.3 (Banach-Alaoglu-Bourbaki, reflexive spaces)**.** Let $(X, \|\cdot\|)$ be a separable and reflexive Banach space. Whenever $(x_j) \subset X$ is a sequence which is bounded for $\|\cdot\|_X$, there exists $x \in X$ and a subsequence $(x_j) \subset X$ such that

$$x_{j(k)} \rightharpoonup x \qquad \text{as } k \to \infty.$$

Proof. Since $\iota\colon X \to X''$ is an isometry, the sequence $(\iota(x_j))$ is bounded in X''. On the other hand, since X is reflexive and separable, so is $(X'', \|\cdot\|_{X''})$ (take a countable dense subset of X and use ι to obtain a countable dense subset of X'' – then ι being an isometry yields the desired countable dense subset of X''). As a consequence of Hahn-Banach, if $(X'', \|\cdot\|_{X''})$ is separable, $(X', \|\cdot\|_{X'})$ must be separable, too.

[7] Using the so-called Krein-Milman theorem, one is able to prove that, e.g., L^1 is not the dual of any normed space, and we will pick up on this fact later on.

We are now in the situation of Theorem 3.2, being applied to X' and X''. Theorem 3.2 implies the existence of some $y'' \in X''$ such that $\iota(x_{j(k)}) \overset{*}{\rightharpoonup} y''$ for some suitable subsequence $(x_{j(k)})$. Since ι is bijective, there exists a unique $x \in X$ with $y'' = \iota(x)$. Therefore, $\iota(x_{j(k)} - x) \overset{*}{\rightharpoonup} 0$ as $k \to \infty$. Now let $x' \in X'$ be arbitrary. Then,

$$x'(x_{j(k)} - x) = \iota(x_{j(k)} - x)(x') \to 0, \qquad k \to \infty,$$

and so $x_{j(k)} \rightharpoonup x$ in X. The proof is complete. □

As we wish to finally apply the foregoing theory to Sobolev spaces, we now identify the reflexive Sobolev spaces.

Remark 3.4 (Reflexivity of $\mathrm{W}^{k,p}$)**.** Let $\Omega \subset \mathbb{R}^n$ be open, $k \in \mathbb{N}$ and $1 \leq p \leq \infty$. Then we have

$$(\mathrm{W}^{k,p}(\Omega); \|\cdot\|_{\mathrm{W}^{k,p}(\Omega)}) \text{ reflexive } \Longleftrightarrow 1 < p < \infty.$$

Proof of Remark 3.4. The proof of the remark was not discussed in class. We start from the fact that $(\mathrm{L}^p(\Omega), \|\cdot\|_{\mathrm{L}^p(\Omega)})$ is a reflexive Banach space if and only if $1 < p < \infty$. Now consider the linear map

$$\Phi\colon\ \mathrm{W}^{k,p}(\Omega) \ni u \mapsto (\partial^\alpha u)_{|\alpha| \leq k} \in (\mathrm{L}^p(\Omega))^{\mathtt{k}},$$

where $\mathtt{k} := \{\alpha \in \mathbb{N}_0^n\colon\ |\alpha| \leq k\}$. We claim:

(a) Φ is a linear isometry for a suitable equivalent norm on $\mathrm{W}^{k,p}(\Omega)$,

(b) $\Phi(\mathrm{W}^{k,p}(\Omega))$ is closed in $\mathrm{L}^p(\Omega)^{\mathtt{k}}$,

(c) Closed subspaces of reflexive Banach spaces are reflexive,

(d) If $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ are isometrically isomorphic and $(Y, \|\cdot\|_Y)$ is reflexive, so is $(X, \|\cdot\|_X)$.

Ad (a). The space $\mathrm{L}^p(\Omega)^{\mathtt{k}}$ is endowed with the norm $\|(f_\alpha)_{|\alpha|\leq k}\|_{\mathrm{L}^p(\Omega)^{\mathtt{k}}} := \sum_{|\alpha|\leq k} \|f_\alpha\|_{\mathrm{L}^p(\Omega)}$. The norm on $\mathrm{W}^{k,p}(\Omega)$ is given by

$$\|u\|_{\mathrm{W}^{k,p}(\Omega)} := \Big(\sum_{|\alpha|\leq k} \|\partial^\alpha u\|_{\mathrm{L}^p(\Omega)}^p\Big)^{\frac{1}{p}},$$

but $\|u\|_{\mathrm{W}^{k,p}(\Omega)}^{\sim} := \sum_{|\alpha|\leq k} \|\partial^\alpha u\|_{\mathrm{L}^p(\Omega)}$ is an equivalent norm on $\mathrm{W}^{k,p}(\Omega)$ for which the latter still is reflexive. Equipping $\mathrm{W}^{k,p}(\Omega)$ with this norm, by definition, Φ then is an isometry as indicated.

Ad (b). Let $(\Phi(u_j))_j$ converge to some $v = (v_\alpha)_{|\alpha|\leq k} \in \mathrm{L}^p(\Omega)^{\mathtt{k}}$ in $\mathrm{L}^p(\Omega)^{\mathtt{k}}$. This means that for every $\alpha \in \mathbb{N}_0^n$ there exists $v^\alpha \in \mathrm{L}^p(\Omega)$ such that $\partial^\alpha u_j \to$

v^α. We have to prove that $v \in \mathrm{W}^{k,p}(\Omega)$ and $\partial^\alpha v = v^\alpha$. Let $\varphi \in \mathrm{C}_c^\infty(\Omega)$ be arbitrary. Then there holds

$$\begin{aligned}\int_\Omega v \partial^\alpha \varphi \,\mathrm{d}x &= \int_\Omega (v - u_j)\partial^\alpha \varphi \,\mathrm{d}x + \int_\Omega u_j \partial^\alpha \varphi \,\mathrm{d}x \\ &= \int_\Omega (v - u_j)\partial^\alpha \varphi \,\mathrm{d}x + (-1)^{|\alpha|} \int_\Omega ((\partial^\alpha u_j) - v^\alpha)\varphi \,\mathrm{d}x \\ &\qquad\qquad + (-1)^{|\alpha|} \int_\Omega v^\alpha \varphi \,\mathrm{d}x \\ &=: \mathrm{I}_j + \mathrm{II}_j + (-1)^{|\alpha|} \int_\Omega v^\alpha \varphi \,\mathrm{d}x.\end{aligned}$$

Now, by Hölder's inequality (where $p' = \frac{p}{p-1}$),

$$|\mathrm{I}_j| + |\mathrm{II}_j| \le \|v - u_j\|_{\mathrm{L}^p(\Omega)} \|\partial^\alpha \varphi\|_{\mathrm{L}^{p'}(\Omega)} + \|\partial^\alpha u_j - v^\alpha\|_{\mathrm{L}^p(\Omega)} \|\varphi\|_{\mathrm{L}^{p'}(\Omega)} \to 0$$

as $j \to \infty$. Therefore,

$$\int_\Omega v \partial^\alpha \varphi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega v^\alpha \varphi \,\mathrm{d}x \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega),$$

and so v belongs to $\mathrm{W}^{k,p}(\Omega)$ with $\Phi(v) = (\partial^\alpha v)_{|\alpha| \le k}$. In consequence, $\Phi(\mathrm{W}^{k,p}(\Omega))$ is closed.

Ad (c). Let U be a closed subspace of a reflexive space $(X, \|\cdot\|_X)$. Then $(U, \|\cdot\|_X)$ is a normed space, too. Let $u'' \in U''$. We need to show that there exists $u \in U$ with $u'' = \iota(u)$. To use the reflexivity of X, we first extend u'' to some x'' by

$$x''(x') := u''(x'|_U), \qquad x' \in X'. \tag{3.9}$$

Note carefully that $x'|_U \in U'$ for any $x' \in X'$: Indeed, $x'|_U$ is linear and

$$\begin{aligned}\|x'|_U\|_{U'} &= \sup\{x'(x) \colon\; x \in U,\; \|x\|_X \le 1\} \\ &\le \sup\{x'(x) \colon\; x \in X,\; \|x\|_X \le 1\} \le \|x'\|_{X'}.\end{aligned}$$

Thus x'' given by (3.9) is well-defined and belongs to X''. By reflexivity of X, there exists $\xi \in X$ with $x'' = \iota(\xi)$. Suppose that $\xi \notin U$. Then, using a corollary of Hahn-Banach and the closedness of U, we find $\widetilde{x}' \in X'$ with $\widetilde{x}'|_U = 0$ and $\widetilde{x}'(\xi) = 1$. Then,

$$1 = \widetilde{x}'(\xi) = x''(\widetilde{x}') = u''(\widetilde{x}'|_U) = 0,$$

which is an obvious contradiction. Thus, $\xi \in U$, and so U is reflexive.

Ad (d). Denote the underlying isometric isomorphism by $\Phi \colon X \to Y$. In this situation, we have for any $y \in Y'$:

$$\|y'\|_{Y'} = \sup\{y'(y) \colon\; \|y\|_Y \le 1\} = \sup\{y'(\Phi(x)) \colon\; \|x\|_X \le 1\} = \|y' \circ \Phi\|_{X'}.$$

In consequence, the duals (and by iteration) the biduals are isomorphic as well, and from here the claim follows.

For $p = 1, \infty$, one may either imitate the construction that leads to the non-reflexivity of L^1 or L^∞. Alternatively, we shall see later by an explicit example that $\mathrm{W}^{k,1}(\Omega)$ is not reflexive. $\square$

3.2.2 (Sequential) Lower semicontinuity in Sobolev spaces

The previous section has provided us with useful (weak) compactness results in Sobolev spaces – at least, if $1 < p < \infty$. Towards the implementation of the direct method, we moreover need a lower semicontinuity result. We begin with

Theorem 3.5 (Weak lower semicontinuity in Lebesgue spaces)**.** Let $F\colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be continuously differentiable and convex. Let $\Omega \subset \mathbb{R}^n$ be open and bounded, $z, z_1, z_2, ... \in \mathrm{L}^1(\Omega;\mathbb{R}^n)$ be such that

$$z_j \rightharpoonup z \qquad \text{in } \mathrm{L}^1(\Omega;\mathbb{R}^n).$$

Then there holds

$$\int_\Omega F(z) \leq \liminf_{j\to\infty} \int_\Omega F(z_j)\,\mathrm{d}x.$$

Lecture 2, April 23, 2021

Lecture 3, April 30, 2021

To prove Theorem 3.5, we require the following result from measure theory[8]:

Lemma 3.6 (Lusin)**.** Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be measurable and be such that $f \equiv 0$ outside an open set A of finite Lebesgue measure. Then, for any $\varepsilon > 0$, there exists $g_\varepsilon \in \mathrm{C}_0(A)$ such that

$$\mathscr{L}^n(\{x \in \mathbb{R}^n\colon\ f(x) \neq g_\varepsilon(x)\}) < \varepsilon.$$

Proof of Theorem 3.5. We split the proof into three steps:

- *Step 1.* Approximation and reduction to continuous functions,
- *Step 2.* Convexity and lower semicontinuity,
- *Step 3.* Coming back to our original setting.

