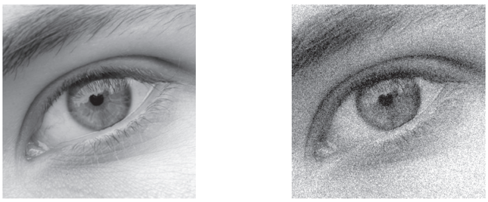
Chapter 1

Introduction

Consider manufactured parts, such as screws, car doors, lenses, or mirrors for lasers, for example. All these manufactured parts have to go through quality inspections checking if there are unwanted bumps or scratches that should not be there. There are different methods to measure the manufactured parts. One that we will consider in this thesis is a deflectometric measurement process that deals with the measurement of specular objects. The output of such measurement processes is given in some raw data depending on the process. The goal is to describe the measured object exactly by the data. This is one example for a so-called inverse problem.

Another example that we want to consider are images. These images can be photographies or MRI scans, for example. A photography can be corrupted by noise. For example these unwanted signals can occur in images where the photo was taken in a too dark environment without a flash light. That can look like the image in Figure 1.1.



(a) Noise-free image. (b) Image corrupted by noise.

Figure 1.1: Comparison of a noise-free image and an image corrupted by noise.

In general, in inverse problems an operator equation modeling a specific process is given. These processes are physical processes and applications include e.g. tomography, medical imaging, or object measurements. The aim is to find an input argument that provides the given results. To this end, an inversion of the operator equation is desired. However, these operators are usually not invertible. Additionally, inverse problems are highly sensitive to errors in the measurement data.

To overcome this issue, mathematical tools are needed to approximate an inversion of the problem. One approach is to construct functionals for which the minimization problem is well-posed in the sense that unique minimizers exist and are close to the unknown solution within a tolerance range. These functionals consist of two parts. The first part is a fidelity term which controls the deviation between the given output data and the data produced by the model for some input data. The second term is a regularization that gives the option to force the input parameters to fulfill certain properties such as how much an image resembles a “natural image”.

The application that we consider in the first part of the thesis is a data fusion process. The given dataset is a result of a deflectometric measurement process [Pet04, Pet06] and is provided by the Institute of Production Measurement Technology (IPROM) at the Technical University Braunschweig.

Deflectometric measurement processes deal with object measurements of specular objects, such as lenses or mirrors. The aim is to calculate a dataset that describes the measured object exactly. The output data consists of two sets of separately measured types of data. There are the measured surface points and the measured surface orientation given by three spatial coordinates and by normal vectors, respectively. Since the direct measurement of the points is more sensitive to noise than the measurement of the normal vectors, the accuracy of data is inconsistent. In detail, the accuracy of the normal vectors is three orders of magnitude higher than the accuracy of the surface points. We resolve this issue with a data fusion process by solving a minimization problem which uses the normal vectors as a reference value. By doing so the accuracy of the surface points is increased.

Taking the gained insights we are able to develop new theories for image denoising. Image denoising, as we realized, is a problem similar to the data fusion process.

In imaging there are different methods to denoise an image. In 1992, Rudin, Osher and Fatemi introduced the total variation as a regularizer [ROF92],

$$\min_u \frac{1}{2} \int_\Omega |u - u_0|^2 \, \mathrm{d}x + \lambda \int\limits_\Omega |\nabla u| \, \mathrm{d}x.$$

One problem in the resulting denoised images is the occurring staircasing effect, i.e. the creation of flat areas separated by jumps. One way to overcome this staircasing was proposed by Lysaker et al. in 2004 [LOT04]. The technique they proposed was a denoising of the image in two separate steps. In a first step, a total variation filter was used to smooth the normal vectors of the level sets of a given noisy image and then, as a second step, a surface was fitted to the resulting normal vectors. The method was designed in a dynamic way, i.e. by solving a certain partial differential equation to steady state.

A similar approach is taken in data fusion process. The measurement device does not only produce approximate point coordinates but also approximate surface normals. It turned out that the incorporation of the surface normals results in an effective, but fairly complicated and non-linear problem. In our approach we switch from surface normals to image gradients which leads to an effective method.

For image denoising we follow the idea of introducing additional information, e.g. gradient information, into the above ROF-model.

We formulate certain minimization problems in which use suitable reference values. In image denoising the reference value we want to use is an approximation of the image gradient vectors. Consequently, our approaches calculate such an approximation and use it as a reference value. Hence, our approaches are two-stage methods.