Step 1. For each $i \in \mathbb{N}$ there exists a measurable set $\widetilde{K}_i$ such that $z|_{\widetilde{K}_i}$ is continuous together with $\mathscr{L}^n(\Omega \setminus \widetilde{K}_i) < \frac{1}{2i}$. By the regularity properties

[8] See Thm. 6.18 in the lecture notes by Robert Denk.

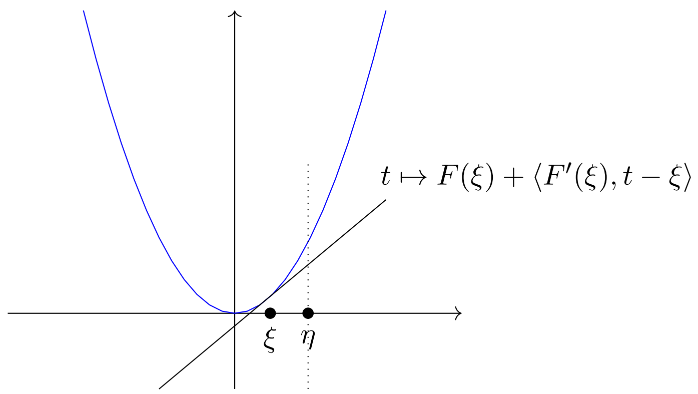

Fig. 5: The idea behind the convexity inequality (3.10): Whenever we form the tangent at some point ξ, the tangent will remain below the graph of F.

of Lebesgue measure, we thus even find a compact set $K_i \subset \widetilde{K}_i$ such that $\mathscr{L}^n(\Omega \setminus K_i) < \frac{1}{i}$. Now,

$$\int_\Omega |\mathbb{1}_\Omega - \mathbb{1}_{K_i}| \,\mathrm{d}x = \mathscr{L}^n(\Omega \setminus K_i) < \frac{1}{i} \to 0, \qquad i \to \infty.$$

Passing to a subsequence if necessary, we may thus assume that $\mathbb{1}_{K_i} \to \mathbb{1}_\Omega$ $\mathscr{L}^n$-a.e. on Ω. Now, by our assumptions on F (i.e., continuity and non-negativity),

$$\int_\Omega F(z) \,\mathrm{d}x \leq \liminf_{i\to\infty} \int_\Omega \mathbb{1}_{K_i} F(z) \,\mathrm{d}x \leq \int_\Omega F(z) \,\mathrm{d}x$$

so that the lim inf is actually a limit. In particular,

(a) $f|_{K_i}$ is continuous, and

(b) there exists $\theta\colon \mathbb{N} \to \mathbb{R}_{\geq 0}$ with $\theta(i) \to 0$ as $i \to \infty$ such that

$$\int_{K_i} F(z) \,\mathrm{d}x \geq \int_\Omega F(z) \,\mathrm{d}x - \theta(i) \qquad \text{for all } i \in \mathbb{N}.$$

Step 2. We now come to the crucial impact of convexity for lower semicontinuity. Namely, if $F\colon \mathbb{R}^n \to \mathbb{R}$ is a convex and continuously differentiable function, then we have

$$F(\xi) + \langle F'(\xi), \eta - \xi \rangle \leq F(\eta) \qquad \text{for all } \xi, \eta \in \mathbb{R}^n. \tag{3.10}$$

Inequality (3.10) means nothing but that, as F is convex, the tangent planes to the graph of F lies *under* the graph of F – see Figure 5. We apply (3.10) pointwisely with $\eta = z_j(x)$ and $\xi = z(x)$ to get

$$\langle F'(z(x)), z_j(x)) - z(x) \rangle \leq F(z_j(x)) - F(z) \qquad \text{for all } x \in \Omega. \tag{3.11}$$

We now use (3.11) to conclude that for each $i \in \mathbb{N}$ there holds

$$\begin{aligned}\int_{K_i} F(z_j)\,\mathrm{d}x &\geq \int_{K_i} F(z)\,\mathrm{d}x + \int_{K_i} F(z_j) - F(z)\,\mathrm{d}x \\ &\geq \int_{K_i} F(z)\,\mathrm{d}x + \int_{K_i} \langle F'(z), z_j - z\rangle\,\mathrm{d}x \\ &= \int_{K_i} F(z)\,\mathrm{d}x + \int_{\Omega} \langle \mathbb{1}_{K_i} F'(z), z_j - z\rangle\,\mathrm{d}x \\ &\longrightarrow \int_{K_i} F(z)\,\mathrm{d}x \quad j \to \infty.\end{aligned} \tag{3.12}$$

For the ultimate step, note that $z|_{K_i}$ is continuous and K_i is compact, so $F' \circ z \colon K_i \to \mathbb{R}^n$ is continuous and bounded. In particular, as $z_j \rightharpoonup z$ in $\mathrm{L}^1(\Omega; \mathbb{R}^n)$,

$$\int_{\Omega} \langle \mathbb{1}_{K_i} F'(z), z_j - z\rangle\,\mathrm{d}x \to 0, \qquad j \to \infty.$$

Step 3. We now combine (3.12) and (b) to find by virtue of $F \geq 0$:

$$\begin{aligned}\liminf_{j\to\infty} \int_{\Omega} F(z_j)\,\mathrm{d}x &\geq \liminf_{j\to\infty} \int_{K_i} F(z_j)\,\mathrm{d}x \\ &\geq \int_{\Omega} F(z)\,\mathrm{d}x - \theta(i) \overset{i\to\infty}{\longrightarrow} \int_{\Omega} F(z)\,\mathrm{d}x.\end{aligned}$$

The proof is complete. □

We now provide a useful corollary for functionals defined on gradients of Sobolev functions.

Corollary 3.7. Let $F \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ be continuously differentiable, convex and satisfy the growth bound

$$c_1|z|^p - c_2 \leq F(z) \leq c_3(1 + |z|^p) \qquad \text{for all } z \in \mathbb{R}^n. \tag{3.13}$$

Let $\Omega \subset \mathbb{R}^n$ be open and bounded and $u, u_1, u_2, \ldots \in \mathrm{W}^{1,p}(\Omega)$ such that

$$u_j \rightharpoonup u \qquad \text{in } \mathrm{W}^{1,p}(\Omega).$$

Then there holds

$$\int_{\Omega} F(\nabla u)\,\mathrm{d}x \leq \liminf_{j\to\infty} \int_{\Omega} F(\nabla u_j)\,\mathrm{d}x.$$

Proof. Note that by (3.13), the functional

$$\mathrm{W}^{1,p}(\Omega) \ni v \mapsto \int_{\Omega} F(\nabla v)\,\mathrm{d}x$$

is well-defined. By Theorem 3.5, it now suffices to prove that $u_j \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$ implies $\nabla u_j \rightharpoonup \nabla u$ in $\mathrm{L}^1(\Omega;\mathbb{R}^n)$. Let $\varphi \in \mathrm{L}^\infty(\Omega;\mathbb{R}^n)$. Consider the functional

$$\Psi\colon\ \mathrm{W}^{1,p}(\Omega) \ni v \mapsto \int_\Omega \langle \nabla v, \varphi\rangle \,\mathrm{d}x$$

Then, since $\Omega \subset \mathbb{R}^n$ is open and bounded,

$$|\Psi(v)| \leq c\|\nabla v\|_{\mathrm{L}^p(\Omega)}\|\varphi\|_{\mathrm{L}^{p'}(\Omega)} \leq c\mathscr{L}^n(\Omega)^{\frac{1}{p'}}\|v\|_{\mathrm{W}^{1,p}(\Omega)}\|\varphi\|_{\mathrm{L}^\infty(\Omega)}$$

and so $\Psi \in (\mathrm{W}^{1,p}(\Omega))'$. Hence $\Psi(u_j) \to \Psi(u)$, and thus $\nabla v_j \rightharpoonup \nabla v$ in $\mathrm{L}^1(\Omega)$. Theorem 3.5 now completes the proof. □

3.2.3 Minimisers for the model functionals, $1 < p < \infty$

We now have gathered all the tools that we need to implement the direct method for our model functionals from the very beginning of this section (cf. (3.7)):

Theorem 3.8. Let $1 < p < \infty$. Moreover, let $\Omega \subset \mathbb{R}^n$ be open and bounded, $u_0 \in \mathrm{W}^{1,p}(\Omega)$ and suppose that $F\colon \mathbb{R}^n \to \mathbb{R}$ is a continuous and convex integrand that satisfies (3.5). Then the following hold:

(a) *Ad* D_1. For each $u_0 \in \mathrm{W}^{1,p}(\Omega)$, the variational principle

$$\text{to minimise } \mathscr{F}[v] = \int_\Omega F(\nabla v)\,\mathrm{d}x \quad \text{over } v \in D_1 = u_0 + \mathrm{W}_0^{1,p}(\Omega)$$

has a solution $u \in D_1$.

(b) *Ad* D_2. If Ω moreover is connected and has boundary of class C^1, then the variational principle

$$\text{to minimise } \mathscr{F}[v] = \int_\Omega F(\nabla v)\,\mathrm{d}x$$
$$\text{over } v \in D_2 = \mathrm{W}_0^{1,p}(\Omega) \cap \{v \in \mathrm{L}^1(\Omega)\colon\ \int_\Omega v\,\mathrm{d}x = 0\}$$

has a solution $u \in D_2$.

We argue via the direct method, cf. Section 3.1, which we implement in the Sobolev space $\mathrm{W}^{1,p}(\Omega)$. We first deal with the case of Dirichlet classes D_1.

- **Boundedness from below.** The growth bounds from (3.5) imply that $\mathscr{F}[v] \geq -c_2\mathscr{L}^n(\Omega)$ for all $v \in \mathrm{W}^{1,p}(\Omega)$. Hence, in particular, $\mathscr{F}$ is bounded below on D_1. We denote $(u_j) \subset D_1$ a corresponding minimising sequence so that

$$\mathscr{F}[u_j] \to \inf_{v\in D_1} \mathscr{F}[v] \qquad \text{as } j \to \infty.$$

In particular, $(\mathscr{F}[u_j])$ is bounded.

- **Coercivity.** For each $j \in \mathbb{N}$, we may write $u_j = u_0 + v_j$ with $v_j \in \mathrm{W}_0^{1,p}(\Omega)$. Now, invoking the Poincaré inequalities (see Section 3.4 for a derivation), that is,

$$\|\varphi\|_{\mathrm{L}^p(\Omega)} \le c_{\mathrm{Poinc}} \|\nabla\varphi\|_{\mathrm{L}^p(\Omega)} \qquad \text{for all } \varphi \in \mathrm{W}_0^{1,p}(\Omega) \tag{3.14}$$

and repeatedly making use of the elementary inequality

$$|a+b|^p = 2^p \left|\frac{a+b}{2}\right|^p \overset{t\mapsto t^p \text{ convex}}{\le} 2^{p-1}(|a|^p + |b|^p) \qquad \text{for all } a, b \in \mathbb{R}^n, \tag{3.15}$$

we succesively obtain by the triangle inequality:

$$\begin{aligned}
\|u_j\|^p_{\mathrm{W}^{1,p}(\Omega)} &\le (\|v_j\|_{\mathrm{W}^{1,p}(\Omega)} + \|u_0\|_{\mathrm{W}^{1,p}(\Omega)})^p \\
&\qquad\qquad (\text{triangle inequality on } \mathrm{W}^{1,p}) \\
&\le 2^{p-1}(\|v_j\|^p_{\mathrm{W}^{1,p}(\Omega)} + \|u_0\|^p_{\mathrm{W}^{1,p}(\Omega)}) \\
&\qquad\qquad (\text{by } (3.15)) \\
&= 2^{p-1}((1 + c^p_{\mathrm{Poinc}})\|\nabla v_j\|^p_{\mathrm{L}^p(\Omega)} + \|u_0\|^p_{\mathrm{W}^{1,p}(\Omega)}) \\
&\qquad\qquad (\text{by } (3.14)) \\
&= 2^{p-1}((1 + c^p_{\mathrm{Poinc}})\|\nabla u_j - \nabla u_0\|^p_{\mathrm{L}^p(\Omega)} + \|u_0\|^p_{\mathrm{W}^{1,p}(\Omega)}) \\
&\le 2^{p-1}(2^{p-1}(1 + c^p_{\mathrm{Poinc}})\|\nabla u_j\|^p_{\mathrm{L}^p(\Omega)} + 2^{p-1}(2 + c^p_{\mathrm{Poinc}})\|u_0\|^p_{\mathrm{W}^{1,p}(\Omega)}) \\
&\qquad\qquad (\text{by } (3.15)) \\
&=: \mathfrak{c}(p,n)\|\nabla u_j\|^p_{\mathrm{L}^p(\Omega)} + \mathfrak{d}(p)\|u_0\|^p_{\mathrm{W}^{1,p}(\Omega)},
\end{aligned}$$

where the constants $\mathfrak{c}$ and $\mathfrak{d}$ are defined in the obvious manner. We then invoke the growth bound (3.5) to find

$$\mathfrak{c}(p,n)\|\nabla u_j\|^p_{\mathrm{L}^p(\Omega)} \le \frac{\mathfrak{c}(p,n)}{c_1}\Big(\int_\Omega F(\nabla u_j)\,\mathrm{d}x + c_2\mathscr{L}^n(\Omega)\Big).$$

In conclusion, we obtain

$$\|u_j\|^p_{\mathrm{W}^{1,p}(\Omega)} \le \frac{\mathfrak{c}(p,n)}{c_1}\Big(\int_\Omega F(\nabla u_j)\,\mathrm{d}x + c_2\mathscr{L}^n(\Omega)\Big) + \mathfrak{d}(p)\|u_0\|^p_{\mathrm{W}^{1,p}(\Omega)}. \tag{3.16}$$

Thus, the boundedness of $(\mathscr{F}[u_j])$ **implies the boundedness of** (u_j) for $\|\cdot\|_{\mathrm{W}^{1,p}(\Omega)}$. Hence, $\mathscr{F}$ is coercive on D_1 for the norm on $\mathrm{W}^{1,p}(\Omega)$.

- **Compactness.** By Remark 3.4 and since $1 < p < \infty$, $\mathrm{W}^{1,p}(\Omega)$ is reflexive. Thus, by Corollary 3.3, there exists $u \in \mathrm{W}^{1,p}(\Omega)$ such that $u_{j(k)} \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$ as $k \to \infty$.

- **Lower semicontinuity.** By Corollary 3.7, $\mathscr{F}$ is lower semicontinuous for weak convergence on $\mathrm{W}^{1,p}(\Omega)$.

We are thus done if we can show that $u \in \mathrm{W}^{1,p}(\Omega)$ in fact belongs to D_1. This is a consequence of the following abstract result[9]:

Lemma 3.9. Let $(X, \|\cdot\|)$ be a normed space and let $M \subset X$ be convex and closed (for the norm topology). Then M is weakly sequentially closed: If $(x_j) \subset M$ satisfies $x_j \rightharpoonup x$ as $j \to \infty$, then $x \in M$.

Proof. If we had $x \notin M$, then the Hahn-Banach separation theorem yields the existence of some $y' \in X'$ and $\alpha \in \mathbb{R}$ such that $y'(y) < \alpha < y'(x)$ for all $y \in M$. Inserting $y = x_j$ and passing to the limit $j \to \infty$, $x_j \rightharpoonup x$ yields the contradictory $y'(x) < y'(x)$. The proof is complete. □

The Dirichlet classes D_1 are affine, closed subspaces of $\mathrm{W}^{1,p}(\Omega)$ and thus, in particular, convex and (norm-)closed subsets of $\mathrm{W}^{1,p}(\Omega)$. Therefore, Lemma 3.9 yields $u \in D_1$. In consequence, all requirements of the direct method are fulfilled, and so Theorem 3.8 follows – i.e.,

$$\boxed{u \in D_1 \text{ is a minimiser.}}$$

The case of D_2 as sketched in the following has not been addressed in class. The functional $\mathscr{F}$ still is bounded on D_2, and we denote (u_j) a corresponding minimising sequence. In this situation, we employ the Poincaré inequality, type 2:

$$\|\varphi\|_{\mathrm{L}^p(\Omega)} \leq c\|\nabla\varphi\|_{\mathrm{L}^p(\Omega)} \qquad \text{for all } \varphi \in \mathrm{W}^{1,p}(\Omega) \cap \{v \in \mathrm{L}^1(\Omega)\colon \int_\Omega v \,\mathrm{d}x = 0\}.$$

Note that it is at this inequality that we need the connectedness of Ω and the higher regularity of $\partial\Omega$ – see the discussion in Chapter 3.4. In analogy with D_1 – see (3.16) – we obtain for two constants $\mathfrak{c}, \mathfrak{d} > 0$ that

$$\|u_j\|^p_{\mathrm{W}^{1,p}(\Omega)} \leq \mathfrak{c}\int_\Omega F(\nabla u_j)\,\mathrm{d}x + \mathfrak{d}\mathscr{L}^n(\Omega) \qquad \text{for all } j \in \mathbb{N}.$$

Then we may equally extract a weakly convergent subsequence $(u_{j(k)})$ and find $u \in \mathrm{W}^{1,p}(\Omega)$ with $u_{j(k)} \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$. The Neumann class D_2 is equally norm-closed and convex, so weakly closed; this could even be seen directly without appealing to Lemma 3.9. In consequence, we also obtain the existence of a minimiser in this case.

Up to now, we have not dealt with the full model functionals from (3.7) but only their gradient parts. We now sketch how the full model cases can be handled.

[9] Also see Theorem 9.18 in the lecture notes by Robert Denk.

Remark 3.10 (The full model functionals, $\lambda > 0$ in (3.7))**.** We work from (3.6). This inequality is certainly fulfilled for $q = p$. The **Sobolev embedding theorem** (which we assume here) asserts that we may even choose

$$q \text{ to be } \begin{cases} \text{any number in } [p, \frac{np}{n-p}] & \text{if } 1 \leq p < n, \\ \text{any number in } [p, \infty) & \text{if } p = n, \\ \text{any number in } [p, \infty] & \text{if } p > n. \end{cases}$$

Moreover, the **Rellich-Kondrachov theorem** asserts that the embedding $\mathrm{W}^{1,p}(\Omega)$ into $\mathrm{L}^p(\Omega)$ is compact (provided Ω is open and bounded with boundary of class C^1).

We now exemplarily argue for D_1. As was done for Theorem 3.8, we find by $\lambda > 0$ that $\mathscr{F}$ as in (3.7) is bounded below. We take a minimising sequence (u_j) of $\mathscr{F}$; by the same argument as above, (u_j) is bounded on $\mathrm{W}^{1,p}(\Omega)$. Extracting $u_{j(k)}$ as above with $u_{j(k)} \rightharpoonup u$ in $\mathrm{W}^{1,p}(\Omega)$, we only need to show that

$$\mathscr{F}[u] \leq \liminf_{k\to\infty} \mathscr{F}[u_{j(k)}].$$

By Corollary 3.7, it suffices to establish that

$$\int_\Omega |u - f|^q \, \mathrm{d}x \leq \liminf_{k\to\infty} \int_\Omega |u_{j(k)} - f|^q \, \mathrm{d}x. \tag{3.17}$$

By passing to a subsequence, we may assume that the lim inf on the right-hand side of (3.17) is a limit indeed. By Rellich-Kondrachov, we may assume that $u_{j(k)} \to v$ **strongly** in $\mathrm{L}^p(\Omega)$. Passing to another subsequence, we may thus assume that $u_{j(k(l))} \to u$ $\mathscr{L}^n$-a.e. as $l \to \infty$. Since $f \in \mathrm{L}^q(\Omega)$, we obtain (3.17) directly by Fatou's lemma, thereby completing the proof.

3.3 Weak*-compactness and vectorial Radon measures

The previous section gives us a satisfactory treatment of the existence of minima for convex p-growth functionals for $1 < p < \infty$. We now turn to $p = 1$.

If $p = 1$, then the chief obstruction is that $\mathrm{W}^{1,1}(\Omega)$ is **not reflexive**. So the Banach-Alagolu-Bourbaki theorem does not yield any compactness result. This might be regarded as the *functional analytic viewpoint* on the matter[10]; namely, if it were, any *bounded sequence* $(u_j) \subset \mathrm{W}^{1,1}(\Omega)$ would allow to extract a weakly convergent subsequence. This is easily seen to **not be the case** by adopting the *measure theoretic viewpoint*. *The following example is slightly different from that as given in class – both do the job.*

[10] Indeed, even more drastically, $\mathrm{W}^{1,1}$ is not even the dual of any normed space.