Another approach to prevent the staircasing effect is to go to higher orders of differentiation within the regularization term. One approach was proposed in 2010 namely the total generalized variation (TGV) functional [BKP10, KBPS11].

We propose different kinds of combinations of these functionals, since we can use the functionals as constraints or penalties. In this way we are able to formulate different minimization problems that are in some sense equivalent to the TGV problem. One advantage of some of these problems lies in the easy parameter choice rules that perform equally well as the TGV problem. Additionally, the duality gaps of these new problems are finite instead of infinite as it is usually the case in the primal-dual gap for the TGV problem. Hence, these can be used to create a reasonable stopping criterion for the optimization process. An additional advantage is the decreased runtime of the two-stage methods, since the problem is divided into two smaller problems.

1.1 Organization of the thesis

Chapter 2 provides underlying theory and notation that is used throughout this thesis. These preliminaries include notions of functional analysis such as functional spaces, e.g., spaces of bounded total variation and bounded total generalized variation, norms and seminorms, and different types of derivatives. Furthermore, the chapter contains some convex analysis, including illustrations of terms that are necessary for the solution theory of convex optimization problems. It also includes a brief overview of inverse problems in general and the direct method of the calculus of variations. The chapter concludes with solution methods for convex optimization problems that are based on [CNCP10]. Chambolle-Pock's primal-dual algorithm for solving minimization problems of the type

$$\min_{x \in X} F(x) + G(Kx)$$

for functionals F, G (e.g. representing the fidelity and the regularization term) with certain properties and a linear and continuous operator K closes this chapter.

The first of the two main parts of this thesis is Chapter 3. This chapter describes an inverse problem for an application in deflectometric measurements. The given data are sets of measured point coordinates and measured normal vectors of an object. These are provided by the same measurement setup but are determined independently. Because of the architecture of the measurement process and the different sensitivity to noise, the types of measured data do not have the same accuracy. In the chapter the geometry of the measurement setup is explained alongside the structure of the measured dataset. Moreover, different approaches to increase the accuracy to a higher order of magnitude which we call data fusion are discussed. Algorithms are proposed and tested on real datasets provided by the Institute of Production Measurement Technology (IPROM) at the Technical University Braunschweig.

The other part of the thesis, Chapters 4 to 6, uses the insights of the previous chapter and applies these to mathematical image denoising.

In Chapter 4 an idea of a two-stage image denoising method is proposed which is inspired by the fact that measured surface orientation is a powerful tool to lift the accuracy of measured surface point coordinates in the data fusion process. Consequently, a two-stage image denoising method is proposed which, in a first step, denoises the image gradients and takes these gradients into a second step as prior information where the image is denoised with respect to the solution from the gradient denoising step. Here, we propose two methods building on the same idea. Within each step the minimization functionals can be formulated via penalization or with constraints. Later on, the advantages and disadvantages of using one or the other formulation will be discussed.

Chapter 5 gives variants of total generalized variation image denoising. Taking the two-stage methods of Chapter 4, both steps can be combined into one optimization problem in various ways. The particular functionals considered in the two-stage methods can be combined into one optimization problem by pure penalization within the minimization functional, resulting in the total generalized variation problem [BKP10, KBPS11], or they can be combined in a mixed type using one or two of the particular functionals as constraints. In Chapter 5 the resulting combined methods are discussed. We investigate advantages and disadvantages of the problem formulation especially with respect to parameter choices.

In Chapter 6, the proposed methods for image denoising are experimentally tested on various images and different noise levels. Not all of the proposed combined methods come with an simple or clear parameter choice rule.The numerical experiments are restricted to those methods that do. Thus, first all methods are evaluated separately according to performance and quality and after that the methods are compared with one another.

Chapter 2

Preliminaries

This chapter introduces the mathematical background needed for this work and fixes the notation. We give a brief overview over Banach and dual spaces with examples and look into function spaces appropriate for solving minimization problems. Afterwards, we will recall convex analysis, inverse problems with variational calculus, and conclude with solving methods for convex optimization problems.

2.1 Functional analysis

In this section we will collect basic results and concepts. For more details and proofs the reader is referred to standard literature such as [Rud91, Bre].

In the following we denote by X a vector space over a field $\mathbb{K}$ ($\mathbb{R}$ or $\mathbb{C}$). A mapping $\|\cdot\|_X : X \to [0,\infty)$ is called *norm*, if the following holds:

1. $\|\lambda x\|_X = |\lambda| \, \|x\|_X \;\; \forall \lambda \in \mathbb{K}, \; \forall x \in X$,
2. $\|x+y\|_X \leq \|x\|_X + \|y\|_X \;\; \forall x, y \in X$,
3. $\|x\|_X = 0 \Rightarrow x = 0$.