Example 3.11 (Concentration of bounded sequences in $\mathrm{W}^{1,1}$)**.** Consider the following sequence $(u_j) \in \mathrm{W}^{1,1}((-1,1))$ given by

$$u_j(x) := \begin{cases} 1 & \text{for } \frac{1}{j} < x < 1, \\ jx & \text{for } -\frac{1}{j} \leq x \leq \frac{1}{j}, \\ -1 & \text{for } -1 < x \leq \frac{1}{j}. \end{cases}$$

The weak gradient of u_j is given by

$$u_j'(x) := \begin{cases} 0 & \text{for } \frac{1}{j} < x < 1, \\ j & \text{for } -\frac{1}{j} \leq x \leq \frac{1}{j}, \\ 0 & \text{for } -1 < x \leq \frac{1}{j}. \end{cases}$$

We may compute explicitely that (u_j) is bounded in $\mathrm{W}^{1,1}((-1,1))$, which we leave as an exercise to the reader. Now, if $u_{j(k)} \rightharpoonup v \in \mathrm{W}^{1,1}((-1,1))$ in $\mathrm{W}^{1,1}((-1,1))$, then the same argument as in the proof of Corollary 3.7 would yield that $u'_{j(k)} \rightharpoonup v'$ in $\mathrm{L}^1((-1,1))$. Now let $\varphi \in \mathrm{C}_c^\infty((-1,1))$. Then

$$\begin{aligned} \int_{(-1,1)} v'\varphi \,\mathrm{d}x &= \lim_{k\to\infty} \int_{(-1,1)} u'_{j(k)}\varphi \,\mathrm{d}x \\ &= \frac{1}{2} \lim_{k\to\infty} 2j(k) \int_{-1/j(k)}^{1/j(k)} \varphi \,\mathrm{d}x = \frac{1}{2}\varphi(0), \end{aligned}$$

where we have used that

$$\frac{1}{2\varepsilon} \int_{x-\varepsilon}^{x+\varepsilon} \varphi(t)\,\mathrm{d}t \to \varphi(0) \qquad \text{as } \varepsilon \searrow 0.$$

Taking φ to be arbitrary but compactly supported off $x = 0$, $v' \equiv 0$ $\mathscr{L}^1$-a.e.. As in Example 2.7, we then arrive at a contradiction by considering $\varphi \in \mathrm{C}_c^\infty((-1,1))$ with $\varphi(0) \neq 0$.

Lecture 3, April 30, 2021

Lecture 4, May 07, 2021

Different from the context of Sobolev spaces with $1 < p < \infty$,

$$\begin{aligned} &(u_j) \text{ bounded in } \mathrm{W}^{1,1}(\Omega) \\ &\qquad \not\Rightarrow (\nabla u_j) \text{ has a weakly convergent subsequence in } \mathrm{L}^1(\Omega;\mathbb{R}^n). \end{aligned}$$

Thus, we wish to come up with a suitable space $\mathcal{X}$ such that

$$(u_j) \text{ bounded in } \mathrm{W}^{1,1}(\Omega)$$

$$\Rightarrow (\nabla u_j) \text{ has a convergent subsequence in } \mathcal{X} \text{ (in a suitable sense)}$$

This consideration will finally lead us to the space BV of functions of bounded variation, which consquently will be defined in terms of $\mathcal{X}$. In the following, we present some ideas of what such a space $\mathcal{X}$ could be.

We want to **gain compactness**. By the concentration displayed in Example 3.11 – recall that the weak*-limit is a multiple of the Dirac measure centered at the origin – a good guess is to work with measures. The key will be the Banach-Alaoglu-Bourbaki theorem (Theorem 3.2), for which we would like to *realise* measures as the dual of another (separable) space. Here we face some obstructions:

- First of all, the gradients of functions are $\mathbb{R}^n$-valued. If we wish to understand them – in some sense – as measures, we need a concept of $\mathbb{R}^n$-valued measures.
- Similarly, the Banach-Alaoglu-Bourbaki theorem forces us to work with vector spaces. The usual measures as known from previous courses do *not* form a vector space.
- Thirdly, if we wish to fruitfully employ the Banach-Alaoglu-Bourbaki theorem, we have to find a *norm* on the $\mathbb{R}^n$-valued measures.

We now introduce the requisite terminology and generalise the concepts from an introductory course on measure and integration theory.

Let (X, Σ) be a measurable space and $m \in \mathbb{N}$. We say that

- $\mu\colon \Sigma \to [0, \infty]$ is a **positive measure** provided it is a measure in the sense of an introductory course on measure theory and integration; this is, $\mu(\emptyset) = 0$ and we have

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i)$$

  for any sequence (A_i) of mutually disjoint elements contained in Σ.

- $\mu\colon \Sigma \to \mathbb{R}^m$ is an $\mathbb{R}^m$**-valued measure** (or **vector-valued measure** or simply **measure**) provided $\mu(\emptyset) = 0$ and

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i)$$

  for any sequence (A_i) of mutually disjoint elements contained in Σ. If $m = 1$, then $\mathbb{R}$-valued measures are also called **real measures**.

For an $\mathbb{R}^m$-valued measure μ as above, we define the associated **total variation measure** $|\mu|$ on Σ by

$$|\mu|(A) := \sup \left\{ \sum_{i=1}^{\infty} |\mu(A_i)| \colon \begin{array}{c} A_i \in \Sigma \text{ for all } i \in \mathbb{N}, \text{ mutually} \\ \text{disjoint with } A = \bigcup_{i=1}^{\infty} A_i \end{array} \right\} \tag{3.18}$$

Here, $|\cdot|$ is an arbitrary but fixed norm on $\mathbb{R}^m$. The particular choice of $|\cdot|$ is non-essential as all norms are equivalent on $\mathbb{R}^m$.

We now work towards the three items from above:

Lemma 3.12. Let (X, Σ) be a measurable space and let $\mu\colon \Sigma \to \mathbb{R}^m$ be a vector measure. Then the following hold:

(a) $|\mu|$ is a positive measure with $|\mu|(X) < \infty$.

(b) The space $\mathrm{M}(X;\mathbb{R}^m)$ of measures $\mu\colon \Sigma \to \mathbb{R}^m$ is a **normed vector space** when endowed with the **total variation norm** $\|\mu\| := |\mu|(\Omega)$.

Proof. We only establish that $|\mu|(X) < \infty$. Here, we may assume that $m = 1$; if $\mu = (\mu_1, ..., \mu_m)$ is an $\mathbb{R}^m$-valued measure, we use that $|\mu|(A) \leq \sum_{i=1}^m |\mu_i|(A)$.

Hence let $m = 1$ and suppose towards a contradiction that $|\mu|(X) = +\infty$. Working from the definition of the total variation measure, cf. (3.18), we then find a sequence (X_j) of mutually disjoint sets contained in Σ and some $N \in \mathbb{N}$ such that

$$\sum_{j=1}^{N} |\mu(X_j)| > 2(|\mu(X)| + 1).$$

This consequently yields

$$A + B := \Big(\sum_{\substack{1\leq j\leq N:\\ \mu(X_j)\geq 0}} \mu(X_j) \Big) + \Big(\sum_{\substack{1\leq j\leq N:\\ \mu(X_j)< 0}} (-\mu(X_j)) \Big) > 2(|\mu(X)| + 1).$$

Clearly, not both of A and B can be smaller than $|\mu(X)| + 1$. Without loss of generality, suppose that $A > |\mu(X)| + 1$. In this case, put

$$E := \bigcup_{\substack{1\leq j\leq N:\\ \mu(X_j)\geq 0}} X_j$$

(if $B > |\mu(X)| + 1$, let E be the union over those X_j with $1 \leq j \leq N$ and $\mu(X_j) < 0$). In any case, we obtain $E \in \Sigma$ with $|\mu(E)| > |\mu(X)| + 1$.

We now define $F := X \setminus E$. Since μ is an $\mathbb{R}^m$-valued measure, we conclude by the inverse triangle inequality:

$$|\mu(F)| = |\mu(X) - \mu(E)| = |\mu(E) - \mu(X)| \geq |\mu(E)| - |\mu(X)| > 1.$$

Since $|\mu|$ is a measure, too, we have $|\mu|(F) + |\mu|(E) = |\mu|(X) = \infty$. Hence, either $|\mu|(F) = \infty$ or $|\mu|(E) = \infty$ (or both). We distinguish two cases:

- If $|\mu|(F) = \infty$, we put $E_1 := E$. Replacing X by F in our above considerations, we then find E_2 and F_1 with $E_2 \cup F_1 = F$, $|\mu|(E_2) > 1$ and $|\mu|(F_1) = \infty$.

- If $|\mu|(E) = \infty$, put $E_1 := F$.

Proceeding in this way, we obtain a sequence (E_i) of mutually disjoint elements of Σ such that $|\mu(E_i)| > 1$ for all $i \in \mathbb{N}$. In consequence, the series $\sum_{i\in\mathbb{N}} \mu(E_i)$ cannot coverge. However, since μ is a measure and thus σ-additive,

$$\sum_{i\in\mathbb{N}} \mu(E_i) = \mu\Big(\bigcup_{i\in\mathbb{N}} E_i\Big) \in \mathbb{R}^m.$$

This is at variance with $\sum_{i\in\mathbb{N}} \mu(E_i)$ not being convergent, and the proof is complete. □

In the following, we need a variant of the Radon-Nikodým theorem for vectorial measures. To state it, let (X, Σ) be a measurable space and $\mu, \widetilde{\mu}$ be positive measures, and $\nu, \widetilde{\nu}$ be $\mathbb{R}^m$-valued measures on (X, Σ). We say that

- ν is **absolutely continuous** for μ (in formulas $\nu \ll \mu$) if we have

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 \qquad \text{for all } A \in \Sigma.$$

- $\mu, \widetilde{\mu}$ are **mutually singular** (in formulas $\mu \perp \widetilde{\mu}$) if there exists $E \in \Sigma$ such that

$$\mu(A) = \widetilde{\mu}(X \setminus A) = 0.$$

- $\nu, \widetilde{\nu}$ are mutually singular if $|\nu| \perp |\widetilde{\nu}|$, and ν, μ are mutually singular if $|\nu| \perp \mu$.

For the following proposition, recall that X is σ-finite for μ if there exists a sequence $A_1 \subset A_2 \subset ...$ of sets in Σ such that $\mu(A_i) < \infty$ for all $i \in \mathbb{N}$ and $X = \bigcup_{i\in\mathbb{N}} A_i$. We also define

$$\mathcal{L}^1_\mu(X; \mathbb{R}^m) := \{f = (f_1, ..., f_m) \colon \Omega \to \mathbb{R}^m \colon \int_X |f_i| \,\mathrm{d}\mu < \infty \text{ for all } i \in \{1, ..., m\}\},$$

and $\mathrm{L}^1_\mu(X; \mathbb{R}^m) := \mathcal{L}^1_\mu(X; \mathbb{R}^m)/_\sim$, $\sim$ being the usual equivalence relation identifying two maps that coincide μ-a.e..

Proposition 3.13 (Radon-Nikodým)**.** Let (X, Σ) be a measurable space, μ, ν be as above and let X be σ-finite for μ. Then there exists a uniquely determined pair ν^a, ν^s of $\mathbb{R}^m$-valued measures on Σ such that

$$\nu = \nu^a + \nu^s, \quad \nu^a \ll \mu \text{ and } \nu^s \perp \mu.$$

Moreover, there is a unique $f = (f_1, ..., f_m) \in \mathrm{L}^1_\mu(X; \mathbb{R}^m)$ such that $\nu^a = f\mu$, that is,

$$\nu^a(A) = (f\mu)(A) := \int_X (f_1, ..., f_m) \,\mathrm{d}\mu$$

$$:= \Big(\int_X f_1 \, \mathrm{d}\mu, \dots, \int_X f_m \, \mathrm{d}\mu \Big) \quad \text{for all } A \in \Sigma.$$

We refer to f as the **density** of ν^a with respect to μ, and also write

$$f = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}.$$

From an introductory course on measure and integration theory, we have a clear understanding of integrating functions against a *positive measure*. Based on the previous proposition, we now generalise this concept to integration with respect to vector-valued measures.