If only 1. and 2. holds, the mapping is called a *seminorm*. Since we will work with norms coming from different vector spaces, lets review a few in the following example:

Example 2.1

1. Let $X = \mathbb{R}^d$. The following mappings define norms on $\mathbb{R}^d$:

$$\|x\|_p = \left(\sum_{k=1}^{d} |x_k|^p \right)^{1/p}, \quad 1 \leq p < \infty,$$
$$\|x\|_\infty = \max_{k=1,\ldots,d} |x_k| .$$

2. Let $X = \ell^p$, i.e. the space of real-valued sequences for which the norm mappings are finite:

$$\|x\|_p = \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{1/p}, \quad 1 \leq p < \infty,$$
$$\|x\|_\infty = \sup_{k=1,\ldots,\infty} |x_k| .$$

Two norms $\|\cdot\|, \|\cdot\|'$ are called equivalent on X, if there are constants $c_1, c_2 > 0$ such that

$$c_1 \|x\| \le \|x\|' \le c_2 \|x\|$$

for all $x \in X$. If X is finite-dimensional, all norms on X are equivalent.

A normed (real) vector space $(X, \|\cdot\|_X)$ is called (real) *Banach space*, if it is *complete*. That means that every Cauchy sequence $(u_n)_{n\in\mathbb{N}}$ converges, i.e. there exists an $u \in X$ such that $\lim_{n\to\infty} \|u_n - u\|_X = 0$. We will write "normed vector space X" instead of the tuple above.

Definition 2.2 (Operator, Functional)
Let X, Y be normed vector spaces. A continuous linear mapping $K : X \to Y$ is called *operator*. If $Y = \mathbb{K}$, the mapping is called *functional*.

Definition 2.3 (Space of linear mappings)
Let X and Y be normed spaces. The *space of continuous linear mappings* is

$$\mathcal{L}(X,Y) := \{K : X \to Y \mid K \text{ is a linear and continuous operator}\}.$$

The operator is bounded by

$$\|K\|_{X\to Y} := \sup_{\|x\|_X \le 1} \|Kx\|_Y < \infty. \tag{2.1}$$

$\mathcal{L}(X,Y)$ is a normed space with *operator norm* $\|\cdot\|_{X\to Y}$.

Theorem 2.4
Let Y be a Banach space. Then $\mathcal{L}(X,Y)$ is a Banach space, independently of the completeness of X.

Definition 2.5 (Lebesgue space $L^p(\Omega)$)
Let ν be the Lebesgue measure in $\mathbb{R}^d$, and $\Omega \subseteq \mathbb{R}^d$ open and nonempty. Then for $1 \le p \le \infty$

$$L^p(\Omega) := \left\{ u : \Omega \to \mathbb{C}, \nu\text{-measurable} \;\middle|\; \|u\|_{L^p(\Omega)} < \infty \right\}$$

equipped with the norm

$$\|u\|_{L^p(\Omega)} := \begin{cases} \left(\int_\Omega |u(x)|^p \, \mathrm{d}\nu(x) \right)^{1/p}, & p \in [1,\infty), \\ \operatorname*{ess\,sup}_{x\in\Omega} |u(x)| := \inf \{\alpha \ge 0 \mid \nu(\{|u| > \alpha\}) = 0\}, & p = \infty \end{cases}$$

is a Banach space. Functions $u \in L^p(\Omega)$ are called p-integrable functions and $u \in L^\infty(\Omega)$ are essentially bounded measurable functions.

Per se the Lebesgue spaces (actually written as $\mathcal{L}^p(\Omega)$) are not normed vector spaces, since it is only equipped with a seminorm $\|\cdot\|^*_{\mathcal{L}^p(\Omega)}$. But by considering $N_p = \{u \mid u = 0 \ \nu\text{-a.e.}\}$, the kernel of $\|\cdot\|^*_{\mathcal{L}^p(\Omega)}$, and identifying functions which are equal ν-a.e., i.e. considering equivalence classes $[u]$ instead of u, we obtain

$$L^p(\Omega) = \mathcal{L}^p(\Omega) \big/ N_p.$$

These are the Banach spaces in Definition 2.5. We write u instead of $[u]$, and when there are no misunderstandings within the context, we will write $\|\cdot\|_p$ instead of $\|\cdot\|_{L^p(\Omega)}$, for $1 \leq p \leq \infty$. If $K : X \to Y$ is linear, then continuity of K is equivalent to

$$\|Kx\|_Y \leq c\,\|x\|_X$$

for all $x \in X$ and a constant $c > 0$. Due to this, a linear continuous operator is also called a *bounded* linear operator. The operator norm $\|K\|_{X\to Y}$ defined in Equation (2.1) is the smallest possible constant c to satisfy this inequality.

Definition 2.6 (Dual space, duality pairing)
Let X be a Banach space. The space $X^* := \mathcal{L}(X, \mathbb{K})$ of linear continuous functionals on a normed space X is called *dual space* of X. It is equipped with the dual norm

$$\|x^*\|_{X^*} := \sup_{\|x\|_X=1} |x^*(x)| = \sup_{\|x\|_X\leq 1} |x^*(x)| = \sup_{x\in X\setminus\{0\}} \frac{|x^*(x)|}{\|x\|_X} \tag{2.2}$$

where

$$\langle x^*\,,\,x\rangle_{X^*\times X} = x^*(x). \tag{2.3}$$

The functional (2.3) is called the *duality pairing*. X^* also is a Banach space (which is a direct consequence of Theorem 2.4).

This definition immediately implies

$$\langle x^*\,,\,x\rangle_{X^*\times X} \leq \|x^*\|_{X^*}\,\|x\|_X \quad \text{for all } x \in X,\ x^* \in X^*. \tag{2.4}$$

Definition 2.7 (Adjoint Operator)
Let X and Y be Banach spaces, and $K \in \mathcal{L}(X, Y)$. The *adjoint operator* $K^* \in \mathcal{L}(Y^*, X^*)$ is defined by the relation

$$\langle K^*y\,,\,x\rangle_X = \langle y\,,\,Kx\rangle_Y \tag{2.5}$$

for all $x \in X$ and $y \in Y^*$. Further, it holds that $\|K^*\|_{Y^*\to X^*} = \|K\|_{X\to Y}$.

Example 2.8 (Dual spaces, dual pairs)

1. Let $X = \ell^p$, $1 < p < \infty$, then the dual space can be identified with $X^* \cong \ell^q$, where $\frac{1}{p} + \frac{1}{q} = 1$. The duality pairing is given by

$$\langle x^*\,,\,x\rangle_{X^*\times X} = \sum_{k=1}^{\infty} x_k^* x_k.$$

   For $p = 1$ it is $(\ell^1)^* = \ell^\infty$, but $(\ell^\infty)^*$ is not a sequence space.
2. Let $X = L^p(\Omega)$, $1 < p < \infty$, then $(L^p(\Omega))^* \cong L^q(\Omega)$ for $\frac{1}{p} + \frac{1}{q} = 1$. The duality pairing is given by

$$\langle u^*\,,\,u\rangle_{L^q(\Omega)\times L^p(\Omega)} = \int_\Omega u^*(x)u(x)\,\mathrm{d}x.$$

   Similar to the sequence space cases above, $(L^1(\Omega))^* \cong L^\infty(\Omega)$, but $(L^\infty(\Omega))^*$ cannot be identified with $L^1(\Omega)$.

Definition 2.9 (Topological terms)
Let X be a normed vector space. $U \subset X$ is called

1. *closed*, if for every convergent sequence $(x_n)_{n\in\mathbb{N}} \subset U$ it holds that $\lim_{n\to\infty} x_n = x \in U$,
2. *compact*, if every sequence $(x_n)_{n\in\mathbb{N}} \subset U$ contains a convergent subsequence $(x_{n_k})_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} x_{n_k} = x \in U$.

Further, the open ball on X is defined as $B_{X,r}(x) := \{y \in X \mid \|x - y\|_X < r\}$ and the closed ball is $\overline{B}_{X,r}(x) := \{y \in X \mid \|x - y\|_X \leq r\}$. (We will drop the "$X$" in the subscript, if the according space is clear, i.e. we write $B_r(x)$.) With this $U \subset X$ is called

3. *open*, if for all $x \in U$ there exists an $r > 0$ with $B_r(x) \subset U$. Therefore, all $x \in U$ are interior points of U and $U = U^\circ$, where U° are all interior points of U,
4. *bounded*, if U is contained in $\overline{B}_r(x)$ for an $r > 0$,
5. *convex*, if for all $x, y \in U$ and $\lambda \in [0, 1]$ it holds that $\lambda x + (1 - \lambda)y \in U$.