Remark 3.14. In the situation of the previous proposition, we **always** have

$$\nu \ll |\nu|. \tag{3.19}$$

With the notation from Proposition 3.13, we then have

$$\nu = \frac{\mathrm{d}\nu}{\mathrm{d}|\nu|} |\nu|.$$

Now, if $g \colon X \to \mathbb{R}$ or $h = (h_1, \dots, h_m) \colon X \to \mathbb{R}^m$ are Σ-$\mathscr{B}(\mathbb{R})$- or Σ-$\mathscr{B}(\mathbb{R}^m)$-measurable, respectively, we put

$$\begin{aligned} \int g \, \mathrm{d}\nu &:= \int g \frac{\mathrm{d}\nu}{\mathrm{d}|\nu|} \, \mathrm{d}|\nu|, \\ \int h \, \mathrm{d}\nu &:= \int \langle h, \mathrm{d}\nu \rangle := \int \Big\langle h, \frac{\mathrm{d}\nu}{\mathrm{d}|\nu|} \Big\rangle \, \mathrm{d}|\nu| \\ &= \sum_{i=1}^m \int h_i \Big(\frac{\mathrm{d}\nu}{\mathrm{d}|\nu|} \Big)_i \, \mathrm{d}|\nu|. \end{aligned} \tag{3.20}$$

Finally, if $\nu \ll \mu$ so that $\nu = f\mu$ for some $f \in \mathrm{L}^1_\mu(X; \mathbb{R}^m)$, the total variation measure of ν can be represented via

$$|\nu|(A) = \int_A |f| \, \mathrm{d}\mu, \qquad A \in \Sigma. \tag{3.21}$$

Now, having a clear definition of vector-valued measure and the underlying integration theory, we introduce the main concepts that turn out crucial for the requisite compactness assertions. In a first step, we generalise the concept of Radon measures (cf. Example 2.7) to the vectorial context:

Definition 3.15 (Vectorial Radon measures)**.** Let $\Omega \subset \mathbb{R}^n$ be open. An $\mathbb{R}^m$-valued set function defined on the relatively compact Borel subsets of Ω, which is a measure on $(K, \mathscr{B}(K))$ for any compact $K \subset \Omega$, is called a **real or vectorial Radon measure** (on Ω). If, moreover, $\mu\colon \mathscr{B}(\Omega) \to \mathbb{R}^m$ is a measure, we call μ a **finite Radon measure**. We put

$$\begin{aligned} \mathrm{RM}(\Omega; \mathbb{R}^m) &:= \{\mu \text{ is an } \mathbb{R}^m\text{-valued Radon measure on } \Omega\}, \\ \mathrm{RM}_{\mathrm{fin}}(\Omega; \mathbb{R}^m) &:= \{\mu \text{ is a finite } \mathbb{R}^m\text{-valued Radon measure on } \Omega\}. \end{aligned}$$

Theorem 3.2 gives us *some* compactness in the space of finite $\mathbb{R}^m$-valued Radon measures provided the latter space is a dual space (of some separable space). The following theorem a lá Riesz precisely allows this conclusion:

Theorem 3.16 (Riesz)**.** Let $\Omega \subset \mathbb{R}^n$ be open and let $\Phi\colon \mathrm{C}_0(\Omega; \mathbb{R}^m) \to \mathbb{R}$ be a linear and bounded functional, so

$$\|\Phi\| := \sup\{|\Phi(\varphi)|:\ \varphi \in \mathrm{C}_0(\Omega; \mathbb{R}^m),\ |\varphi| \leq 1\} < \infty.$$

Then there exists a uniquely determined finite, $\mathbb{R}^n$-valued Radon measure $\mu = (\mu_1, ..., \mu_n) \in \mathrm{RM}_{\mathrm{fin}}(\Omega; \mathbb{R}^n)$ such that

$$\Phi(\varphi) = \int_\Omega \varphi \,\mathrm{d}\mu := \sum_{i=1}^m \int_\Omega \varphi_i \,\mathrm{d}\mu_i$$

holds for all $\varphi = (\varphi_1, ..., \varphi_m) \in \mathrm{C}_0(\Omega; \mathbb{R}^m)$. We then moreover have

$$\|\Phi\| = |\mu|(\Omega). \tag{3.22}$$

In other words, the finite $\mathbb{R}^m$-valued Radon measures on Ω are nothing but the dual of the $\mathbb{R}^m$-valued continuous functions vanishing at the boundary:

$$(\mathrm{RM}_{\mathrm{fin}}(\Omega; \mathbb{R}^m), |\cdot|(\Omega)) \cong (\mathrm{C}_0(\Omega), \mathbb{R}^m), \|\cdot\|_{\sup})'. \tag{3.23}$$

Let us now return to the situation that we addressed in the main part of the chapter. Namely, let $(u_j) \subset \mathrm{W}^{1,1}(\Omega)$ be a sequence which is bounded in $\mathrm{W}^{1,1}(\Omega)$ – in particular, the sequence of gradients (∇u_j) is bounded in $\mathrm{L}^1(\Omega; \mathbb{R}^n)$, so $\sup_{j\in\mathbb{N}} \|\nabla u_j\|_{\mathrm{L}^1(\Omega)} \leq C < \infty$.

We define measures $\mu_j \in \mathrm{RM}_{\mathrm{fin}}(\Omega; \mathbb{R}^m)$ via $\mu_j = \nabla u_j \mathscr{L}^n$, i.e.,

$$\mu_j(A) = \int_A \nabla u_j \,\mathrm{d}x \qquad \text{for all } A \in \mathscr{B}(\Omega).$$

By Remark 3.14 and our assumption $\sup_{j\in\mathbb{N}}\|\nabla u_j\|_{\mathrm{L}^1(\Omega)} \leq C < \infty$, we then have

$$|\mu_j|(\Omega) = \int_\Omega |\nabla u_j|\,\mathrm{d}x \leq C < \infty.$$

By Theorems 3.16 and 3.2[11], there is $\mu \in \mathrm{RM}_{\mathrm{fin}}(\Omega;\mathbb{R}^n)$ and a subsequence $(\mu_{j(k)}) \subset (\mu_j)$ such that

$$\mu_{j(k)} \overset{*}{\rightharpoonup} \mu \qquad \text{as } k \to \infty.$$

This last line means nothing but

$$\int_\Omega \varphi\,\mathrm{d}\mu_{j(k)} \to \int_\Omega \varphi\,\mathrm{d}\mu \qquad \text{for } k \to \infty,$$

and recalling the definition of $\mu_{j(k)}$, $\mu_{j(k)} = \nabla u_j \mathscr{L}^n$, we have

$$\int_\Omega \langle \varphi, \nabla u_{j(k)}\rangle\,\mathrm{d}x \to \int_\Omega \varphi\,\mathrm{d}\mu \qquad \text{for } k \to \infty.$$

Therefore, **if we assume** that $u_{j(k)} \to u$ e.g. in $\mathrm{L}^1(\Omega)$, we obtain

$$\begin{aligned}\int_\Omega u\,\mathrm{div}(\varphi)\,\mathrm{d}x &= \lim_{k\to\infty}\int_\Omega u_{j(k)}\,\mathrm{div}(\varphi)\,\mathrm{d}x \\ &= -\lim_{k\to\infty}\int_\Omega \langle\varphi, \nabla u_{j(k)}\rangle\,\mathrm{d}x = -\int_\Omega \varphi\,\mathrm{d}\mu.\end{aligned} \tag{3.24}$$

Generalising the concept of measure-regular distributions from Example 2.7 to the situation at our disposal, we see that the ultimate equation just tells us that μ is the (measure regular representative of the) distributional gradient of u; for more detail, also see Remark 3.17 below.

We may summarise the above discussion by means of the following.

The moral: If we want to get compactness for sequences bounded in $\mathrm{W}^{1,1}$, then we should work those L^1-functions for which the **distributional gradients are finite $\mathbb{R}^n$-valued measures**.

The following remark had not been addressed in the lecture and is only intended to clarify the discussion after (3.24):

[11] Note that $\mathrm{C}_0(\Omega;\mathbb{R}^m)$ is a separable Banach space when endowed with the supremum norm. Recall the argument: We may regard $\mathrm{C}_0(\Omega;\mathbb{R}^m)$ as a subset of $\mathrm{C}(\overline{\Omega};\mathbb{R}^m)$, and the latter space is separable by the Weierstraß approximation theorem – the $\mathbb{R}^m$-valued polynomials with rational coefficients are dense.

Remark 3.17. In Definition 2.8, we defined the distributional gradient of some $T \in \mathscr{D}(\Omega)$ as $(\partial_1 T, ..., \partial_n T) \in (\mathscr{D}'(\Omega))^n$, where $\partial_i T$ is the i-th ditributional partial derivative of T. Equivalently – and this is precisely what underlies (3.24) – an element $S = (S_1, ..., S_n) \in (\mathscr{D}'(\Omega))^n$ is the distributional gradient of u (or, more precisely, T_u) if and only if

$$\int_\Omega u\mathrm{div}(\varphi)\,\mathrm{d}x = -S(\varphi) := -\sum_{i=1}^n S_i(\varphi_i)$$

for all $\varphi = (\varphi_1, ..., \varphi_n) \in (\mathscr{D}(\Omega))^n$.

Above, we also called μ the *measure regular representative* of u (or, more precisely, T_u). This is analogous to Example 2.7, every real Radon measure μ on Ω induces a distribution via

$$T_u(\varphi) := \int_\Omega \varphi\,\mathrm{d}\mu, \qquad \varphi \in \mathrm{C}_c^\infty(\Omega).$$

Similarly as above, if $T \in \mathscr{D}'(\Omega)$ satisfies $T = T_\mu$ for some real Radon measure μ, then we also call T *measure regular*. This terminology directly inherits to $\mathbb{R}^n$-valued Radon measures, now defined on $(\mathscr{D}'(\Omega))^n$.

3.4 Recap on some facts from Sobolev space theory

Poincaré inequalities allow to estimate the L^p-norm of a weakly differentiable function by that of its gradient. We used them in the proof of the main existence result of this chapter, Theorem 3.8, and now give a brief recap.

Poincaré-type inequalities, type 1: Let $\Omega \subset \mathbb{R}^n$ be open and bounded. Then, for each $1 \leq p < \infty$, there exists a constant $c = c(p, n, \mathrm{diam}(\Omega)) > 0$ such that

$$\|u\|_{\mathrm{L}^p(\Omega)} \leq c\,\|\nabla u\|_{\mathrm{L}^p(\Omega)} \qquad \text{holds for all } u \in \mathrm{W}_0^{1,p}(\Omega). \tag{3.25}$$

First of all, note carefully that inequality (3.25) does not hold for all Sobolev functions $u \in \mathrm{W}^{1,p}(\Omega)$ (note our requirement $u \in \mathrm{W}_0^{1,p}(\Omega)$). Indeed, (3.25) is false for any non-zero, constant function – such functions are precisely ruled out by $u \in \mathrm{W}_0^{1,p}(\Omega)$.

Let us briefly see how inequalities (3.25) can be derived in $n = 1$ dimensions. To this end, let $\Omega \subset \mathbb{R}$ be open and bounded. Given $u \in \mathrm{C}_c^\infty(\Omega)$, we think of u to be extended to the entire $\mathbb{R}$ by zero and suppose that $\Omega = (a, b)$ with $-\infty < a < b < \infty$. By the fundamental theorem of calculus, we obtain for any $x \in \Omega$

$$u(x) = u(x) - \underbrace{u(a)}_{=0} = \int_a^x u'(t)\,\mathrm{d}t,$$

and so, by Hölder's inequality,

$$\int_\Omega |u(t)|^p\,\mathrm{d}t = \int_a^b |u(x)|^p\,\mathrm{d}x \le \int_a^b \left|\int_a^x u'(t)\,\mathrm{d}t\right|^p \mathrm{d}x$$
$$\le (b-a)^p \int_a^b |u'(t)|^p\,\mathrm{d}t = (b-a)^p \int_\Omega |u'(t)|^p\,\mathrm{d}t.$$

Now let $u \in \mathrm{W}_0^{1,p}(\Omega)$. By definition of $\mathrm{W}_0^{1,p}(\Omega)$, there exists $(u_j) \subset \mathrm{C}_c^\infty(\Omega)$ such that $\|u - u_j\|_{\mathrm{W}^{1,p}(\Omega)} \to 0$ as $j \to \infty$. Then, based on what we have already established,

$$\begin{aligned}
\|u\|_{\mathrm{L}^p(\Omega)} &\le \|u - u_j\|_{\mathrm{L}^p(\Omega)} + \|u_j\|_{\mathrm{L}^p(\Omega)} \\
&\le \|u - u_j\|_{\mathrm{L}^p(\Omega)} + c\|u_j'\|_{\mathrm{L}^p(\Omega)} \\
&\le \|u - u_j\|_{\mathrm{L}^p(\Omega)} + c\|u_j' - u'\|_{\mathrm{L}^p(\Omega)} + c\|u'\|_{\mathrm{L}^p(\Omega)} \\
&\to c\|u'\|_{\mathrm{L}^p(\Omega)}, \qquad j \to \infty.
\end{aligned}$$

This establishes the claim for $n = 1$; the general case can be reduced to the one-dimensional case by Fubini's theorem.

Poincaré-type inequalities, type 2: Let $\Omega \subset \mathbb{R}^n$ be connected, open and bounded with boundary of class C^1. Then, for each $1 \le p < \infty$, there exists a constant $c = c(p, n, \mathrm{diam}(\Omega)) > 0$ such that

$$\|u - (u)_\Omega\|_{\mathrm{L}^p(\Omega)} \le c\,\|\nabla u\|_{\mathrm{L}^p(\Omega)} \qquad \text{holds for all } u \in \mathrm{W}_0^{1,p}(\Omega), \tag{3.26}$$

where

$$(u)_\Omega := \fint_\Omega u\,\mathrm{d}x := \frac{1}{\mathscr{L}^n(\Omega)} \int_\Omega u\,\mathrm{d}x$$

is the **mean of u over** Ω.

A sketch of proof for these sorts of inequalities will be inserted soon.

4 Functions of bounded variation: Definition & elementary properties

In Chapter 3 we saw the crucial need of *compactness* to ensure the existence of minima by means of the direct method. Since compactness displays a major issue for p-growth functionals with $p = 1$, we argued in Section 3.3 to study functions for which the distributional gradients are finite, vectorial Radon measures.

This is precisely the space BV of *functions of bounded variation*. In this chapter, we thoroughly introduce this class of functions and study their various properties. As a solid basis, we start with the following

Theorem and Definition 4.1 (Functions of bounded variation)**.** Let $\Omega \subset \mathbb{R}^n$ be open. Then the following are equivalent for $u \in \mathrm{L}^1(\Omega)$:

(a) The distributional gradient of u (or, more precisely, T_u) is a measure regular distribution in the sense that there exists a finite, $\mathbb{R}^n$-valued Radon measure $\mu = (\mu_1, ..., \mu_n) \in \mathrm{RM}_{\mathrm{fin}}(\Omega; \mathbb{R}^n)$ such that

$$\int_\Omega u \operatorname{div}(\varphi)\,\mathrm{d}x = -\int_\Omega \varphi\,\mathrm{d}\mu\Big(:= \sum_{i=1}^n \int_\Omega \varphi_i\,\mathrm{d}\mu_i\Big)$$

holds for all $\varphi = (\varphi_1, ..., \varphi_n) \in (\mathrm{C}_c^\infty(\Omega))^n$.

(b) The **total variation** $|Du|(\Omega)$ satisfies

$$|Du|(\Omega) := \sup\left\{\int_\Omega u \operatorname{div}(\varphi)\,\mathrm{d}x\colon\ \varphi \in \mathrm{C}_c^\infty(\Omega; \mathbb{R}^n),\ |\varphi| \le 1\right\} < \infty.$$

If (a) or, equivalently, (b) hold, then we say that u is **of bounded variation**, and we put

$$\mathrm{BV}(\Omega) := \{u \in \mathrm{L}^1(\Omega)\colon\ |Du|(\Omega) < \infty\}. \tag{4.1}$$

Lecture 4, May 07, 2021

Lecture 5, May 14, 2021

Proof of Theorem 4.1. Ad '(a)⇒(b)'. Let $\varphi = (\varphi_1, ..., \varphi_n) \in \mathrm{C}_c^\infty(\Omega; \mathbb{R}^n)$ be such that $|\varphi| \le 1$. By (b), we have

$$\begin{aligned}\int_\Omega u \operatorname{div}(\varphi)\,\mathrm{d}x &= -\int_\Omega \varphi\,\mathrm{d}\mu \\ &\le \left|\int_\Omega \varphi\,\mathrm{d}\mu\right| = \left|\int_\Omega \left\langle \varphi, \frac{\mathrm{d}\mu}{\mathrm{d}|\mu|}\right\rangle \mathrm{d}|\mu|\right| \le \int_\Omega |\varphi|\,\mathrm{d}|\mu| \le |\mu|(\Omega).\end{aligned}$$

Passing to the supremum over all such φ, we obtain $|Du|(\Omega) < \infty$, which is (b). Ad '(b)⇒(a)'. Consider the linear functional

$$\Phi(\varphi) := -\int_\Omega u \operatorname{div}(\varphi)\,\mathrm{d}x, \qquad \varphi = (\varphi_1, ..., \varphi_n) \in \mathrm{C}_c^\infty(\Omega;\mathbb{R}^n).$$

By (b), Φ is bounded as a functional on $(\mathrm{C}_c^\infty(\Omega;\mathbb{R}^n), \|\cdot\|_{\sup})$, where

$$\|\varphi\|_{\sup} := \sum_{i=1}^n \|\varphi_i\|_{\sup}, \qquad \varphi = (\varphi_1, ..., \varphi_n).$$

Since $\mathrm{C}_c^\infty(\Omega;\mathbb{R}^n)$ is dense in $\mathrm{C}_0(\Omega;\mathbb{R}^n)$ for $\|\cdot\|_{\sup}$, Φ uniquely extends to a bounded linear functional $\overline{\Phi}\colon \mathrm{C}_0(\Omega;\mathbb{R}^n) \to \mathbb{R}$ with $\|\overline{\Phi}\| = \|\Phi\|$. By the Riesz representation theorem, Theorem 3.16, there exists a uniquely determined finite, $\mathbb{R}^n$-valued Radon measure μ on Ω such that

$$\overline{\Phi}(\varphi) = \int_\Omega \varphi\,\mathrm{d}\mu \qquad \text{for all } \varphi \in \mathrm{C}_0(\Omega;\mathbb{R}^n).$$

Since Φ and $\overline{\Phi}$ coincide on $\mathrm{C}_c^\infty(\Omega;\mathbb{R}^n)$, we obtain by the definition of Φ

$$-\int_\Omega u \operatorname{div}(\varphi)\,\mathrm{d}x = \int_\Omega \varphi\,\mathrm{d}\mu \qquad \text{for all } \varphi \in \mathrm{C}_c^\infty(\Omega;\mathbb{R}^n).$$

This is (a), and the proof is complete. □

Remark 4.2. With $|Du|(\Omega)$ as in item (b) from above, we may equally write

$$|Du|(\Omega) = \sup\left\{\int_\Omega u \operatorname{div}(\varphi)\,\mathrm{d}x\colon\ \varphi \in \mathrm{C}_c^1(\Omega;\mathbb{R}^n),\ |\varphi| \le 1\right\} \tag{4.2}$$

(note the difference in the smoothness of admissible competitors). The right-hand side of (4.2) is certainly larger or equal than $|Du|(\Omega)$. Now let $\varphi \in \mathrm{C}_c^1(\Omega;\mathbb{R}^n)$ be arbitrary with $|\varphi| \le 1$. Pick a standard mollifier ρ and its ε-rescaled variant ρ_ε. We then have $\operatorname{div}(\rho_\varepsilon * \varphi) = \rho_\varepsilon * \operatorname{div}(\varphi) \to \operatorname{div}(\varphi)$ in $\mathrm{C}_0(\Omega;\mathbb{R}^n)$ as $\varepsilon \searrow 0$ (as φ is C_c^1 and so $\operatorname{div}(\varphi)$ is C_c). Note that

$$\left|\int_\Omega u \operatorname{div}(\varphi - \rho_\varepsilon * \varphi)\,\mathrm{d}x\right| \le \|u\|_{\mathrm{L}^1(\Omega)} \|\operatorname{div}(\varphi) - \rho_\varepsilon * \operatorname{div}(\varphi)\|_{\mathrm{L}^\infty} \to 0 \tag{4.3}$$

as $\varepsilon \searrow 0$. Since, moreover, $|\rho_\varepsilon * \varphi| \le 1$, we obtain

$$\int_\Omega u\operatorname{div}(\varphi)\,\mathrm{d}x = \underbrace{\int_\Omega u\operatorname{div}(\varphi - \rho_\varepsilon * \varphi)\,\mathrm{d}x}_{\to 0 \text{ by } (4.3)} + \underbrace{\int_\Omega u\operatorname{div}(\rho_\varepsilon * \varphi)\,\mathrm{d}x}_{\le |Du|(\Omega)}$$

and so (4.2) follows.