The definition of the norm directly gives us the property that open and closed balls are convex sets. In normed vector spaces it also holds that the complement of a closed set is open and vice versa.

Definition 2.10 (Weak(-*) convergence)
Let X be a normed vector space. A sequence $(x_n)_{n\in\mathbb{N}} \subset X$ is said to *converge weakly* to $x \in X$, if

$$\lim_{n\to\infty} \langle x^*, x_n\rangle_{X^*\times X} = \langle x^*, x\rangle_{X^*\times X}$$

for all $x^* \in X^*$; we write $x_n \rightharpoonup x$.
A sequence $(x_n^*)_{n\in\mathbb{N}} \subset X^*$ *converges weak-** to a $x^* \in X^*$, if

$$\lim_{n\to\infty} \langle x_n^*, x\rangle_{X^*\times X} = \langle x^*, x\rangle_{X^*\times X}$$

for all $x \in X$; we write $x_n^* \overset{*}{\rightharpoonup} x^*$.

Note, that the convergence $x_n \to x$ is also called *strong convergence* in X. Further, it holds that if $x_n \to x$ and $x_n^* \overset{*}{\rightharpoonup} x^*$ or $x_n \rightharpoonup x$ and $x_n^* \to x^*$, then $\langle x_n^*, x_n\rangle_{X^*\times X} \to \langle x^*, x\rangle_{X^*\times X}$. The duality pairing, however, does not converge in general, i.e. not for $x_n^* \overset{*}{\rightharpoonup} x^*$ and $x_n \rightharpoonup x$. Terms like continuity, closedness of mappings and topological terms like closedness of sets and compactness in case of strong convergence transfer directly to weak(-*) continuity, etc.

Definition 2.11 (Separable, reflexive)
A normed vector space X is called

1. *separable*, if it contains a countable dense subset,
2. *reflexive*, if the canonical linear isometric mapping $i : X \to X^{**}$, $(i(x))(x^*) = x^*(x)$ is surjective.

By the Weierstraß approximation theorem, the spaces $L^p(\Omega)$, for $1 < p < \infty$, and $C(\overline{\Omega})$ are separable. For reflexive spaces it holds that $X \cong X^{**}$, but this property is not sufficient. The spaces ℓ^p and L^p for $1 < p < \infty$ are reflexive, but ℓ^1 is not.

Theorem 2.12 (Eberlein-Šmulyan)
Let X be a normed vector space. Then X is reflexive if and only if $\overline{B_1(0)}$ is weakly compact.

Theorem 2.13 (Banach-Alaoglu)
If X is a separable normed vector space, $\overline{B}_{X^*,1}(0)$ is weakly- compact.*

Definition 2.14 (Hilbert space)
Let X be a $\mathbb{K}$-vector space. A normed vector space X is called a *pre-Hilbert space*, if there is an inner product $\langle \cdot \, , \, \cdot \rangle_X$ defined on $X \times X$ with $\langle x \, , \, x \rangle_X^{1/2} = \|x\|_X$ for all $x \in X$. If $(X, \|\cdot\|_X)$ is complete, the space is called a *Hilbert space*.

We will drop the "X" if the context is clear. Inner products satisfy the *Cauchy-Schwarz inequality*:

$$|\langle x \, , \, y \rangle_X| \leq \|x\|_X \, \|y\|_X \, . \tag{2.6}$$

Remark 2.15. For $p = 2$ the space $L^2(\Omega)$ equipped with the inner product

$$\langle u \, , \, v \rangle_{L^2(\Omega)} = \int_\Omega u(x) \, \overline{v(x)} \, \mathrm{d}\nu(x)$$

is a Hilbert space.

Theorem 2.16 (Fréchet-Riesz)
Let X be a Hilbert space and $x^* \in X^*$. Then there exists a unique $y \in X$ with $x^*(x) = x_y^*(x) = \langle x^* \, , \, x \rangle_{X^* \times X} = \langle x \, , \, y \rangle_X$ for all $x \in X$. Further, $\|x^*\|_{X^*} = \|y\|_X$.

Definition 2.17 (Directional derivative)
Let $F : X \to Y$ be an operator or functional. The *directional derivative* at $x \in X$ in direction $h \in X$ is defined as

$$D_h F(x) := \lim_{t \searrow 0} \frac{F(x + th) - F(x)}{t} \in Y,$$

if the limit exists.

Definition 2.18 (Gâteaux derivative)
Let X, Y be normed vector spaces. A mapping $F : X \to Y$ is called *Gâteaux-differentiable at* $x \in X$, if $D_h F(x)$ exists for all $h \in X$ and

$$DF(x) : X \to Y, \, h \mapsto D_h F(x)$$

is a linear and bounded operator. $DF \in \mathcal{L}(X, Y)$ is its *Gâteaux-derivative*. Further, F is called *Gâteaux-differentiable*, if it is at every $x \in X$.

Definition 2.19 (Fréchet derivative)
Let X and Y be normed vector spaces. $F : X \to Y$ is called *Fréchet-differentiable at* $x \in X$, if there is $DF \in \mathcal{L}(X, Y)$ with

$$\lim_{\|h\|_X \to 0} \frac{\|F(x + h) - F(x) - DF(x)h\|_Y}{\|h\|_X} = 0.$$

The *(Fréchet-)derivative of F (at x)* is $F'(x) = DF(x)$ and if F is Fréchet-differentiable in every $x \in X$, F is called *Fréchet-differentiable*.

Note, that if $F : X \to \mathbb{R}$ then $F'(x) \in \mathcal{L}(X, \mathbb{R}) = X^*$, hence, for functionals the derivative is an element of the dual space. Another class of function vector spaces we will need are Sobolev spaces and the associated notions.

Definition 2.20 (Spaces of test functions)
Let $\Omega \subset \mathbb{R}^d$ be non-empty and open. The space $\mathcal{D}(\Omega)$ is called *space of test functions* and is defined by

$$\mathcal{D}(\Omega) = \{f \in C^\infty(\Omega) \,|\, \mathrm{supp} f \subset \Omega \text{ compact}\}\,.$$

Definition 2.21 (Weak derivative)
Let $\Omega \subset \mathbb{R}^d$ be non-empty, open and connected, $u \in L^1_{\mathrm{loc}}(\Omega)$, and $\alpha \in \mathbb{N}^d$ a multiindex. A function $v \in L^1_{\mathrm{loc}}(\Omega)$ is called the α-th *weak derivative* of u if for every test function $\varphi \in \mathcal{D}(\Omega)$ it holds that

$$\int_\Omega v(x)\varphi(x)\,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u(x)\partial^\alpha \varphi(x)\,\mathrm{d}x.$$

If all weak derivatives with $|\alpha| \leq m$ exist, u is said to be m times weakly differentiable.

Derivatives in the classic way are also weak derivatives.

Example 2.22

1. Let $\Omega = \mathbb{R}$ and $u(x) = |x|$. If we define the sign function as

$$\mathrm{sign}(x) = \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ 1, & x > 0 \end{cases}$$

   it is clear that $v(x) = \mathrm{sign}(x)$ satisfies

$$\int_\Omega |x|\,\varphi'(x)\,\mathrm{d}x = -\int_\Omega \mathrm{sign}(x)\varphi(x)\,\mathrm{d}x$$

   with integration by parts.

2. Now, we want to differentiate the sign function, i.e. $u(x) = \mathrm{sign}(x)$. For $x \neq 0$ the derivative is 0. But the constant zero function is not a good candidate as the derivative of sign, since the fundamental theorem of calculus is not satisfied for all $x \in \mathbb{R}$. For $a, b \in \mathbb{R}$ with $a < 0$ and $b > 0$ it is

$$\int_a^b \mathrm{sign}'(x)\,\mathrm{d}x = 0 \neq 2 = \mathrm{sign}(b) - \mathrm{sign}(a).$$

   However, we can calculate the following with integration by parts and $\varphi(x) \to 0$, $|x| \to \infty$:

$$\int_\mathbb{R} v(x)\varphi(x) = -\int_\mathbb{R} \mathrm{sign}(x)\varphi'(x)\,\mathrm{d}x = 2\varphi(0) = 2\delta_0(x),$$

   where $\delta_0 = \delta$ is the *delta distribution* in 0.

Definition 2.23 (Sobolev spaces)
For $1 \leq p \leq \infty$ and $m \in \mathbb{N}$

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) \,|\, D^\alpha \in L^p(\Omega),\, 0 \leq |\alpha| \leq m\}$$

is a *Sobolev space* of order m, p. Equipped with the norm

$$\|u\|_{W^{m,p}(\Omega)} = \begin{cases} \left(\sum\limits_{|\alpha| \leq m} \|D^\alpha u\|^p_{L^p(\Omega)} \right)^{1/p}, & 1 \leq p < \infty, \\ \max_{|\alpha| \leq m} \|D^\alpha u\|_{L^\infty(\Omega)}, & p = \infty, \end{cases}$$