Now let $u \in \mathrm{BV}(\Omega)$. By the Radon-Nikodým theorem, Proposition 3.13, we may decompose Du into an absolutely continuous and singular part for $\mathscr{L}^n$:

$$\begin{aligned} Du &= D^a u + D^s u \\ &= \frac{\mathrm{d}D^a u}{\mathrm{d}\mathscr{L}^n}\mathscr{L}^n + \frac{\mathrm{d}D^s u}{\mathrm{d}|D^s u|}|D^s u|. \end{aligned}$$

A good deal of the present chapter will be to get a deeper insight on the single parts $D^a u$ and $D^s u$. We now turn to some examples, and start with the case where the singular part is not present:

Remark 4.3. Let $u \in \mathrm{BV}(\Omega)$. Then we have

$$u \in \mathrm{W}^{1,1}(\Omega) \Leftrightarrow D^s u = 0.$$

This is a direct consequence of Theorem and Definition 4.1.

Motivated by the questions of the preceding chapters, we first verify that BV-functions are indeed allowed to have jumps. We thereby obtain a first instance of a BV-function that does not belong to $\mathrm{W}^{1,1}$:

Example 4.4 (Jumps)**.** Let $n \geq 2$ and consider the function

$$u := \mathbb{1}_{\mathrm{B}_1(0)\cap\{x_n>0\}}\colon \mathrm{B}_1(0) \to \mathbb{R}$$

with the open unit ball $\mathrm{B}_1(0)$ in $\mathbb{R}^n$. Let $\varphi = (\varphi_1, ..., \varphi_n) \in \mathrm{C}_c^\infty(\mathrm{B}_1(0); \mathbb{R}^n)$. Using the Gauß-Green theorem (i.e., integration by parts), we then obtain

$$\begin{aligned} \int_{\mathrm{B}_1(0)} u \operatorname{div}(\varphi)\,\mathrm{d}x &= \int_{\mathrm{B}_1(0)\cap\{x_n>0\}} \operatorname{div}(\varphi)\,\mathrm{d}x \\ &= \int_{\partial\mathrm{B}_1(0)\cap\{x_n>0\}} \langle\varphi,\nu\rangle\,\mathrm{d}\mathscr{H}^{n-1} \\ &= -\int_{\{x=(x',x_n)\colon\ |x'|<1,\,x_n=0\}} \langle\varphi,e_n\rangle\,\mathrm{d}\mathscr{H}^{n-1}. \end{aligned}$$

Denoting the restriction of a measure μ to some A by $\mu \llcorner A$, we thus obtain

$$Du = e_n\mathscr{H}^{n-1} \llcorner \{x = (x', x_n)\colon\ |x'| < 1,\, x_n = 0\}$$

and so $u \in \mathrm{BV}(\mathrm{B}_1(0))$. Since the Hausdorff measure restricted to $\{x = (x', x_n)\colon\ |x'| < 1, x_n = 0\}$ is not absolutely continous for $\mathscr{L}^n$, $u \in (\mathrm{BV}\setminus\mathrm{W}^{1,1})(\mathrm{B}_1(0))$.

The previous example gives an instance of a *jump discontinuity*, and in this case, $D^s u$ does not vanish. However, $D^s u$ does not need to vanish when u

has no jump discontinuities. This can be seen by the Cantor ternary function, which we only discuss; the proof of the corresponding properties will be added soon.

Example 4.5 (The Cantor ternary function)**.** We define the **Cantor ternary function** as follows: For $x \in [0,1]$, we

- express x in base 3, i.e., $x = 0.c_1c_2c_3c_4...$, where

$$x = \sum_{j=1}^{\infty} c_j 3^{-j}, \qquad c_j \in \{0,1,2\}.$$

- If some digit c_j equals 1, then we replace any digit after the first 1 by 0.
- We replace all remaining coefficients 2 by 1.
- Finally, we interpret the number obtained in this way as a binary number – the outcome of this procedure is the Cantor ternary function $c(x)$.

Let us exmplarily compute two values of c:

- If $x \in [\frac{1}{3}, \frac{2}{3})$, then $x = 1 \cdot 3^{-1} + 0 \cdot 3^{-1} + 0 \cdot 3^{-2} + ...$ We then directly pass to the base-2-interpretation and obtain $c(x) = 1 \cdot 2^{-1} + 0 \cdot 2^{-2} + ...$, so $c(x) = \frac{1}{2}$.
- If $x \in [\frac{1}{9}, \frac{2}{9})$, then $x = 0 \cdot 3^{-1} + 1 \cdot 3^{-2} + 0 \cdot 3^{-3} + 0 \cdot 3^{-4} + ...$ We then form $0 \cdot 2^{-1} + 1 \cdot 2^{-2} + 0 \cdot 3^{-3} + 0 \cdot 3^{-4} + ...$, and so the value of $c(x)$ is $\frac{1}{4}$.

The graph of c is depicted in Figure 6. Based on the above definition, one is in position to show that c is continuous and increasing; in one dimension, this already implies that $c \in \mathrm{BV}((0,1))$; moreover, one can establish that $c \notin \mathrm{W}^{1,1}((0,1))$. However, by continuity, c cannot have jump discontinuities.

In conclusion, the singular part might not only stem from jump discontinuities but also parts which reveal some Cantor-type behaviour. The **aim of the chapter** thus is to

get a good understanding of the above gradient decomposition

and the corresponding densities.

4.1 Notions of convergence and approximation by smooth functions

In this subsection we present three sorts of convergence on $\mathrm{BV}(\Omega)$ and study both their advantages and disadvantages in view of **smooth approximation** and **compactness**. This particularly comprises

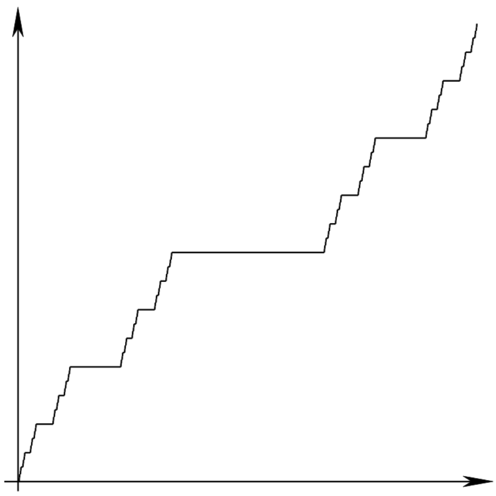

Fig. 6: A plot of the Cantor ternary function from Example 4.5.

- the norm convergence,
- the strict convergence,
- the weak*-convergence.

As an important auxiliary tool, we require the following proposition.

Proposition 4.6 (Lower semicontinuity of the total variation)**.** Let $\Omega \subset \mathbb{R}^n$ be open and suppose that $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ and $u_1, u_2, \ldots \in \mathrm{BV}(\Omega)$ are such that $u_j \to u$ in $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$. Then there holds

$$|Du|(\Omega) \leq \liminf_{j \to \infty} |Du_j|(\Omega). \tag{4.4}$$

Proof. Let $\varphi \in \mathrm{C}^\infty_c(\Omega; \mathbb{R}^n)$ be arbitrary with $|\varphi| \leq 1$. We then have

$$\begin{aligned} \int_\Omega u \operatorname{div}(\varphi) \,\mathrm{d}x &= \lim_{j \to \infty} \int_\Omega u_j \operatorname{div}(\varphi) \,\mathrm{d}x \\ &= - \lim_{j \to \infty} \int_\Omega \varphi \,\mathrm{d}Du_j \\ &\leq \liminf_{j \to \infty} |Du_j|(\Omega). \end{aligned}$$

We then pass to the supremum over all such φ to conclude. □

Norm convergence

Lemma 4.7. Let $\Omega \subset \mathbb{R}^n$ be open and define $\|u\|_{\mathrm{BV}(\Omega)} := \|u\|_{\mathrm{L}^1(\Omega)} + |Du|(\Omega)$. Then the following hold:

(a) $(\mathrm{BV}(\Omega), \|\cdot\|_{\mathrm{BV}(\Omega)})$ is Banach space.

(b) $(\mathrm{C}^\infty \cap \mathrm{BV})(\Omega)$ is **not dense** in $\mathrm{BV}(\Omega)$ for $\|\cdot\|_{\mathrm{BV}(\Omega)}$.

Proof. Let $(u_j) \subset \mathrm{BV}(\Omega)$ be a Cauchy sequence for $\|\cdot\|_{\mathrm{BV}(\Omega)}$. Then, since $\mathrm{L}^1(\Omega)$ is complete, there exists $u \in \mathrm{L}^1(\Omega)$ such that $u_j \to u$ in $\mathrm{L}^1(\Omega)$. By Proposition 4.6, $u \in \mathrm{BV}(\Omega)$. Moreover, for any fixed j,

$$|D(u - u_j)|(\Omega) \leq \lim_{m\to\infty} |D(u_m - u_j)|(\Omega) \to 0, \qquad j \to \infty,$$

and so the Banach space property follows. As a consequence of Example ??, for every open $\Omega \subset \mathbb{R}^n$, there exists $u \in (\mathrm{BV} \setminus \mathrm{W}^{1,1})(\Omega)$. Now, if $(\mathrm{C}^\infty \cap \mathrm{BV})(\Omega)$ were dense in $\mathrm{BV}(\Omega)$, we would find a sequence $(u_j) \subset (\mathrm{C}^\infty \cap \mathrm{BV})(\Omega)$ such that $\|u_j - u\|_{\mathrm{BV}(\Omega)} \to 0$ as $j \to \infty$. In consequence, (u_j) is Cauchy for $\|\cdot\|_{\mathrm{BV}(\Omega)}$. Now note that on $\mathrm{C}^\infty \cap \mathrm{BV} = \mathrm{C}^\infty \cap \mathrm{W}^{1,1}$, the norms $\|\cdot\|_{\mathrm{W}^{1,1}(\Omega)}$ and $\|\cdot\|_{\mathrm{BV}(\Omega)}$ coincide. Thus (u_j) is already Cauchy in $\mathrm{W}^{1,1}(\Omega)$, and since $(\mathrm{W}^{1,1}(\Omega), \|\cdot\|_{\mathrm{W}^{1,1}(\Omega)})$ is complete, $u_j \to v \in \mathrm{W}^{1,1}(\Omega)$. Now,

$$\begin{aligned}\int_\Omega \varphi \,\mathrm{d}Du &= -\int_\Omega u \,\mathrm{div}(\varphi)\,\mathrm{d}x = -\lim_{j\to\infty}\int_\Omega u_j \,\mathrm{div}(\varphi)\,\mathrm{d}x \\ &= -\int_\Omega v\,\mathrm{div}(\varphi)\,\mathrm{d}x = \int_\Omega \langle \varphi, \nabla v\rangle\,\mathrm{d}x\end{aligned}$$

for all $\varphi \in \mathrm{C}_c^\infty(\Omega;\mathbb{R}^n)$ and so

$$Du = \nabla v \mathscr{L}^n, \quad \text{so} \quad Du \ll \mathscr{L}^n.$$

This is a contradiction to $u \in (\mathrm{BV} \setminus \mathrm{W}^{1,1})(\Omega)$, and the proof is complete. ☐

↑ Lecture 5, May 14, 2021

↓ Lecture 6, May 21, 2021

Strict convergence

Definition 4.8. Let $\Omega \subset \mathbb{R}^n$ be open. The **strict metric** on $\mathrm{BV}(\Omega)$ is defined as

$$d_s(u,v) := \|u - v\|_{\mathrm{L}^1(\Omega)} + \big|\,|Du|(\Omega) - |Dv|(\Omega)\big|, \qquad u, v \in \mathrm{BV}(\Omega).$$

Theorem 4.9. Let $\Omega \subset \mathbb{R}^n$ be open. Then the following hold:

(a) $(\mathrm{BV}(\Omega), d_s)$ is a complete metric space.

(b) The metric d_s is not translation-invariant. In particular, there is no norm $\|\cdot\|$ on $\mathrm{BV}(\Omega)$ such that $d_s(u,v) = \|u-v\|$ for all $u, v \in \mathrm{BV}(\Omega)$.

(c) $(\mathrm{C}^\infty \cap \mathrm{BV})(\Omega)$ is dense in $\mathrm{BV}(\Omega)$ for d_s.

Proof. Ad (b). Translation invariance of a metric $d\colon X \to X \to \mathbb{R}_{\geq 0}$ defined on a vector space X means that $d(u+z, v+z) = d(u,v)$ holds for all $u, v, z \in X$. In particular, if $u_j \to u$ for a translation-invariant metric, then $u_j - u \to 0$, which can be seen by $d(u_j - u, 0) = d(u_j, u) \to 0$. Let $u \in (\mathrm{BV} \setminus \mathrm{W}^{1,1})(\Omega)$ and choose, due to (c), a sequence $(u_j) \subset (\mathrm{C}^\infty \cap \mathrm{BV})(\Omega)$ such that $d_s(u, u_j) \to 0$. If d_s were translation-invariant, then $d_s(u_j - u) \to 0$. This, in turn, is equivalent $\|u_j - u\|_{\mathrm{BV}(\Omega)} \to 0$ which we know to be impossible by Lemma 4.7.

The key to almost all smooth approximation results is mollification. Here we do not necessarily work on the entire $\mathbb{R}^n$ and thus follow the usual scheme to *firstly localise, secondly mollify* and *finally patch together*. To this end, let $\varepsilon > 0$ be arbitrary.

Step 1. Constructing the smooth approximation. We choose $m \in \mathbb{N}$ so large such that with

$$\Omega_k := \Big\{x \in \Omega\colon \ \operatorname{dist}(x, \partial\Omega) > \frac{1}{m+k}\Big\} \cap \mathrm{B}(0, m+k)$$

there holds

$$|Du|(\Omega \setminus \Omega_1) < \frac{\varepsilon}{4}. \tag{4.5}$$

We then put $U_0 := \varnothing$ and inductively define $U_k := \Omega_{k+1} \setminus \overline{\Omega_{k-1}}$ for $k \in \mathbb{N}_{\geq 1}$. For future reference, we remark that by construction, at most three U_k's overlap each.

In a next step, let (ρ_k) be a partition of unity subordinate to $(U_k)_{k \in \mathbb{N}}$. By this we understand that

(P1) $\rho_k \in \mathrm{C}_c^\infty(U_k; [0,1])$ for all $k \in \mathbb{N}$ and

(P2) $\sum_k \rho_k \equiv 1$ in Ω.

For each $k \in \mathbb{N}$, we pick $\varepsilon_k \in (0,1)$ such that $\operatorname{spt}(\varphi_{\varepsilon_k} * (\rho_k u)) \subset U_k$,

$$\int_\Omega |\varphi_{\varepsilon_k} * (\rho_k u) - \rho_k u| \,\mathrm{d}x < \frac{\varepsilon}{2^{k+2}}, \tag{4.6}$$

together with

$$\int_\Omega |\varphi_{\varepsilon_k} * (u \otimes \nabla \rho_k) - (u \otimes \nabla \rho_k)| \,\mathrm{d}x < \frac{\varepsilon}{2^{k+3}}. \tag{4.7}$$

Our candidate for the requisite smooth approximation then is given by

$$u_\varepsilon := \sum_{k=1}^{\infty} \varphi_{\varepsilon_k} * (\rho_k u).$$

Note that this is a *locally finite* sum: For each $x \in \Omega$ there exists a neighbourhood U such that only finitely many (namely, three) summands in the infinite sum defining u_ε actually contribute to $u_\varepsilon(y)$ for all $y \in U$. Since each of the summands is clearly of class C^∞, we have $u_\varepsilon \in \mathrm{C}^\infty(\Omega;\mathbb{R}^N)$.

Step 2. The L^1*-part.* We recall (P2) from above to find

$$\begin{aligned}\|u - u_\varepsilon\|_{\mathrm{L}^1(\Omega;\mathbb{R}^N)} &= \int_\Omega \left| (\sum_{k=1}^{\infty}\rho_k)u - \sum_{k=1}^{\infty}\varphi_{\varepsilon_k} * (\rho_k u)\right| \mathrm{d}x \\ &\leq \sum_{k=1}^{\infty}\int_{U_k} |\rho_k u - \varphi_{\varepsilon_k} * (\rho_k u)| \,\mathrm{d}x < \frac{\varepsilon}{4}.\end{aligned}$$

Thus $u_\varepsilon \to u$ in $\mathrm{L}^1(\Omega;\mathbb{R}^N)$ as $\varepsilon \searrow 0$.

Step 3. The total variation part. As established in step 2, $u_\varepsilon \to u$ in $\mathrm{L}^1(\Omega;\mathbb{R}^N)$ as $\varepsilon \searrow 0$ and thus, by Lemma **??**, $|Du|(\Omega) \leq \liminf_{\varepsilon\searrow 0}|Du_\varepsilon|(\Omega)$. We thus must show that $\liminf_{\varepsilon\searrow 0}|Du_\varepsilon|(\Omega) \leq |Du|(\Omega)$ to conclude the proof. To this end, we first recall the equality

$$\int_{\mathbb{R}^n} (f * g)h \,\mathrm{d}x = \int_{\mathbb{R}^n} f(g * h)\,\mathrm{d}x \tag{4.8}$$

for all f, g, h. Aiming to employ the dual characterisation (**??**) of the total variation, we let $\varphi \in \mathrm{C}_c^1(\Omega;\mathbb{R}^{N\times n})$ with $|\varphi| \leq 1$ be arbitrary. We then rewrite

$$\begin{aligned}\int_\Omega u_\varepsilon \,\mathrm{div}(\varphi)\,\mathrm{d}x &= \sum_{k=1}^{\infty}\int_\Omega (\varphi_{\varepsilon_k} * (\rho_k u))\,\mathrm{div}(\varphi)\,\mathrm{d}x \\ &= \sum_{k=1}^{\infty}\int_\Omega (\rho_k u)\,\mathrm{div}(\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x \\ &= \sum_{k=1}^{\infty}\int_\Omega u\,\mathrm{div}(\rho_k\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x - \int_\Omega (u \otimes \nabla\rho_k)(\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x) = (*).\end{aligned}$$

By (P2), $\sum_k \nabla\rho_k = \nabla\sum_k \rho_k = 0$ in Ω and thus $\sum_k u \otimes \nabla\rho_k = 0$ in Ω. Therefore,

$$\begin{aligned}(*) &= \sum_{k=1}^{\infty}\int_\Omega u\,\mathrm{div}(\rho_k\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x - \int_\Omega (u \otimes \nabla\rho_k)(\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x \\ &= \sum_{k=1}^{\infty}\int_\Omega u\,\mathrm{div}(\rho_k\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x - \int_\Omega (u \otimes \nabla\rho_k)(\varphi_{\varepsilon_k} * \varphi - \varphi)\,\mathrm{d}x = \mathbf{I} + \mathbf{II},\end{aligned}$$

with an obvious definition of $\mathbf{I}$ and $\mathbf{II}$.

The map $\rho_1(\varphi_{\varepsilon_1}*\varphi)$ is compactly supported in Ω and satisfies $|\rho_1(\varphi_{\varepsilon_1}*\varphi)| \leq 1$. Since at most three U_k's overlap each, we thus obtain

$$\begin{aligned}\mathbf{I} &\leq \int_\Omega u \operatorname{div}(\rho_1\varphi_{\varepsilon_1} * \varphi)\,\mathrm{d}x + \sum_{k=2}^{\infty}\int_\Omega u \operatorname{div}(\rho_k\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x \\ &\leq |Du|(\Omega) + \sum_{k=2}^{\infty}\int_\Omega u \operatorname{div}(\rho_k\varphi_{\varepsilon_k} * \varphi)\,\mathrm{d}x \\ &\leq |Du|(\Omega) + \frac{3}{4}\varepsilon,\end{aligned}$$

where we have used the dual characterisation of the total variation, cf. (**??**), in the second and assumption (4.5) in the third step. Ad **II**. Arguing similarly as above, cf. (4.8), we find by $|\varphi| \leq 1$

$$\begin{aligned}\mathbf{II} &= \sum_{k=1}^{\infty}\left|\int_\Omega (u\otimes\nabla\rho_k)(\varphi_{\varepsilon_k} * \varphi - \varphi)\,\mathrm{d}x\right| \\ &\leq \sum_{k=1}^{\infty}|\varphi_{\varepsilon_k} * (u\otimes\nabla\rho_k) - u\otimes\nabla\rho_k|\,\mathrm{d}x \leq \frac{\varepsilon}{8}\sum_{k=1}^{\infty}2^{-k} = \frac{\varepsilon}{8}.\end{aligned}$$

Summarising, since the estimates on $\mathbf{I}, \mathbf{II}$ do not depend on the specific choice of φ,

$$|Du_\varepsilon|(\Omega) \leq |Du|(\Omega) + \varepsilon.$$

Now send $\varepsilon \searrow 0$ to conclude. The proof is complete. □

Remark 4.10.

Example 4.11.

Example 4.12.

List of figures

References

[1] Ambrosio, L.; Fusco, N.; Pallara, D.: Functions of bounded variation and free discontinuity problems. Oxford University Press, 2000.

[2] Denk, R.: Lecture notes on functional analysis (in German), 2017.

[3] Evans, L.C.: Partial differential equations. Graduate Studies in Mathematics, Vol. 19. American Mathematical Society, 1998.

[4] Rudin, L. I.; Osher, S.; Fatemi, E. (1992). "Nonlinear total variation based noise removal algorithms". Physica D. 60 (1–4): 259–268.